1

The modular proteome and its significance

Rudolf Aebersold[1, 2]

1. Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

2. University of Zurich, Zurich, Switzerland

Biological or clinical phenotypes and the cellular response to stimuli arise from the biochemical state of a cell or tissue which, in turn, is the result of the composition of biomolecules and their organization in the cell. At present, there are neither a comprehensive theory nor computational models that generally predict phenotypes or cellular responses to signals. Nevertheless, such predictions are frequently attempted, particularly in clinical research, exemplified by personalized/precision medicine. It is therefore an important question which type(s) of molecular information, either by themselves or integrated with other data, will increase the ability to predict phenotypes from molecular measurements beyond what is possible today.

The biochemical literature indicates that proteins are particularly rich in biological information. To date, most proteomic studies have focused on the identification and quantification of proteins. However, most proteins associate with other proteins and/or other types of biomolecules and carry out their function as protein modules, and a multitude of such functional modules constitutes the biochemical state of the cell. We refer to a specific instance of proteome composition and organization as the proteotype.

In this presentation we will discuss computational and mass spectrometry based experimental methods to infer or measure the modular organization of the proteotype. We will then examine to what extent the proteotype correlates with quantitative phenotypes and how the cell reorganizes the proteotype in response to genetic or external signals. Selected examples will be used to illustrate the general concepts.

2

Deadly proteomes: the central role of proteomics in dissecting the chemical arsenal of animal venoms

Glenn F King[1]

1. Institute for Molecular Bioscience, The University of Queensland, St Lucia, QLD, Australia

We are literally surrounded by venomous animals, which make up about 15% of all animal species on the planet. Venoms are unique model systems for studying protein evolution because the ongoing battle between venomous animals and their prey and predators places a constant selection pressure on venom toxins. As a result, most venoms are complex chemical cocktails, and they have proven to be an extremely valuable source of pharmacological tools, eco-friendly bioinsecticides, and human therapeutics.

Venoms vary greatly in composition because the animals that produce them have evolved to target different prey or deter different predators (or competitors in the case of platypus venom). Yet we still know very little about the composition of many animal venoms. Prior to this century, venom research focused on animals that produce large amounts of venom (e.g., snakes and large spiders) because these were amenable to the low-sensitivity analytical tools available at the time. The introduction of high-sensitivity, high-resolution MS-based proteomics in the early 2000s completely transformed the study of venoms. Tandem MS, in combination with venom-gland transcriptomics, now allows the assembly of high fidelity venom proteomes from microgram amounts of venom and nanogram amounts of venom-gland mRNA. In recent years, these technical developments have enabled the first comprehensive overview of the venom proteome of a wide variety of small venomous invertebrates, including ants, assassin bugs, centipedes, parasitoid wasps, pseudoscorpions, and robber flies. MS-based methods have also been critical for cataloguing the extraordinary diversity of post-translational modifications used to augment venom toxins. Although still in its infancy, the recent introduction of MS imaging has unveiled enormous potential to shed light on the location and mechanism of toxin production, and the functional role of toxins.

I will show seminal examples of how MS-based proteomics has enabled new insights into venom composition, venom evolution, venom production, and toxin structure and function.

3

Proteomics driven precision medicine for the early-stage hepatocellular carcinoma

Fuchu He[1], **Fuchu He**[1]

1. National Center for Protein Sciences(Beijing), State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Lifeomics, Beijing, 102206, China

Hepatocellular carcinoma is the third leading cause of deaths from cancer worldwide. Infection with the hepatitis B virus is one of the leading risk factors for developing hepatocellular carcinoma, particularly in East Asia. Although surgical treatment may be curative in the early stages, the five-year overall survival rate is only 50-70%. Advances in genomics and transcriptomics have greatly improved our understanding of the molecular mechanisms of HCC genesis and progression, HCC treatment remains a

unique clinical challenge, and targeted therapies for HCC subpopulations and corresponding predictive biomarkers are urgently required. Here, we present the largest characterization effort involving the proteomic and phosphoproteomic profiling of 110 paired clinical HBV-related early HCC tumor and non-tumor tissues (Barcelona Clinic Liver Cancer stages 0 and A). The quantitative proteomic data highlighted the heterogeneity in early-stage HCC and were used to stratify the cohort into three subtypes (i.e., S-I, S-II and S-III) with different clinical outcomes. S-III, which was characterized by disrupted cholesterol homeostasis, had the worst overall survival and greatest risk of a poor prognosis after first-line surgery. Moreover, the knockdown of the S-III specific signature SOAT1 altered the distribution of cellular cholesterol and effectively suppressed HCC proliferation and migration. Finally, on the basis of a patient-derived tumour xenograft mouse model of hepatocellular carcinoma, we found that treatment with avasimibe, an inhibitor of SOAT1, markedly reduced the size of tumours that had high levels of SOAT1 expression. Thus, our study identifies the proteomic landscape in early HCC, and the patterns of protein signatures and pathways that are altered in proteomic subtypes of HCC. The drug-targetable proteins that are identified by proteomic alterations may provide a powerful tool for identifying patients with HCC subtypes associated with a poor prognosis, and who might benefit from further targeted treatment, moving us towards the era of proteomics-driven precision medicine.

4

The elephant in the room: glycomics and glycoproteomics

Nicolle H. Packer[1, 2]

1. GLYCO @ MQ, Department of Molecular Sciences, Biomolecular Discovery & Design Research Centre, Faculty of Science & Engineering, Macquarie University, North Ryde, NSW, Australia

2. ARC Centre of Excellence for Nanoscale BioPhotonics (CNBP), Macquarie University , Sydney, NSW, Australia

Too often the glycosylation modifications to proteins have been overlooked as “being too hard” to analyse. Increasingly however, these extensive protein modifications are being found to have crucial and varied influence on, dare I say, most biological molecular interactions.

The analysis of the glycan structures, their sites on the proteins and their relative abundance is now able to be achieved using much the same tools as are used in proteomics analysis; at the single protein, protein complex and potentially at high throughput glycoprotein detailed characterisation [1,2,3,6,10]. Although improvements in informatics are still required, the current methods that we use for glycomics (the analysis of the sequence and linkage of glycan structures) and glycoproteomics (the analysis of the glycan composition, site occupation, and abundance of intact glycopeptides) are now able to address most biological research questions.

The availability of this glycoanalytical capability is in fact, increasingly being applied to a range of biological systems with the resultant increase in our understanding of biomolecular interaction mechanisms [4,8,9,12] as well as producing new biomarkers, imaging agents and drug targets. In this regard, the importance of the glycosylation of different experimental models used to investigate biomolecular questions will be emphasised. The talk will also present new data on some examples of where a knowledge of the glycan structural changes that have been observed to occur under biological perturbations such as cancer and pain can be usefully targeted by sensitive microscopic and spectroscopic detection approaches [5,7,11] at the biopsy, cell and tissue level.

1. Pia H. Jensen et al (2012) Structural Analysis of N- and O- glycans released from glycoproteins. Nature Protocols. 7 , 1299.
2. Daniel Kolarich, et al (2012) Determination of site specific glycan heterogeneity on glycoproteins. Nature Protocols. 7, 1285.
3. Benjamin Parker et al (2013) Site-specific glycan-peptide analysis for determination of N-glycoproteome heterogeneity. J. Proteome Res. 12:5791.
4. Ian Loke et al (2015) Complementary LC-MS/MS-Based N-Glycan, N-Glycopeptide, and Intact N-Glycoprotein Profiling Reveals Unconventional Asn71-Glycosylation of Human Neutrophil Cathepsin G. Biomolecules 5:1832
5. Arun Everest-Dass, et al (2016). N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues. Molecular Cell Proteomics 15:3003
6. Ling Lee et al (2016) Towards Automated N-glycopeptide identification in glycoproteomics. J. Proteome Res. 15:3904
7. Nima Sayyadi et al (2016). Sensitive Time-Gated Immunoluminescence Detection of Prostate Cancer Cells Using a TEGylated Europium Ligand. Analytical Chemistry. 88:9564
8. Phil Bokiniec et al (2017) Polysialic Acid Regulates Sympathetic Outflow By Facilitating Information Transfer Within The Nucleus Of The Solitary Tract. J.Neuroscience, 37:6558
9. Merrina Anugraham et al (2017) Tissue glycomics distinguish tumour sites in women with advanced serous adenocarcinoma. Mol Oncol. 11:1595.
10. Arun Everest-Dass et al (2018) Human disease glycomics: technology advances enabling analysis of specific glycan structures on proteins – Part 1&2. Expert Reviews in Proteomics 15:165
11. Nicole Cordina et al (2018) Reduced background autofluorescence for cell imaging using nanodiamonds and lanthanide chelates. Scientific Reports 8:4521.
12. David De Oliveira et al (2019) Human glycan expression patterns influence Group A streptococcal colonization of epithelial cells The FASEB Journal (accepted)

5

Gene editing based analysis of functional protein networks in human disease

Marius Ueffing[1], Tina Beyer[1], Katrin Junger[1], Nicola Horn[1], Yasmin Wissinger[2], Franziska Worz[1], Mohamed Ali Jarbour[1], Janina D'Alvise[1], Sylvia Bolz[1], Rob Russell[3], Ronald Roepman[4], Karsten Boldt[1]

1. University of Tübingen, Tübingen, Germany

2. Thermo Fisher Scientific, Clinical Diagnostics Systems, Hennigsdorf, Germany

3. University of Heidelberg, Germany

4. Radboud University Medical Center, The Netherlands

The enormous number of variants of unknown significance in humans as well as allelic variation of causative mutations result in a high degree of interindividual variance in clinical presentation of disease. Genotype-phenotype correlations for allelic variations are therefore hard to identify by genetics alone. This also applies to rare diseases, where a mixture of splice-site, nonsense/frameshift and missense point mutations in combination with putative modifiers define penetrance and severity of disease. To analyze interindividual variation in phenotypic severity and assign pathogenic mechanisms to different disease phenotypes we analyze the impact of allelic variance on protein interaction networks, using ciliopathies, a group of rare diseases as a paradigm. Combining CRISPR-Cas based knock-down, affinity proteomics to isolate protein interactions, quantitative mass-spectrometry and computational modeling (Boldt et al., Nature Comm. 2016) we analyze the impact of mutations on protein-protein interaction networks. This allows to identify variants that affect molecular functions relevant to disease pathogenesis, and exclude those that are not. (Beyer et al., Mol.Cell.Prot 2018). CRISPR-Cas mediated gene-editing can further be used to study the impact of a specific variance in a protein of interest under physiological and in pathophysiological conditions. This allows analysis of genetic threshold effects of specific variants like missense mutations within protein interaction networks. We can further investigate the role of protein diversity generated through splice events and gene-regulatory-activities. The versatility of the approach enables a multiplicity of possible investigations linking functional proteomic analysis to human molecular genetics as well as identify pathogenic mechanisms for individual clinical features.

6

Proteomic Tools to Decipher Mechanisms of Senescence in Aging and Age-related Diseases

Natan Basisty[1], Abhijit Kale[1], Okhee Jeon[1], Christopher Wiley[1], Su Liu[1], Chisaka Kuehnemann[1], Anja Holtz[1], Pankaj Kapahi[1], Luigi Ferrucci[2], Judith Campisi[1], Birgit Schilling[1]

1. Buck Institute for Research on Aging, Novato, CA, United States

2. National Institute on Aging, Bethesda, Maryland , United States

Cellular senescence is a striking example of a prime driver of aging phenotypes and pathologies across multiple tissues. This complex stress response causes an essentially irreversible arrest of cell proliferation and the development of a multi-component senescence-associated secretory phenotype (SASP). We hypothesize that, via the SASP, senescent cells exert cell non-autonomous effects that can disrupt cells and tissues locally and at a distance and contribute to neurodegeneration, thrombosis, and multiple age-related pathologies. Using quantitative SILAC workflows as well as more comprehensive data-independent acquisitions approaches, we have assessed the composition and functions of the SASP in aging and disease contexts. Recently, our proteomic screens have identified a novel role for senescent cells and SASP in hemostasis and blood coagulation. Since senescent cells accumulate with age, we speculate that the SASP is at least partly responsible for thrombotic events that increase with age. The general role of senescent cells as driver of age-related diseases has moved forward potential therapeutics (senolytics) to remove senescent cells to improve health span. Overall, it will be of key importance to identify senescence markers, both as biomarkers for aging and age-related diseases, in order to monitor any therapeutic interventions to eliminate senescent cells.

This work was supported by the National Institute on Aging (U01 AG060906-01, PI: Schilling; and P01AG017242 and R01AG051729, PI: Campisi) and postdoctoral fellowships by the Glenn Foundation (to NB and OJ).

10

Endurance training of human skeletal muscle results in extensive mitochondrial biogenesis and remodelling of the mitochondrial proteome

Cesare Granata[1], Nikeisha J Caruana[2], Boris Reljic[2], Javier Botella[3], Nicholas Jamnick[3], Jujiao Kuang[3], Hans Janssen[3], Adrienne Laskowski[1, 4], Tegan Stait[4], Melinda T Coughlan[1], David R Thorburn[4, 5, 6], Ann Frazier[4, 5], David J Bishop[3], David A Stroud[2]

1. Department of Diabetes, Monash University, Melbourne, Victoria, Australia

2. Department of Biochemistry and Molecular Biology, Bio21, Parkville, Victoria, Australia

3. Institute for Health and Sport, Victoria University, Melbourne, Victoria, Australia

4. Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Victoria, Australia

5. Department of Paediatrics, University of Melbourne, Melbourne, Victoria, Australia

6. Victorian Clinical Genetic Services, Royal Children's Hospital, Melbourne, Victoria, Australia

In addition to generating the bulk of cellular energy, mitochondria direct a vast array of biological functions essential for cellular homeostasis. Mutations affecting mitochondrial function and biogenesis cause mitochondrial diseases which affect tissues of high energy demand such as heart, skeletal muscle and brain. Mitochondrial dysfunction has also been implicated in various

cancers and aging. A long-standing question in biology concerns the biogenesis of mitochondria and its regulation in response to stress and the metabolic needs of the cellular environment. Exercise represents a major challenge to both these pathways while also being arguably one of the most 'natural' perturbations available. Despite this, there has been little study into exactly how mitochondria adapt to variable exercise conditions. In order to further demonstrate the effects exercise has on the mitochondria, ten participants underwent three different training volume phases over 12 weeks. Tissue biopsies were taken prior to commencing the study and after each phase and then subjected to a panel of bioenergetic assessments to measure oxidative capacity. Mitochondria were isolated from muscle biopsy material and their proteomes analysed by label-free quantitative mass-spectrometry. Training phases included a combination of normal-, high- and reduced-volume training regimens. By incorporating a combination of variable volumes of exercise we illustrated the effects these training volumes have on mitochondrial biogenesis, energetic capacity and the mitochondrial proteome more broadly. We observed extensive mitochondrial biogenesis in response to changing volumes of exercise training. While this was met with an overall increase in oxidative capacity, mitochondria underwent extensive remodelling of energetic pathways. Cessation of high-volume exercise reversed some, but not all of these changes. Our findings suggest that training volume is an important determinant of changes in mitochondrial content and function and is a useful model to help to further our understanding of the fundamental mechanisms of mitochondrial biogenesis.

11

Proteomic analysis of lysine acetylation dynamics and stoichiometry

Chuna Choudhary[1]

1. University of Copenhagen, Copenhagen, COPENHAGEN, Denmark

Lysine acetylation is a major posttranslational modification. Acetylation is implicated in diverse cellular functions, most prominently including the epigenetic regulation of gene transcription. Our laboratory is using mass spectrometry (MS)-based quantitative proteomics to map the scope of acetylation as well as to investigate its dynamic regulation in response to genetic and environmental perturbations. Furthermore, we developed novel proteomic methods to accurately quantify the site-specific stoichiometry of acetylation on a proteome-wide scale. Our results show that acetylation can occur through both enzymatic and non-enzymatic mechanisms, and enzyme-catalyzed acetylation appears to have a higher stoichiometry. I will discuss our recently published and ongoing efforts in understanding acetylation signaling.

12

Dissecting signaling pathways using PTMomics

Martin R. Larsen[1]

1. University of Southern Denmark, Odense M, FYN, Denmark

Over the years we have developed a large number of methods for enrichment of post-translational modified (PTM) peptides, from complex biological matrices, in order to make it possible to study these in nature. A PTM is the attachment of a chemical group to a protein after or during protein synthesis and the global analysis of PTMs is termed PTMomics. Many PTMs are reversible and numerous enzymes for their controlled attachment and removal exist. In fact, about 5% of proteome in an eukaryote cell are enzymes capable of add or remove PTMs. Often a complex interplay between these enzymes modulates the PTM level of a protein to fine-tune the protein activity, function and interaction, allowing a level of delicate control of signaling pathways not possible in any other way. Characterization of the PTMome requires specialized and sensitive strategies aiming not only at monitoring one single PTM but rather a repertoire of PTMs in order to obtain a larger picture of the signaling pathways. Here we will discuss strategies for large scale PTMomics and show an example of one of the workflows we have developed for assessment of phosphorylated peptides, sialylated glycopeptides, lysine acetylated peptides and free and reversibly modified cysteine-containing peptides from the same sample. We will illustrated the workflow by showing the characterization of short time T-cell signaling and ultra-short time stimulation of isolated nerve terminals.

13

Subcellular distribution of post-translational modifications in human proteome

Anna Andrejeva[1], **Tom S Smith**[1], **Rayner M L Queiroz**[1], **Eneko Villanueva**[1], **Kathryn Lilley**[1]

1. University of Cambridge, Cambridge, CAMBRIDGESHIRE, United Kingdom

Protein post-translational modifications (PTMs) regulate many processes, including subcellular localisation, protein function and protein-protein interaction. Phosphorylation on serine, threonine and tyrosine remains the most studied PTM but more than 200 PTMs have been discovered. Lysine succinylation is a relatively understudied modification despite it being widespread and expected to have large effect on protein function and activity through the change in lysine charge.[1,2] Succinylation of mitochondrial proteins may be explained by the production of succinyl-CoA as an intermediate in the mitochondrial citric acid cycle. Yet, so far there is no study addressing the true subcellular localisation of succinylated proteins and the hypothesised correlation of succinyl-CoA subcellular concentration and protein succinylation. Furthermore, the discovery of biologically relevant PTMs is greatly assisted through the establishment of their subcellular distributions, especially where the modified proteins are found at distinct locations from their unmodified counterparts.

We have pioneered the development of the novel high-resolution subcellular proteomics technique LOPIT (Localisation of Organelle Proteins by Isotope Tagging), as well as machine-learning based approaches to address protein high-throughput

subcellular classification.[3,4] LOPIT-DC utilises 10-plex TMT tags to quantify peptide profiles over cellular fractions and enables protein classification to subcellular niches.[5] In the work presented here, we combine LOPIT-DC with sequential enrichment of Lysine succinylated and phosphorylated peptides to address fundamental questions including: Which modifications co-occur with differential localisation? Where in the cell are the proteins with stable succinylated sites?

Here we present data of subcellular localisation of the phosphorylated and succinylated peptides alongside the unmodified proteome derived from the same sample. Our preliminary experiments suggest that the majority of succinylated proteins are not mitochondrial and identify succinylated proteins which are differently located relative to their unmodified proteins. This study will be critical to further our understanding of the role of these modifications and their interplay on protein localisation.

1. M.D. Hirschey and Y. Zhao, Metabolic Regulation by Lysine Malonylation, Succinylation, and Glutarylation, Molecular and Cellular Proteomics, 14 (9) 2308 – 2315, 2015
2. J. Park, Y. Chen., D. X. Tishkoff, C. Peng, M. Tan, L. Dai, Z. Xie, Y. Zhang, B. M. M. Zwaans, M. E. Skinner, D. B. Lombard and Y. Zhao, SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways, Molecular Cell, 50 (6), 919 – 930, 2013
3. A. Geladaki, N. K. Britovsek, L. M. Breckels, T. S. Smith, O. L. Vennard, C. M. Mulvey, O. M. Crook, L. Gatto and K. S. Lilley, Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics, Nature Communications, 10, Article number 331, 2019
4. L. Gatto., L. M. Breckels, T. Burger, D. J. H. Nightingale, A. J. Groen, C. Campbell, N. Nikolovski, C. M. Mulvey, A. Christofou, M. Ferro and K. S. Lilley, A foundation for reliable spatial proteomics data analysis, Molecular and Cellular Proteomics, 13 (8) 1937 – 1952, 2014
5. L. Gatto, L. M. Breckels, S. Wieczorek, T. Burger and K. S. Lilley, Mass-spectrometry-based spatial proteomics data analysis using pRolooc and pRolocdata, Bioinformatics, 30 (9) 1322 – 1324, 2014

14

Systems-wide analysis of ADP-ribosylation in human cells using quantitative mass spectrometry

Sara C. Larsen[1], Ivo A. Hendriks[1], Alexandra Stripp[1], Michael L. Nielsen[1]

1. university of Copenhagen, Copenhagen, COPENHAGEN, Denmark

ADP-ribosylation is a reversible posttranslational modification involved in a range of cellular processes. Although ADP-ribosylation was described more than 50 years ago, identification of proteins modified with ADP-ribosylation have remained notoriously difficult due to their low abundance and high structural heterogeneity. Previous approaches using chemical approaches or engineered cell culture systems either lack efficiency or sensitivity for systems-wide studies of ADP-ribosylation in complex biological samples. Alleviating these limitations we describe a proteomic strategy for comprehensive mapping of the human ADP-ribosylome, using quantitative high-resolution mass spectrometry in combination with our Af1521 enrichment approach and augmented ETD/EThcD fragmentation.

To benchmark our methodology, we characterized the ADP-ribosylation response in human cells exposed to oxidative stress yielding quantitative identification of >7.000 ADP-ribosylation sites residing on >2.200 nuclear proteins. Our data demonstrate that ADP-ribosylation catalyzed by PARP1/PARP2 is able to modify more than one-third of all nuclear proteins, confirming that ADP-ribosylation is a widespread modification with a regulatory scope comparable to other extensive posttranslational modifications. We find that serine residues are the major induced target of ADP-ribosylation during oxidative stress (>6.300), and our data provides novel insights into extensive and site-specific crosstalk between serine ADP-ribosylation and phosphorylation.

We additionally investigated the crosstalk between ADP-ribosyaltion and SUMOylation as both modifications occurs in the nucleus of cells leading to extensive chain-like modification. However, despite that SUMO and ADP-ribosylation being nuclear modifications targeting same biological processes, the extent of proteins and biological pathways regulated by their individual crosstalk remains enigmatic. Hence we applied our proteomics methodologies for systems-wide analysis, and describe a regulatory crosstalk between SUMO conjugates and ADP-ribosylation during oxidative stress.

By applying our analytical strategy for the detection and quantification of ADP-ribosylation patterns under various cellular perturbations and in tissue samples, we showcase the potential of the described methodology for both basic research and diagnostic purposes.

15

Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition (DIA) without the need for spectral libraries

Dorte B. Bekker-Jensen[1], Oliver M. Bernhardt[2], Alexander Hogrebe[1], Ana Martinez del Val[1], Lynn Verbeke[2], Tejas Gandhi[2], Christian D. Kelstrup[1], Lukas Reiter[2], Jesper Olsen[1]

1. University of Copenhagen, NNF CPR, Copenhagen N, DANMARK, Denmark

2. Biognosys, Schlieren, Switzerland

Quantitative phosphoproteomics has in recent years revolutionized understanding of cell signaling, but it remains a challenge to scale the technology for high-throughput analyses. Here we present a rapid and reproducible phosphoproteomics approach to systematically analyze hundreds of samples by fast liquid chromatography tandem mass spectrometry using data independent acquisition (DIA). To overcome the inherent issue of positional phosphopeptide isomers in DIA-based phosphoproteomics, we developed and employed an accurate site localization scoring algorithm, which is incorporated into the Spectronaut software tool.

Using a library of synthetic phosphopeptides spiked-in to a yeast phosphoproteome in different ratios we show that it is on par with the top site localization score for data-dependent acquisition (DDA) based phosphoproteomics. Single-shot DIA-based phosphoproteomics achieved an order of magnitude broader dynamic range, higher reproducibility of identification and improved sensitivity and accuracy of quantification compared to state-of-the-art DDA-based phosphoproteomics. Importantly, direct DIA without the need of spectral libraries performed almost on par with analyses using specific project-specific libraries. Moreover, we implemented and benchmarked an algorithm for globally determining phosphorylation site stoichiometry in DIA. Finally, we demonstrate the scalability of the DIA approach by systematically analyzing the effects of thirty different kinase inhibitors in context of epidermal growth factor (EGF) signaling showing that a large proportion of EGF-dependent phospho-regulation is mediated by a specific set of protein kinases.

16

Pathogenic mutations in ALS/FTD gene, *CCNF,* causes increased Lys48-ubiquitylation and defective autophagy leading to proteostasis dysfunction

Albert Lee[1], Stephanie L Rayner[1], Serene SL Gwee[1], Alana De Luca[1], Hamideh Shahheydari[1], Vinod Sundaramoorthy[1], Marco Morsch[1], Bingyang Shi[1], Shu Yang[1], Alison L Hogan[1], Emily K Don[1], Kelly L Williams[1], Justin J Yerbury[2], Lezanne Ooi[2], Angela S Lard[1], Ian P Blair[1], Julie D Atkin[1], Mark P Molloy[1, 3], Roger S Chung[1]

1. Macquarie University, North Ryde, NSW, Australia

2. University of Wollongong, Wollongong

3. Sydney University, Sydney

Neurodegenerative disorders, such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), are increasing in prevalence but lack targeted and mechanism-based therapies. Despite various biological stresses and processes that appear to contribute to the pathogenesis of these diseases, the molecular events and converging pathways remain unclear. We recently identified mutations in the CCNF gene as a novel cause of ALS/FTD, with the Ser621Gly mutation found to segregate across multiple generations in an Australian family. CCNF encodes cyclin F, an E3 ubiquitin ligase that forms a part of a SCF complex that binds to proteins for ubiquitylation and degradation by the ubiquitin-proteasome system (UPS). Experimental expression of the cyclin F^{S621G} mutation led to defective protein degradation, motor axonopathy, and features of ALS pathogenesis in vitro and in vivo. We investigated the effect of the cyclin F^{S621G} mutation on Lys48-specific ubiquitylation, and how this mutation alters its E3 ligase activity and stability that contributes to the ubiquitylation of neuronal proteins and causes proteostasis dysfunction. Additionally, we examined the phosphorylation status of cyclin F at Ser621 and how this site regulates the Lys48-specific ubiquitylation activity of the $SCF^{(Cyclin\ F)}$ complex. Proteomic analysis of immunoprecipitated Lys48-ubiquitylated proteins from mutant cyclin F^{S621G} identified proteins that clustered to the autophagy pathway, including sequestosome-1 (p62/SQSTM1), heat shock proteins (HSPs) and chaperonin complex components. Examination of autophagy markers p62, LC3 and Lamp2 in mutant cyclin F^{S621G} revealed autophagy defects specifically resulting in impairment in autophagosomal-lysosome fusion. We also identified a mechanism by which cyclin F hyperubiquitylates p62, the receptor responsible for transporting substrates for autophagic degradation. These findings demonstrate that a single missense mutation in ALS/FTD-causing cyclin F disrupts Lys48-specific ubiquitylation, leading to accumulation of substrates and defects to autophagy, that are pathological events leading up and/or contribute to ALS and FTD progression.

17

Contextualized functions of glycans in human tissue formation

Hans H Wandall[1]

1. University of Copenhagen, København N, DENMARK, Denmark

More than 200 different types of post-translational modifications (PTMs) finetune the structure and function of proteins in human tissue. The most diverse and abundant subtype of PTM is believed to be glycosylation. Glycans exhibit a large structural diversity with cell-type specificities that underlie defined biological functions. However, our functional understanding of the glycome is limited, partially due to the lack of simple model systems that allow for open-ended, unbiased screens of glycan function in human tissues. We here present the first human organotypic platform to systematically interrogate glycan functions in tissue formation. Using CRISPR-Cas9 and a 3D organotypic model of human skin, we have generated a human tissue library with truncation of the key glycan structures, thus providing a platform contextualized to a human setting with broad discovery potential. In combination with global transcriptomic, differential glycoproteomic, and differential phosphoproteomic analyses, the library demonstrates distinct phenotypes associated with loss of individual glycosylation pathways. The platform can help define the roles of the glycome in epithelial homeostasis, epithelial transformation, cell-cell and cell-matrix adhesion, signaling and host pathogen interactions, enabling glycobiology to move beyond the constrains of 2D cell culture assays, printed glycan arrays, and animal models.

18

Generation and Degradation of Free Oligosaccharides

Tadashi Suzuki[1]

1. RIKEN, Wako, SAITAMA, Japan

Glycosylation is known to be one of the most prevalent co- and post-translational modification of proteins, affecting the physicochemical/physiological properties of cognate proteins. It is also known that glycans can also occur as a “free”, unconjugated form. Recent studies have clarified that there are multiple different pathways involved in the formation of free oligosaccharides (1). For instance, intracellular free N-glycans, free oligosaccharides structurally related to asparagine-linked (N-linked) glycans, are mainly formed by hydrolysis of oligosaccharides in mammalian cells (2, 3), while they are predominatly formed by the action of the cytoplasmic peptide:N-glycanase in budding yeast, *S. cerevisiae* (4). There are also “extracellular” free oligosaccharides found in various animal sera (5). Their formation mechanism should be distinct from that for intracellular glycans. In this talk, I will summarize the current state of our knowledge concerning the formation of free oligosaccharides in mammalian cells and yeast.

References

- Harada, Y. et al., ***Cell Mol Life Sci* 72**, 2509-2533 (2015)
- Harada, Y. et al., ***Biol. Chem.* 288**, 32673-32684 (2013)
- Harada, Y. et al., ***Glycobiology* 25**, 1196-1205 (2015)
- Hirayama, H., et al., ***Biol. Chem.* 285**, 12390-12404 (2010)
- Seino, J., et al., ***Glycobiology* 26**, 1072-1085 (2016)

19

Highly efficient and precise glycoproteomic analysis

Pengyuan Yang[1], Weiqian Cao[1], Minqi Liu[1]

1. Fudan University, Shanghai, SHANGHAI, China

Protein glycosylation is a heterogeneous post-translational modification with great diversity and gives rise to functional variance. Highly precise characterization of protein glycosylation at site-specific level and the proteome scale is critical. In our recent work, we proposed a series of identification strategies and searching engines: pGlyco1.0, pGlcyo2.0, pGlyco3.0 and pGlycoQuant. pGlyco 1.0 is a pipeline for the identification of intact glycopeptides by integrating HCD, CID-MS/MS and MS3, and using a novel target-decoy method to estimate the false discovery rate of the glycan identification. pGlyco 2.0 is a one-step tandem MS strategy for intact glycopeptide identification with optimized stepped-energy fragmentation and a dedicated search engine. pGlyco 2.0 is the first search engine to conduct comprehensive quality control including false discovery rate evaluation at all three levels of matches to glycans, peptides and glycopeptides, improving the current level of accuracy of intact glycopeptide identification. With pGlyco 2.0, we reported a large-scale glycoproteome dataset consisting of 10,009 distinct site-pecific N-glycans on 1988 glycosyltaion sites from 955 glycoproteins in five mouse tissues.

We then developed pGlyco 3.0, a glycan database-free algorithm. pGlyco 3.0 is the first pipeline to achieve large-scale, universal glycopeptide analysis on diverse model organisms. We have identified a total of 44,261 distinct intact glycopeptides in seven model organisms (Budding yeast, Fission yeast, Fruit fly, Nematode, Zebrafish, Mouse and Arabidopsis). pGlycoQuant was proposed to be a robust intact glycopeptide quantification tool. It enables quantification of intact glycopeptides with chemical labeling, metabolic labeling and label free method. Accurate, comprehensive quantification was conducted using pGlycoQuant and routine proteomic methods. We achieved quantification of 9140 proteins, 5384 N-glycosylation sites and 5806 intact glycopeptide in three hepatocellular carcinoma cell lines. The series glycopeptide analysis strategies enables us to be highly efficient profiling of intact glycopeptides, and provides insight into the N-glycosylation status in biological organism.

20

Glycan analysis from tissue to serum - identification and validation of a biomarker for the early detection of hepatocellular carcinoma

Anand Mehta[1], Mengjun Wang[1], Harmin Herrera[2], Connor West[1], Alyson Black[1], Peggi Angel[1], Richard Drake[1]

1. Medical University of South Carolina, Charleston, SOUTH CAROLINA, United States

2. Microbiology and Immunology, Drexel University , Philadelphia , PA, USA

Goal: Identify a serum biomarker(s) of hepatocellular carcinoma that originates from cancer tissue.

Methods: Formalin-fixed human liver tissue from patients classified as healthy, cirrhotic, or hepatocellular carcinoma (HCC) were evaluated using N-glycan MALDI imaging mass spectrometry on a Solarix dual source 7T MALDI-FTICR (Bruker-Daltonics). A coating of PNGase F was applied to the tissue using a TMSprayer (HTX Technologies LLC). Data was analyzed using FlexImaging and SCiLS Lab software (Bruker-Daltonics). In total, 188 HCC tissue samples and 145 control tissue samples were analyzed. Subsequently, serum glycoproteomics was performed using a recombinant lectin with greater affinity toward branched and fucosylated glycan. Proteins identified as containing the same glycans observed in HCC tissue were further examined in four independent sample sets consisting of 225 patients with liver cirrhosis and 326 with HCC including 240 with early stage cancer.

Results: Through tissue-based glycan imaging, increased fucosylation was observed in 96% of HCC tissue. The glycan most often observed was a tetra-antennary glycan with one to three fucose residues. Using a recombinant AAL lectin that has increased affinity towards fucosylated and branched glycan, we identified low molecular weight kininogen (LMWK) as a serum protein that contained this glycan. Subsequently a plate based assay using high affinity antibodies, modified to lack glycosylation,

and recombinant lectins with high affinity towards the specific fucose change observed was used to test the performance of fucosylated LMWK as a biomarker of HCC. In four studies, a fucosylated LMWK based biomarker algorithm had a median AUC of 0.9575 in the detection of all HCC and 0.945 for early stage HCC.

Conclusions: Using a novel tissue-based glycan imaging platform, we were able to identify glycan changes that occur directly in cancer tissue and used this information to identify serum based biomarkers of HCC that are superior to those currently used.

21

Functional characterization of complement component C9 *C*-mannosylation in oesophageal adenocarcinoma

Julie Webster[1], Alok K Shah[1], Vojtech Franc[2], Bradley Spicer[3], Paul Morgan[4], Michelle Dunstone[4], Albert J.R Heck[2], Michelle Hill[1]

1. QIMR Berghofer Medical Research Institute, Herston, QLD, Australia

2. University of Utrecht, Netherlands

3. Monash University, Melbourne, Australia

4. Cardiff University, Cardiff, UK

Oesophageal cancer is globally ranked as the ninth highest cancer, with a 5 year survival of less than 25%. Previously we reported altered glycosylation of serum complement component C9 in oesophageal adenocarcinoma (OAC). C9 is the terminal protein of the complement pathway, which polymerizes to form the membrane attack complex (MAC) leading to cell lysis. Intriguingly, we detected C9 protein in OAC tissues, including high level deposition in high grade dysplasia. These results led us to hypothesize that differentially glycosylated C9 promotes OAC proliferation and progression, instead of cell lysis. In this study, we investigated the C9 glycosylation sites associated with OAC, and developed recombinant glycosylated C9 mutants for functional analysis on OAC cells. Glycopeptide and glycoform analysis of immunopurified C9 from patient sera showed a 30% increase in the dual occupancy of W27 and W30 C-mannosylation in OAC patients compared to healthy controls. We expressed his-tagged glycomutant C9 in HEK 293 cells (W27F and W30F), and purified by His tag. The impact of differential C-mannosylation on C9 lytic and/or proliferative function was assayed by dose titration of C9 and cell viability analysis. Endocytosis and shedding of the MAC from the plasma membrane was also examined. In summary, this project determined the regulatory role of C9 C-mannosylation in OAC development.

23

Clinical evaluation of a multiplexed protein panel to discriminate patients with psoriatic arthritis from those with rheumatoid arthritis.

Stephen R Pennington[1], Angela McArdle[1], Anna Kwasnik[1], Gerry Wilson[1], Oliver FitzGerald[1]

1. University College Dublin, Belfield, Dublin, CO DUBLIN, Ireland

Currently, the diagnosis of psoriatic arthritis (PsA)is based on the knowledge and expertise of the examining physician. One significant challenge they face is the overlapping clinical features of PsA with rheumatoid arthritis, osteoarthritis and other spondyloarthropathies. The development of a diagnostic test, which reliably distinguishes PsA from RA and other arthropathies would be of considerable clinical benefit. As a first step to the development of such a test, we evaluated the performance of a panel of candidate biomarker proteins, named PAPRICA, for its ability to discriminate between patients with PsA from those with RA. The PAPRICA assay is a multiplexed targeted proteomics (MRM) assay which includes 206 candidate biomarker proteins targeted by 423 peptides measured using reverse phase liquid chromatography coupled to triple quadrupole mass spectrometry (Agilent 1290 Infinity II chromatography system and 6495A QqQ). Peptide peaks in the assay were quantified using Skyline (v.3.7.0.11317) and data analysed by both univariate and multivariate statistical methods. Evaluation of the analytical performance revealed highly reproducible peptide separations (mean retention time %CV=0.1%)and quantification (mean peak area %CV of less than 9%) for all proteins. Application of the PAPRICA assay to PsA (n=94) and RA (n=72) samples and subsequent Random Forest analysis revealed an AUC of 0.90 and an overall accuracy of 83%. Of the 94 PsA patients, the PAPRICA assay suggested 9 may have RA and of the 72 RA patients that 19 may have PsA. Further examination of the clinical presentation of these patients will help establish the authenticity of the PAPRICA data. These findings demonstrate the ability of the PAPRICA method to discriminate patients with PsA from those with RA. In ongoing studies, the method will be validated by applying it to additional cohorts of PsA and RA patients.

24

Affinity proteomics for array based profiling of autoantibody repertoires

Peter Nilsson[1]

1. Div Affinity Proteomics, Dept Protein Science, SciLifeLab & KTH Royal Institute of Technology, Stockholm, Sverige

The quest to find novel disease associated autoantigens could and should lead to increased understanding and also enhance the long term possibilities to develop drugs for diagnosis, treatment and monitoring of progression. In order to contribute to this, we utilize in-house developed affinity proteomics resources and technologies for high-throughput and highly multiplex array-

based profiling of proteins and autoantibody repertoires in CSF and plasma within a range of inflammatory, neurodegenerative and/or psychiatric disorders as well as within healthy individuals.

The general understanding of the global reactivity patterns in the human autoantibody repertoires is still at an early phase. More and more diseases and conditions are speculated to have autoimmune components but very few novel targets are clearly associated to disease conditions.

We are utilizing a unique resource of affinity reagents created within the Human Protein Atlas (HPA, www.proteinatlas.org). HPA has generated more than 42 000 unique protein fragments of in average 80 amino acids and the atlas currently contains protein expression and localization data for 26 000 antibodies targeting 17 000 human proteins.

Through a combination of various planar and bead-based microarray formats, including an array with 42 000 protein fragments representing 19 000 unique proteins, assays are set up for untargeted screening of autoantibody repertoires and targeted protein profiling and verification assays. The latter part utilizes a suspension bead array format with magnetic color-coded beads functionalized with either antibodies to generate protein profiles from labeled samples or antigens to capture autoantibodies. This format enables the analysis of up to 384 samples in parallel on 384 antigens or antibodies.

Within a large effort of neuroproteomic profiling in large numbers of both plasma and CSF samples are we comparing and exploring the global autoantibody repertoires as well as protein profiles of brain associated proteins. The focus is on the analysis of samples from patients with frontotemporal dementia, amyotrophic lateral sclerosis, multiple sclerosis, Alzheimer's disease and psychiatric disorders.

We see in general a large degree of heterogeneity between individuals and within diseases and also often relatively high numbers of antigens targeted by each individuals repertoires of IgGs, which is also the case for healthy individuals. There is certainly a need for careful and extensive characterization of identified autoantigens in order to understand the disease associated mechanisms.

1. http://www.publicationslist.org/nipe

29

Exploring farm animal proteomes, and their relevance to human health.

Emøke Bendixen[1]

1. Aarhus University, Aarhus, DENMARK, Denmark

The OneHealth perspective reminds us that human life and health depends on the health and biology of countless animals, plants and microbes. This has become particularly urgent and clear with the rapid spread of microbial resistance to antibiotics (AMR), which currently poses a major threat to global human health. AMR is a direct consequence of a massive use of antibiotics for both humans and farm animals. With 70 % of the total global antibiotics consumption being used for farm animals, solving the antibiotics crisis will depend on solving the current crisis in farm animal health. Moreover, 60 % of all human pathogens originate from animals, and while some are highly publicized such as SARS and H5N1, the area of zoonosis is largely neglected. Cross-species studies and comparative biology will be essential to deliver the best possible measures to control zoonotic infectious diseases. Fundamental proteome research of farm animal species is in particular essential for characterizing specific host-pathogen crosstalk at the molecular level, and for monitoring the health state of farm animals, and how their health affects human health.

This talk will give a status on progress in farm animal proteome research and present our ongoing studies of host pathogen interactions in farm animals, including studies of how gene variants can support pigs resistance to *E. coli*and sheeps resistance to helminth pathogens. Our cross species studies provide important opportunities to study comparative host pathogen interactions in both animals and humans, and to characterize specific virulence factors that makes *E.coli*a major pathogen challenge for both human and farm animal health.

30

Proteomics applied to beef productions: from the discovery of biomarkers to the development of evaluation tools

Brigitte BP Picard[2, 1], **Christine CC Carapito**[3], **Muriel MB Bonnet**[2, 1]

1. Université Clermont Auvergne, INRA, VetAgro Sup, UMR Herbivores, Saint-Genès-Champanelle, 63122 , France

2. INRA, Theix, SAINT-GENèS-CHAMPANELLE, France

3. Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, 67000, France

Research of biomarkers of beef qualities, tenderness and intramuscular fat content (IMF), follows the workflow of human disease biomarkers studies aimed to be clinically used for diagnosis (Rifai et al., 2006). The first step was the discovery of candidate biomarkers by comparing extreme groups of tenderness and IMF. Two-dimensional electrophoresis associated with mass spectrometry was first used (Picard and Gagaoua 2018), and further combined with a "label-free" quantitative shotgun (Bazile et al., 2019). Samples of *semimembranosu*s muscle (SM) were collected on 89 cows of the French Maine Anjou Protected Denomination of Origin. Shotgun analysis revealed 875 proteins with a unique ID identified by at least 2 peptides among the 89 SM muscles. Comparative analysis of groups of high *versus* low tenderness and high *versus* low IMF (n=5 cows for each group), revealed respectively 53 and 77 proteins differentially abundant between groups. A second step consisted in predicting the tenderness and IMF on the whole population of the 89 cows. A logistic regression allowed to perfectly distinguish (100% well classified) the SM samples of high, medium or low tenderness with 7 proteins, and of high or low IMF with 5 proteins. Bioinformatic analysis with the ProteINSIDE tool (Kaspric et al., 2015) allows highlighting the corresponding biological functions. The next step is to select the method to be used in diagnostic tools to accurately qualify and quantify these biomarkers. Several MS-based

methods allowing absolute quantification were compared including selected reaction monitoring (SRM), parallel reaction monitoring (PRM) and sequential windowed acquisition of all theoretical spectral SWATH-MS methods (Bons et al., 2018). PRM has qualified 6 proteins over 10 candidate biomarkers as differentially abundant between groups of muscles divergent by the tenderness or IMF values. A semi-quantitative immunological assay (Reverse Phase Protein Array) is currently developed to accurately and quickly assay these muscular proteins.

1. Bazille J., Picard B., Chambon C., Valais A., Bonnet M. 2019. Pathways and biomarkers of marbling and carcass fat deposition in bovine revealed by a combination of gel-based and gel-free proteomic analyses. Meat Sci, in press
2. Bons J., Husson G., Delalande F., Cianférani S., Picard B., Bonnet M., Carapito C. 2018. Targeted proteomics method comparison: SRM, PRM and SWATH-MS to quantify proteins in bovine muscle tissues. In the procedding of XIIth Eupa Congress, 16-20 june 2018, Santiago de Compostela, (SP), in press.
3. Kaspric N., Picard B., Reichstadt M., Tournayre J., Bonnet M. 2015. ProteINSIDE to easily investigate proteomics data from ruminants: Application to mine proteome of adipose and muscle tissues in bovine foetuses. Plos One, 10 (5), e0128086.
4. Picard B, Gagaoua M. Chapter 11 - Proteomic investigations of beef tenderness. In: Colgrave M, editor. Proteomics in Food Science: From Farm to Fork: Academic Publisher; 2017. p. 177-97.
5. Rifai N., Gillette M.A., Carr S.A. 2006. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol. 24(8):971-83.

35

High throughput Single Cell Chemical Characterization of the Cells in the Brain

Jonathan V Sweedler[1]

1. University of Illinois at Urbana Champaign, Urbana, IL, United States

In the postgenomic era, one expects the suite of chemical players in a brain region to be known and their functions uncovered. Perhaps surprisingly, many brain chemicals remain poorly characterized and for those that are known, their localization, dynamics and function are oftentimes unknown. Several approaches for assaying the chemical content within targeted brain regions and from individual brain cells are highlighted, including mass spectrometry imaging (MSI) and single cell measurements. Using these approaches, we can measure lipids, fatty acids, neurotransmitters, neuropeptides, and proteins. For single cell measurements, the cells of interest are scattered across a microscope slide, the exact cell positions determined via optical microscopy, and mass spectra are acquired only at the cell positions. The single cell assays allow differences in the metabolome and peptidome from supposedly homogeneous populations of cells to be explored. By obtaining information from tens of thousands of individual cells, rare cells are found and unusual neurochemicals are discovered. For select cells, follow-up capillary electrophoresis-mass spectrometry and other information rich assays can be performed. Several applications of single cell mass spectrometry are highlighted from the discovery of unusual metabolites to characterizing the both known and previously unknown neuropeptides and hormones in single cells. Our overarching goal is to uncover the complex chemical mosaic of the brain and pinpoint key cellular players involved in a range of physiological and pathological processes.

36

Single cell proteome variability

Emma Lundberg[1]

1. Science for Life Laboratory (KTH), Solna, STOCKHOLMS LAN, Sweden

Compartmentalization of biological reactions in time and space is an important mechanism to allow multiple cellular reactions to occur in parallel. Thus, resolving the spatiotemporal cellular distribution of the human proteome would greatly increase our understanding of human biology and disease. We have previously generated a high-resolution map of the subcellular distribution of the human proteome as part of the open access Human Protein Atlas database. We have shown that as much as half of all proteins localize to multiple compartments. Such proteins may have context specific functions and 'moonlight' in different parts of the cell, thus increasing the functionality of the proteome and the complexity of the cell from a systems perspective. Recently we turned to single cell analysis to identify proteins with potential temporal variability in expression. We identify 17% of the human proteome to display cell-to-cell variability, of which we could attribute 25% as correlated to cell cycle progression and present the first evidence of cell cycle association for 258 proteins. A key finding is that the variance, of many of the cell cycle associated proteins, is only partially explained by the cell cycle, which hints at extensive cross-talk between the cell cycle and other signaling pathways. Single cell sequencing data further demonstrates that only 18 % of these proteins are temporally regulated at the transcript level, indicating that the majority of the novel cell cycle proteins are subjected to translational or post-translational temporal regulation.

In summary, I will demonstrate the importance of spatial proteomics data for improved single cell biology.

41

The Human Protein Atlas - Implications for Human Biology, Drug Development and Precision Medicine

Mathias Uhlen[1]

1. KTH Royal Institute of Technology, Stockholm, NA, Sweden

The Human Protein Atlas (HPA) is an international program with the aim to map of all the human proteins in cells, tissues and organs using integration of various omics technologies, including genomics, transcriptomics, antibody-based imaging, mass spectrometry-based proteomics and systems biology. The current version (www.proteinatlas.org) consists of three separate parts, each focusing on a particular aspect of the genome-wide analysis of the human proteins; (1) the Tissue Atlas showing the distribution of the proteins across all major tissues and organs in the human body, (2) the Cell Atlas showing the subcellular localization of proteins in single cells, and (3) Pathology Atlas showing the impact of protein levels for survival of patients with cancer. This year we plan to also launch a new Brain and a new Blood Atlas. All the data in the knowledge resource is open access to allow scientists both in academia and industry to freely access the data for exploration of the human proteome. We have used this resource to launch various efforts in the field of Precision Medicine. We have also launched a Human Secretome Project to produce in mammalian cell cultures all human secreted proteins (4). The progress of this project will be discussed and the results from phenotypic screening of this resource will be presented.

1. Uhlen et al (2015) *Science* 347: 1260419
2. Thul et al (2017) *Science*356 (6340): eaal3321
3. Uhlen et al (2017) *Science*357 (6352): eaan2507
4. Uhlen et al (2018) BioRxiv, https://doi.org/10.1101/465815

42

Studies of innate immune signaling regulation through quantification of proteoforms and modeling of the TLR pathway

Aleksandra Nita-Lazar[1]

1. National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MARYLAND, United States

Toll-like receptor (TLR) signaling in macrophages is essential for generating effective innate immune responses. Quantitative differences dependent on the dose and timing of the stimulus critically affect cell function and have been found to involve proteins that are not components of widely shared transduction pathways. Mathematical modeling is an important approach to better understand how these signaling networks function in time and space.

We have successfully modeled the S1P signaling pathway in macrophages using selected reaction monitoring (SRM) to measure the absolute abundance of the pathway proteins and were able to use the resulting values as parameters in a computational pathway model. RNA-seq was performed to identify expressed transcripts. Shotgun mass spectrometry was used to identify proteotypic peptides. Now, to model the TLR signaling networks SRM assays for the canonical TLR signaling pathway and related proteins and phosphoproteins have been developed. SRM with heavy-labeled internal peptide standards was used to quantify protein and phosphorylated protein molecule numbers per cell in both untreated and LPS-stimulated macrophages. These absolute protein abundance values were entered into a model of the TLR pathway that has been developed using Simmune, the rule-based modeling tool with a visual interface. To reach beyond basal level quantification to further develop and test the TLR signaling network model we use global proteomic approaches to discover biologically important proteins, protein complexes and PTMs involved in this pathway. The protein and PTM levels are quantified in macrophages under diverse, but well-defined conditions. Our data will allow us to parameterize and test the TLR network model under a variety of conditions. Tohe model will improve our understanding of the regulation of the immune signaling pathways activated during an infection, and enable immune system modulation for an appropriate immune response. This research was supported by the Intramural Research Program of NIAID, NIH.

43

Value and limitations of LC-MS HLA-ligandome data for antigen discovery and vaccine development

Nicola Ternette[1]

1. University of Oxford, Oxford, UK, United Kingdom

Value and limitations of LC-MS HLA-ligandome data for antigen discovery and vaccine development

48

The *in silico* human surfaceome & technologies for the elucidation of the surfaceome nanoscale organization

Bernd Wollscheid[1]

1. ETH Zurich, Zurich, ZH, Switzerland

Cell surface proteins are of great biomedical importance as demonstrated by the fact that 66% of approved human drugs listed in the DrugBank database target a cell surface protein. Despite this biomedical relevance, there has been no comprehensive assessment of the human surfaceome, and only a fraction of the predicted 5,000 human transmembrane proteins have been shown to be located at the plasma membrane. To enable analysis of the human surfaceome, we developed the surfaceome predictor SURFY, based on machine learning. As a training set, we used experimentally verified high-confidence cell surface proteins from the Cell Surface Protein Atlas (CSPA) and trained a random forest classifier on 131 features per protein and specifically, per topological domain. SURFY was used to predict a human surfaceome of 2,886 proteins with an accuracy of

93.5%, which shows excellent overlap with known cell surface protein classes (i.e., receptors). In deposited mRNA data, we found that between 543 to 1,100 surfaceome genes were expressed in cancer cell lines and maximally 1,700 surfaceome genes were expressed in embryonic stem cells and derivative lines. Thus, the surfaceome diversity depends on cell type and appears to be more dynamic than the non-surface proteome. To make the predicted surfaceome readily accessible to the research community, we provide visualization tools for intuitive interrogation (wlab.ethz.ch/surfaceome). The *in silico* surfaceome enables the filtering of data generated by multi-omics screens and supports the elucidation of the surfaceome nanoscale organization using proximity-based tagging strategies. Proximity Radical Tagging (PRT) and LUX-MS technology are two new chemical proteomics-based strategies which enable the identification of ligand-receptor interactions, but also the elucidation of lateral/cis interactions within the surfaceome. Proximity-tagging by LUX-MS and PRT provide new molecular spatial relationship information which could be exploited for drug targeting.

54

Mass spectrometry imaging applications for neurosurgery and neurooncology

Nathalie Agar[1]

1. Harvard Medical School, Brigham and Womens' Hospital, Boston, MA, United States

Mass spectrometry provides multiple options for the direct characterization of tissue to support surgical decision-making, and provides significant insight in the development of drugs targeting tumors of the central nervous system (CNS). Using an array of mass spectrometry (MS) applications, we rapidly analyze specific tumor markers ranging from small metabolites to proteins from surgical tissue for rapid diagnosis and surgical guidance. With similar clinical protocols and in pre-clinical animal studies, we visualize drug and metabolites penetration in brain tumor tissue and correlate with tumor heterogeneity and response to support drug development.

60

Targeting respiratory viruses using structure-guided inhibitor design on glycoenzymes

Mark von Itzstein[1]

1. Griffith University, Southport, QLD, Australia

Viruses have a long history of causing significant disease throughout the centuries. Immense impact on world socioeconomics has occurred as a consequence of virus-caused diseases, for example Hong Kong's borders were closed as a result of the appearance of the SARS coronavirus and shut down trade and all travel within this region.

Viruses transmitted in aerosolised particles from human to human by coughing and sneezing, such as influenza virus, remain of utmost concern given the potential of rapid pandemic waves due to international travel.

Influenza virus continues to cause both epidemics and pandemics. Successful inhibition of the viral sialidase (neuraminidase, NA) hinders the release of new virus progeny from an infected host cell and significantly reduces further virus spread. We have recently described the discovery of highly potent sialosyl sulfonate inhibitors of influenza virus sialidase.[1,2] One of the designer sialosyl a-sulfonate derivatives is a nanomolar inhibitor[2] in a cell-based influenza virus replication assay and has comparable activity to that of anti-influenza drugs zanamivir and oseltamivir carboxylate. Furthermore, we have undertaken a protein X-ray crystallographic study that provides atomic-level detail of the binding mode of these sialosyl a-sulfonate derivatives.[2]

1. Á. Hadházi, M. Pascolutti, B. Bailly, J. C. Dyason, A. Borbás, R. J. Thomson, M. von Itzstein, *Biomol. Chem.* 2017, *15*, 5249-5253.
2. Á. Hadházi, L. Li, B. Bailly, A. Maggioni, G. Martin, L. Dirr, J. C. Dyason, R. J. Thomson, G. F. Gao, A. Borbás, T. Ve, M. Pascolutti, M. von Itzstein, *ChemMedChem*. 2018, *13*, 785-789.

61

Systematically decoding glycosylation in disease.

Lara K. Mahal[1]

1. University of Alberta, Edmonton, ALBERTA, Canada

Glycans both encode cellular information, such as cell-cell interactions and cell state (Signal), and must avoid being targeted by pathogens (Noise). This leads to a system in which the sugar code (i.e. the glycan motifs controlling function) is hidden within the noisy milleu of larger heterogenous glycan structures. This talk focuses on use of high-throughput analytical methods, including our lectin microarray technology and newly developed miRNA-proxy approach, in tandem with data integration, to decode structure-function relationships in the glycome. Our work is identifying glycan drivers of disease by focusing on clinical samples and relevant systems. We are breaking new ground in areas including melanoma, pancreatic cancer, and host-response to pathogens (e.g. influenza), providing new targets for small molecule intervention in these disease states.

66

Chamber and cell type specific views of the human heart glycoproteome and glycome in health and disease

Rebekah L Gundry[1]

1. Medical College of Wisconsin, Milwaukee, WI, United States

Cell surface glycoproteins and glycans play critical roles in a range of biological functions and disease processes, from maintaining cellular structure and adhesion to controlling how cells send and receive exogenous signals in a complex environment. In the heart, for example, proper expression and glycosylation of ion channels are essential for propagating action potentials and proper contraction of the myocardium. Despite their critical roles in cardiac development, disease, and drug uptake, we do not yet have a detailed cell type- or chamber-resolved view of the cell surface glycoproteome or glycome of the adult human heart. Combining advanced strategies to isolate individual cardiac cell types with the recently developed CellSurfer Platform, which integrates a microscale Cell Surface Capture method for the identification of cell surface glycoproteins from small sample sizes, automated data processing workflows, and SurfaceGenie for cell-type specific marker prioritization, new views of the human heart are emerging. To date, we have identified >650 cell surface *N*-glycoproteins on primary human cardiac myocytes and fibroblasts, including proteins not previously described in these cell types and putative cell-type and chamber-specific markers. Additionally, a structure-based glycomics approach reveals >110 glycan structures in cardiac myocytes, which complement and inform glycoproteomic efforts. Overall, these data represent the first major step towards a comprehensive, cell-type, subtype, and chamber-resolved reference map of cell surface glycoproteins and glycans in the adult human heart and reveal new potential targets for immunophenotyping, *in vivo* imaging, drug delivery, and benchmarking cardiomyocytes derived from human pluripotent stem cells. Moreover, these data inform caveats regarding the use of explanted cardiac fibroblast models, reveal new molecular targets to study in the context of cardiac fibrosis and heart failure, and will promote studies aimed at gaining a better understanding of cross-talk among cardiac cell types in health and disease.

67

Delving Deeper into the Cardiac Proteome – Analyzing the Heart "*Modificome*"

Desmond Li[1, 2], Lauren E Smith[1], Alexander W Rookyard[3, 2], Stuart J Cordwell[3, 1, 2, 4], Melanie Y White[1, 2]

1. Discipline of Pathology , The University of Sydney, NSW, Australia

2. The Charles Perkins Centre, The University of Sydney, NSW, Australia

3. School of Molecular Bioscience, The University of Sydney, Sydney, NSW, Australia

4. Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia

The power of proteomics lies in the simultaneous and unbiased analysis of multiple protein species. The heart, which is comprised largely of cardiomyocytes, has a broad dynamic range, with a relatively small number of highly abundant sarcolemma / contractile and mitochondrial proteins accounting for ~50% of total protein. This can be further complicated by the role of the heart in distribution of blood, which is found in the chambers and coronary circulation, and contributes a similarly and well known dynamic range effect. Thus many studies assume the heart to be relatively low complexity in terms of unique protein species, because standard proteomics identifies only the most abundant proteins. We have utilized two approaches to limit these dynamic range effects and thus identify the low abundance myocardial proteome. We employ *ex vivo* perfusion to facilitate blood wash out and to provide a controlled environment to study, in contracting animal model hearts, several diseases (e.g. ischemia / reperfusion [I/R] injury). Secondly, we utilize several enrichment techniques for specific post-translational modifications (PTMs). We describe how this approach has been used to investigate how type II diabetes mellitus (T2DM) alters the cardiac proteome. Clinically, diabetic patients are a 2-4x greater risk of I/R leading to acute myocardial infarction (AMI), and furthermore surviving T2DM patients also recover from AMI more poorly than their otherwise healthy counterparts. This suggests that T2DM hearts undergo molecular and cellular adaptations even prior to I/R that makes them more susceptible to damage. Langendorff perfused hearts from T2DM and control animals are subjected to an integrated 'omics approach including proteomics, metabolomics, lipidomics and PTM analysis. Using strategies for the enrichment of modified peptides including phosphorylation and redox PTMs we have mapped >7500 cardiac proteins, in comparison to ~3500 proteins using traditional proteomics alone. Changes at the proteomics and 'modificomics' levels were validated via metabolomics, lipidomics and *ex vivo* functional assays. In the setting of T2DM, we have observed changes in basal signaling and redox protection that likely impact the ability of these hearts to recover post-I/R compared to otherwise healthy models. The ability to see beyond changes in proteins regulating metabolic processes may provide a key to understanding the poor recovery of T2DM patients following AMI.

72

The Central Role Of Milk Proteomics In One Health Approach

Paola Roncada[1]

1. University Magna Græcia of Catanzaro, Germaneto Catanzaro, CATANZARO, Italy

In last two decades, great efforts have been addressed to increase the study of milk proteomics (especially in human and ruminants) which remains a bioactive biological fluid of great interest. Because of the complexity and multiplicity of milk components, different research techniques have been combined to explore genetic aspects, molecular pathways, microbiota and cellular functions involved in milk production, quality, and safety to gain a multifaceted picture addressing this complexity. Milk is a heterogeneous body fluid occurring in nature to address the nutritional and defensive needs of the mammal's newborns. It could be considered one of the major feeding resources for humans if considering all the milk products like cheese, fermentation and transformation products. In contrast to human milk, that is a nutrient only in the early life, animal milk and dairy products are nutrients for the entire life of humans.
It has a great biological and nutritional value , milk is the center of the dairy industry where it is mainly transformed into cheese and a diverse array of other milk-by products. Interest in understanding the origin, composition and role of the milk of different

animal origin is steadily increasing in the last years, along with the consciousness of the role played by the microbiota in the development and diversification of the myriad of dairy products and functional foods. Besides reducing the pathogenic population, milk treatments exert a profound inhibition of the indigenous non-pathogenic microflora. The large-scale study of proteins (proteomics) can provide information about different protein profiles, characteristics of milk, nutrients, lactation stage, and the relative health status of the animal. More extensively, proteomics is useful for the assessment of safety and quality of both milk and dairy products. Moreover, ruminants' milk is considered a diagnostic fluid, a biosensor, and metaproteomics investigation of raw milk could represent an index of antibiotic resistance of the environment, giving information about contamination and epidemiology of AMR.

80

Perspectives on International Consortium on Proteogenomic: Interactions between Funders and Investigators

Henry Rodriguez[1]

1. National Cancer Institute, National Institutes of Health, Bethesda, MARYLAND, United States

-

81

HubMap: Partnership between the Funders and Investigators

Michael Snyder[1]

1. Stanford University, Stanford, CA, United States

-

82

Impact of the glycoproteome on B cell responses: From checkpoint inhibitors to HIV vaccines.

James C Paulson[1]

1. The Scripps Research Institute, La Jolla, CA, United States

The mammalian immune system has evolved the ability use glycan binding proteins to regulate immune cell responses, and to use glycans in generating an immune response to pathogens. Receptors that recognize glycans help immune cells to distinguish between self and non-self and help identify and generate immune responses against pathogens that carry 'non-self' glycans. For pathogens that decorate themselves with 'self-like' glycans to shield immune attack, such as HIV, the immune system can use glycans as part of the immune epitope or avoid the glycan shield to develop pathogen specific immune responses. While current tools to study the roles of glycans in immuno-biology are limiting, only mass spectrometry has the sensitivity needed to systematically extract information about the glycome proteo-glycome in biological samples. In our own work, we have turned to the power of mass spectrometry-based glycan profiling and glyco-proteomics whenever possible. Examples include identification the glycoprotein ligands of the B cell inhibitory protein CD22 using *in situ* glycan-protein crosslinking, analysis of the impact of glycosyltransferase inhibitors on the cellular glycome, and analysis of site-specific glycosylation of HIV envelope glycoprotein, the target for all known broadly neutralizing antibodies of HIV. As will be seen from these examples, the existing tools for studying the glycome are powerful, yet they lag far behind those available for studying the genome and proteome. In the foreseeable future, mass spectrometry-based methods will undoubtably provide the basis for transformative technologies needed to elucidate the mammalian glycome and its roles in biology (Supported by NIH research grants: AI113867; AI100663; AI050143).

83

Taiwan Cancer Proteogenomics Moonshot: Pathway to Next Generation Precision Medicine in Cancer

Yi-Ju Chen[1], Theodoros I Roumeliotis[2], Ya-Hsuan Chang[3], Ching-Tai Chen[4], Chia-Li Han[5], Miao-Hsia Lin[1], Ze-Shiang Lin[6], Wei-Hung Chang[7], Gee-Chen Chang[8, 9], Kuen-Tyng Lin[1], Yet-Ran Chen[7], Ana I Robles[10], Henry Rodriguez[10], Jin-Shing Chen[11], Sung-Liang Yu[6], Ting-Yi Sung[4], Jyoti S Choudhary[2], Hsuan-Yu Chen[3], Pan-Chyr Yang[12], Yu-Ju Chen[1]

1. Institute of Chemistry, Academia Sinica, Nankang, Taipei, TAIWAN, Taiwan

2. The Institute of Cancer Research, London, UK

3. Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

4. Institute of Information Science, Academia Sinica, Taipei, Taiwan

5. Master Program in Clinical Pharmacogenomics and Pharmacoproteomics, College of Pharmacy, Taipei Medical University, Taipei, Taiwan

6. School of Medical Technology, National Taiwan University, Taipei, Taiwan

7. Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

8. Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

9. Comprehensive Cancer Center, Taichung Veterans General Hospital, Taichung, Taiwan

10. Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

11. Division of Thoracic Surgery, Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan

12. Center of Genomic Medicine, National Taiwan University, Taipei, Taiwan

Combing proteomics with the long standing success of genomics, the new initiative of large-scale tissue proteogenomics demonstrated discovery to reveal new molecular subtypes of cancer in recent years. The different genetic background and environmental factors contribute to unique features of cancers in diversity of population, awaiting full delineation of genomic-to-proteomic network to identify the fundamental drivers and underlying mechanism. With the aim of accelerating the progress toward prevention, control and treatment for cancer, Taiwan joined the global effort of International Cancer Proteogenome Consortium (ICPC) to apply proteogenomics as a precision approach to delineate the connection of genomic abnormalities and protein alteration in individual cancer patient, which subsequently stimulates academia-government-industry collaboration to jointly map the pathway for next generation precision medicine for Taiwan/Asia cohort.

On the pilot study of early stage and never smoking lung adenocarcinoma (LUAD), the genomic landscape confirmed the distinct mutational profile of our cohort compared to previously reported studies. We observed gender-specific differences in driver and passenger mutations, which likely contribute to the disease heterogeneity and different clinical outcomes. We discovered mutation signatures predominantly associated with carcinogen as well as early onset females. Proteome subtypes highlight molecular differences that extend the classification beyond the level of clinical staging and genomic driver mutation, which signatures may provide clues on patient outcome and progression. Proteomics landscape also revealed the stage-specific progression signatures characterized by dramatic molecular reorganization at early stage to regulate cancer cell survival, migration and proliferation. Further validations by retrospective cohort not only supported the new molecular staging but also nominated biomarker candidates associated with the poor overall survival in patients. In summary, the integrated proteogenomics profile may provide molecular map for next generation precision medicine to address the unmet clinical need of NSCLC in Taiwan/Asia.

84

Characterization and turnover of RNA-binding proteins: novel insights into ribosome maintenance

Jakob Trendel[1, 2], Matilde Bertolini[1, 3], Guenter Kramer[1, 3], Bernd Bukau[1, 3], Jeroen Krijgsveld[1, 4]

1. German Cancer Research Center (DKFZ), Heidelberg, Germany

2. European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

3. Center for Molecular Biology of Heidelberg University (ZMBH), Heidelberg, Germany

4. Heidelberg University, Medical Faculty, Heidelberg, Germany

To understand the functional interplay of protein and RNA, we have recently developed a methodology to globally identify proteins that interact with RNA, independent of the RNA-biotype. The method, termed XRNAX, purifies protein-RNA complexes, serving as a starting point for detailed exploration of the proteins and RNA they contain. Specifically, we have used this to globally characterize RNA-interactomes in various cell lines, to determine dynamic changes in interaction networks upon cellular stress, and to identify protein-RNA interaction sites. In addition, we used pulsed-SILAC labelling to investigate if binding to RNA influences protein stability. Indeed, we found that in MCF7 cells the half-life of RNA-bound proteins was on average 1.6 fold longer than the same proteins in the overall proteome, and up to 5-fold for individual cases, suggesting that protein stabilization can be a general function of RNA when associating with protein. Likewise, ribosomal proteins are overall stabilized on RNA, however half-lives within the ribosome span more than one order of magnitude, suggesting protein-specific exchange from the complex. To investigate the dynamics of ribosome maintenance and turnover in more detail, we combined pulsed-SILAC labelling with sucrose gradient fractionation to determine half-lives of ribosomal proteins in the 40S, 60S, 80S and polysome fraction, demonstrating profound differences in protein stability between these subunits, especially when translation is inhibited. Collectively these data demonstrate ribosome maintenance by exchange of individual proteins from the ensemble, thereby adding a novel dimension to the classical model describing ribosome production at a defined stoichiometry and destruction by subunit.

85

Sugar Code Cracking of Serum Haptoglobin for the Diagnosis of Gastric Cancer by Targeted Glycoproteomic Approach

Hyun Joo An[1, 2]

1. Chungnam National University, Daejeon, South Korea

2. Asia Glyomics Reference Site, Daejeon, South Korea

Gastric cancer (GC) is one of the leading causes of cancer-related death worldwide, largely because of difficulties in early diagnosis. Despite accumulating evidence indicating that aberrant glycosylation is associated with GC, the development of diagnostic platform to capture changes in glycosylation to increase specificity and sensitivity for clinical use is still an analytical challenge. Haptoglobin is the major serum components and a positive acute-phase protein with immunomodulatory properties. Here, we created an analytical platform with a targeted glycoproteomic approach for GC biomarker discovery. Glycosylation alteration was monitored by intact analysis of Hp using liquid chromatography–mass spectrometry followed by immunoaffinity purification with the serum samples. Unlike conventional glycomic approach with untargeted mass spectrometric profiling of released glycan, our platform based on direct analysis of intact glycoprotein has merits such as simple, easy handling of sample preparation and time saving for analysis. Age- and sex-matched 200 serum samples (100 cancer patients and 100 healthy controls) were used to explore the clinically altered glycosylation for gastric cancer diagnosis. This study suggested that glycosylation variation of serum haptoglobin were associated with patients with gastric cancer and might be a promising marker

for the cancer screening. Our platform, which provides biological information as well as high sensitivity and reproducibility, may be useful for GC biomarker discovery.

90

Protein TAILS Tell Remarkable Tales: Positional Proteomics Reveals Diverse N-Terminomes and Proteolytic Landscapes in Disease

Christopher M Overall[1]

1. University of British Columbia, Vancouver, BC, Canada

Identification of protein terminal peptides provides key information on protein stability and function. Our degradomics methods enrich and annotate terminomes—Terminal Amine Isotopic Labelling of Substrates (TAILS, *Kleifeld etal 2010)*, Amino Terminal Orientated Mass Spectrometry (ATOMS), Proteome-wideIdentification of Protease Cleavage Specificity (PICS, *Schilling_&_Overall 2008*), LysargiNase, the new protease for proteomics we discovered (Huesgen *etal* 2015), and our N and C-termini database TopFIND (http://clipserve.clip.ubc.ca/topfind) our studies reveal widespread truncation and generation of termini in normal and diseased tissues.

Certain N terminal semi-tryptic peptides exhibit beneficial m/z, ionization and fragmentation properties over tryptic peptides, rendering these peptides and proteins identifiable. In the C-HPP, TAILS was used to identify proteoforms and also provide MS evidence for the expression of PE2-4 'missing proteins', in rare tissues and cells.

Using PICS we discovered active protease domains in bacterial flagella of >200 species.~1,000 cleavage sites for "flagellinolysin" being identified. With ~20,000 flagellin copies/~10-µm flagella this assembles the largest proteolytic complex known with potential for numerous roles in saprophytic bacteria and in pathogens.

ATOMS was used to identify function-modifying cleavages by metalloproteinases in moonlighting extracellular tRNA synthetases. When moonlighting outside the cell, "intracellular" tyrosyl tRNA synthetase was proteolytically activated as a proinflammatory mediator signaling through Toll-like receptor-2 (TLR2), resulting in NF-κB activation and TNα/chemokine release from macrophages. In tryptophan tRNA synthetase, cleavage inactivated TLR signalling and proinflammatory functions.

By TAILS we find that the innate immune cell MMPs orchestrate leukocyte chemotaxis by cleavage of most human chemokines. For CCL chemokines, inactivation/generation of antagonists are common sequalae, whereas CXCL activation by the neutrophil specific MMP8 is followed by inactivating cleavages within the ELR motif by MMP12 regulates neutrophil chemoattraction in vivo. More recently IFNα and IFNγ, and complement proteins were discovered as new MMP substrates. MMP processing of these bioactive substrates dampens inflammation essential for terminating inflammatory responses.

1. 1. Schilling, O. and Overall, C.M. 2008. Proteome-derived Database Searchable Peptide Libraries for Identifying Protease Cleavage Sites. Nature Biotechnology 26, 685-694.
2. 2. Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Starr, A., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. 2010. Isotopic labeling of Terminal Amines in Complex Samples Identifies Protein N-termini and Protease Cleavage Products. Nature Biotechnology 28, 281-288.
3. 3. Huesgen, P.F., Lange, P.F., Rogers, L.D., Solis, N., Eckhardt, U., Kleifeld, O., Goulas, T., Gomis-Rüth, F.X., and Overall, C.M. 2015. LysargiNase Mirrors Trypsin for Protein C-Termini and Methylation Sites Identification. Nature Methods 12, 55-58.
4. 4. Quancard, J., Klein, T., Fung, S-Y., Renatus, M., Hughes, N., Israël, L., Priatel, J.J., Kang, S., Blank, M.A., Viner, R.I., Blank, J., Schlapbach, A., Erbel, P., Kizhakkedathu, J., Villard, F., Hersperger, R., Turvey, S.E., Eder, J., Bornancin, F., and Overall, C.M. 2019. An Allosteric MALT1 Inhibitor is a Molecular Corrector Rescuing Function in an Immunodeficient Patient. Nature Chemical Biology 15, 304 – 313.
5. 5. Doucet, A. and Overall, C.M. 2011. Broad Coverage Identification of Multiple Proteolytic Cleavage Sites in Complex High Molecular Weight Proteins Using ATOMS Quantitative Proteomics as a Complement to Edman Sequencing. Molecular & Cellular Proteomics 10, M110.003533 1-12

91

Proteomics of Extracellular Matrix Remodeling Following Myocardial Infarction

Merry L Lindsey[1]

1. University of Nebraska Medical Center, Omaha, NE, United States

Following myocardial infarction (MI), the left ventricle (LV) undergoes a series of cardiac wound healing responses that involve stimulation of robust inflammation to clear necrotic myocytes and tissue debris and induction of extracellular matrix (ECM) protein synthesis to generate a scar. Proteomic strategies provide us with a means to index the ECM proteins expressed in the LV, quantify levels, determine molecular and cellular physiology, and explore interactions. This talk will focus on the major cell types that coordinate cardiac wound healing, namely the infiltrating leukocytes and the cardiac fibroblasts. We will discuss efforts in proteomics research that have expanded our understanding of post-MI LV remodeling, concentrating on the strengths and limitations of different proteomic approaches to glean information that is specific to ECM turnover in the post-MI setting. We will discuss how recent advances in sample preparation and labeling protocols increase our successes at identifying components of the cardiac ECM proteome. We will summarize how proteomic approaches, focusing on the ECM compartment, have progressed over time to current gel-free methods using decellularized fractions or labeling strategies. In summary, this talk will provide an overview of how cardiac ECM proteomics has evolved over the last decade and will provide insight into future directions that will drive our understanding of cardiac ECM turnover in the post-MI LV.

96

neXtProt: a SPARQLing light in the dark human proteome

Lydie Lane[2, 1]

1. Department of microbiology and molecular medicine, Faculty of medicine, University of Geneva, Geneva, Switzerland

2. SIB Swiss Institute of Bioinformatics, Geneva, GENEVA, Switzerland

neXtProt (www.nextprot.org) provides a large, coherent, up-to-date human protein dataset, unmatched in terms of scope and quality[1]. This dataset combines manually curated data from the literature and high throughput genomic, transcriptomic and proteomic data from selected resources using a single inter-operable format and community-approved standards. Full traceability is ensured by extensive use of metadata. By providing a SPARQL-based advanced query tool with extensive documentation and examples, neXtProt users can easily carry out complex searches on this data corpus and perform federated queries encompassing data found elsewhere (doi.org/10.7490/f1000research.1116829.1).

neXtProt serves as reference knowledgebase for the HUPO Human Proteome Project (HPP) since 2013 and provides specific tools to design and analyze proteomics experiments. It integrates mass spectrometry data from PeptideAtlas and displays it in the Proteomics and Peptide views. Based on this and other information, neXtProt establishes the yearly reference set of "missing proteins", predicted from genome analysis but never experimentally confirmed[2]. Identification of those missing proteins is one of the main goals of the C-HPP; it is challenging and often requires targeted experiments. The Peptide Uniqueness Checker[3] and the new Protein Digestion tool can be combined to plan such experiments by determining which unique peptides can theoretically be obtained by digestion of a target protein with a given protease. Finally, nextProt provides sequences, PTM and variation data in the PEFF format[4], allowing to take into account a wealth of proteoforms when analyzing MS spectra.

neXtProt also provides the reference set of "proteins with unknown function", either predicted or experimental[5]. Their characterization is the goal of a new initiative from the C-HPP[6]. Information about protein-protein interactions, subcellular location or expression encapsulated in neXtProt can help to build functional hypotheses that can be experimentally validated[5].

neXtProt is constantly optimizing its tools to serve the proteomics community. Feedback is welcome.

97

Computational proteomics enhancements in MaxQuant by (deep) machine learning and ion mobility awareness.

Juergen Cox[1]

1. Max Planck Institute of Biochemistry, ., Germany

Machine learning has shown its great potential of improving identification and quantification in proteomics in many respects. We report on advances in the MaxQuant software for proteomics data analysis by the integration of conventional and deep learning algorithms. The prediction of MS/MS fragmentation spectra has reached accuracies that are only limited by the technical reproducibility of spectrum acquisition. The utility of these prediction models in conjunction with retention time prediction is illustrated by applying them to the analysis of both data-dependent and data-independent acquisition datasets. In the former case, we observe a q-value-dependent increase in the total number of peptide identifications. In the latter case, we confirm that the use of predicted tandem mass spectrometry spectra is nearly equivalent to the use of spectra from experimental libraries. We also describe the ion mobility aware MaxQuant software, which utilizes the data dimension added by ion mobility to LC-MS/MS data. A new matching between runs (MBR) algorithm that utilizes collisional cross section (CCS) values of MS1 features in the matching process significantly gains specificity from the extra dimension. Prerequisite for using CCS values in MBR is a relative alignment of the ion mobility values between the runs. The missing value problem in protein quantification over many samples is greatly reduced by CCS aware MBR.

98

Intact N-glycopeptide database search using GPSeeker

Zhixin TIAN[1]

1. Tongji University, Shanghai, SHANGHAI, China

N-Glycosylation, being one of the most common and complex protein post-translational modifications (PTMs), is known to have microheterogeneity with the presence of different N-glycan structures at a single specific glycosite. These different structures may have exactly the same monosaccharide composition but totally different differential expressions and pathological relevance. Mass spectrometry-based N-glycoproteomics has so far been successful in large-scale characterization of these N-glycans at the composition level, and structure-level identification and quantitation is urgently needed. Here we report our development of the intact N-glycopeptide search engine GPSeeker and the GPSeeker-centered quantitative structural N-glycoproteomics pipeline. In analysis of differential N-glycosylaiton in MCF-7 cancer stem cells (relative to MCF-7 cells) with RPLC-MS/MS (HCD) analysis, 2,558 intact N-glycopeptides were identified corresponding to 727 N-glycosites, 724 unique peptides, 136 putative N-glycan linkages from 53 monosaccharide compositions and 640 intact N-glycoproteins; with isotopic diethyl labeling, 144 differentially expressed intact N-glycopeptides with >=1.5-fold change ($p<0.05$) were quantified. For the 20 known CD series CSC markers, 13 (CD13, CD14, CD20, CD4, CD90, CD63, CD49F, CD151, CD97, CD44, CD49b, CD109 and CD133) were either quantified, identified, or observed as GPSMs.

1. 1. Yue Wang, Feifei Xu, Kaijie Xiao, Yun Chen* and Zhixin Tian*. Chemical Communications, 2019, DOI: 10.1039/C9CC04114A. 2. Kaijie Xiao and Zhixin Tian*. Journal of Proteome Research, 2019, 10.1021/acs.jproteome.9b00191.

100

ReScoring peptide-to-spectrum-matches based on predicted fragment ion intensities leads to an increased identification rate in metaproteomics

Tim Van Den Bossche[2, 1], Ana S. C. Silva[2, 1], Bernhard Y. Renard[3], Thilo Muth[3], Lennart Martens[2, 1]

1. CompOmics, VIB - UGent Center for Medical Biotechnology, Ghent, Belgium

2. Dept. of Biomolecular Medicine, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium

3. MF1 - Bioinformatics Unit, Robert Koch Institute, Berlin, Germany

The field of metaproteomics, the study of the collective proteome of whole (microbial) ecosystems, has seen substantial growth over the past few years. This growth comes from an increased awareness that metagenomics and metatranscriptomics can be powerfully supplemented by analysis of the proteins that can be found in the community, as clearly illustrated by e.g. the Integrative Human Microbiome Project. Despite its high relevance, the field still suffers from low identification rates in comparison to single-species proteomics. The underlying challenge here, is a lack of sequence resolution and statistical validation in the current identification algorithms, which are typically designed for single-species proteomics (Colaert et al. 2011, Muth et al. 2015).

To solve this issue, we applied the recently developed, machine learning-based ReScore algorithm on several multi-species, metaproteomics datasets (Silva et al., 2019). ReScore is a post-processing tool that re-evaluates peptide-to-spectrum-matches (PSMs) based on predicted fragment ion peak intensities. To achieve this, ReScore combines two, well-established machine learning-based algorithms: Percolator, which re-scores PSMs based on the search engine output (Käll et al., 2007), and MS2PIP, which predicts fragment ion peak intensities given a peptide's sequence, charge and modifications (Degroeve et al., 2013). In the ReScore algorithm, the search engine-dependent features of Percolator are replaced with intensity features of MS2PIP. When ReScore is applied on metaproteomics datasets, it performs similar to Percolator. However, when both feature sets from Percolator and MS2PIP are combined, a significant improvement can be achieved.

When the updated ReScore algorithm is applied on metaproteomics datasets, our results show that ReScore leads to an increased identification rate, ranging from the number of PSMs to the taxonomical level, while the false discovery rate (FDR) remains under full control as validated in an entrapment experiment with *Pyrococcus furiosus* (Vaudel et al., 2012).

101

Reactome Pathway Analysis and Visualization

Chuqiao Gong[1], Guilherme Viteri[1], Henning Hermjakob[1], Johannes Griss[1]

1. EMBL-EBI, Cambridge, CAMBRIDGESHIRE, United Kingdom

Reactome (https://reactome.org) is a free, open-source, open-data, curated and peer-reviewed knowledge base of biomolecular pathways, currently covering 10,833 protein coding genes supported by 30,027 literature references. Pathways are arranged in a hierarchical structure, allowing the user to navigate from high level concepts like immune system to detailed pathway diagrams showing biomolecular events like membrane transport or phosphorylation. For the higher levels of the hierarchy, Reactome now provides scalable, interactive textbook-style diagrams in SVG format, which are also freely downloadable and editable. Repeated diagram elements like 'mitochondrion' or 'receptor' are freely available as a library of graphic elements at https://reactome.org/icon-lib. Detailed lower-level diagrams are downloadable in editable PPTX format as sets of interconnected objects, as well as in standard png format. Pathway analysis capabilities have been extended to include quantitative GSEA analysis, an R interface, and a new, visually attractive genome-wide results overview based on Voronoi maps.

102

Drilling into the N-glycomes of parasites and their vectors

Katharina Paschinger[1], Francesca Martini[2], Barbara Eckmair[1], Shi Yan[3], Iain B. H. Wilson[1]

1. Department für Chemie, Universität für Bodenkultur, Wien, Austria

2. Institut für Parasitologie, ETH, Zürich, Switzerland

3. Institut für Parasitologie, Veterinärmedizinische Universität, Wien, Austria

Despite years of research, the glycomes of invertebrates continue to surprise. Gone are the days when it could be said that "simple" organisms have "simple" glycomes. Typically, any biological sample will yield a complex mixture of 100 N-glycans or more, whereby low-abundance structures are still underestimated. On the other hand, as there are variations in structures or abundance of glycoconjugates between species, it appears that speciation correlates with glycomic alterations and thereby with the evolution of special ecological niches, such as being parasites which infect hosts and are transmitted through other invertebrates. This also means that the glycome of each parasitic and host species may act as a "passport" to allow infection or transmission, which leads to the question of which glycans can bind which receptors. In our recent studies on parasites and vectors, as well as comparisons to related species, we see potential for glycomimickry as being a factor. For instance, the canine heartworm *Dirofilaria immitis* is transmitted by the mosquito *Aedes aegypti*; both display glucuronic acid modifications of their N-glycans. On the other hand, the heartworm also expresses phosphorylcholine-modified structures which, as shown by glycan array experiments, can interact with mammalian C-reactive protein. Overall, the heartworm N-glycome contrasts with those reported to date for other parasitic nematodes. For instance, neither *Trichuris suis*, *Haemonchus contortus* nor *Oesophagostomum dentatum* have glucuronylated N-glycans, which are all gut parasites, but the latter two have trifucosylated core chitobiose modifications found neither in *Trichuris suis* or *Dirofilaria immitis.* This leads us to wonder about the functional repercussions of these variations and whether parasite tropism (in terms of vector, host and tissue) and mode of

immunomodulation is affected by glycan structures. Thereby, a mix of approaches including in depth glycomics, glycan arrays, glycoenzymology and chemoenzymatic synthesis are key to informing parasitological, vaccinological and immunological studies.

103

Comprehensive characterization of protein glycosylation in *Leishmania spp.*

Giuseppe Palmisano[1]

1. University of Sao Paulo, Sao Paulo, SAO PAULO, Brazil

The interactions of parasites with their different hosts are critical for the completion of each life stage and carbohydrates are important molecules in this process. The roles of protein glycosylation in host-parasite interactions, such as adhesion, invasion, survival and immune system recognition, determine the diversity of outcome of infection. The surface of trypanosomatids is covered by a dense array of heavily glycosylated glycoproteins and glycolipids attached to the membrane via glycosylphosphatidylinositol anchors. These molecules are important in the activation of the innate immune system and initiation of the acquired host immune response in the vertebrate host. Moreover, the remodeling of the parasite surface coat is one key aspect underlying differentiation processes. The study of protein glycosylation in trypanosomatids has unprecedented promise for the discovery of vaccine candidates and for the development of novel chemotherapy approaches and diagnostic tools. However, analytical tools for in depth characterization of site-specific glycan composition and structure linked to proteins still face challenges due to the extensive glycan micro- and macro-heterogeneity. This lecture aims to provide a brief history of the main findings of glycan structures and enzymes involved in *Leishmania spp*. N- and O-linked protein glycosylation pathways as well as to discuss how analytical methodologies based on mass spectrometry (MS) can be useful as a tool for large-scale characterizing the set of glycans and glycoproteins in parasites. The glycoproteome of three Leishmania species, L. L. amazonensis, L. V. braziliensis and L. L. infantum chagasi, indicated a specie-specific glycosylation which was associated to phylogenetic traits. These methodologies offer novel candidates for biomarker discovery and therapeutics in parasitic diseases.

104

Application of glycoproteomics to identify the plant Golgi localized UDP-GlcNAc transporter and mining the data for novel plant *O*-linked glycans

Joshua L Heazlewood[1]

1. The University of Melbourne, Melbourne, VIC, Australia

Glycosylation require activated glycosyl donors in the form of nucleotide sugars to drive processes such as post-translational modifications and polysaccharide biosynthesis. Many of these reactions occur in the endomembrane using cytosolic-derived nucleotide sugars. These are actively transported into the lumen by nucleotide sugar transporters (NSTs). We recently identified a plant UDP-GlcNAc transporter responsible for the delivery of substrate for the maturation of *N*-glycans and sphingolipids within the endomembrane. To determine the biochemical phenotype of the UDP-GlcNAc transporter loss-of-function mutants, we have applied glycoproteomic profiling. This necessitated the development of a HILIC enrichment and mass spectrometry-based workflow to detect, identify and quantify *N*-glycopeptides. Results indicated that that *N*-glycopeptides containing complex *N*-glycans (e.g. GlcNAc) were significantly reduced in plants lacking the UDP-GlcNAc transporter. In contrast, complex *N*-glycan structures from wild-type plants comprised a considerable proportion (35 %) of observed *N*-glycans i.e. those containing GlcNAc. Our findings indicate that the reference plant Arabidopsis contains a single UDP-GlcNAc transporter responsible for the maturation of complex *N*-glycans in Golgi lumen. The resultant N-glycoproteomic data was further mined to determine whether we could find any evidence for *O*-linked glycans in pant proteins. Since plants encode many proteins with epidermal growth factor (EGF) domains, we hypothesized that, similar to mammals, small *O*-linked glycans may also exist in plants. We have identified a number of high-confidence matches that would indicate that plant proteins may indeed harbour small *O*-linked glycan post translation modifications.

105

Glycomic approach for detection of *Enterococcus* infection in chronic pancreatic diseases

Saki Itoyama[1], Tomohiro Maekawa[1], Risako Fukaya[1], Emika Noda[1], Tomoya Fukuoka[1], Shinji Takamatsu[1], Yoshihiro Kamada[1], Eiji Miyoshi[1]

1. Osaka University Graduate School of Medicine, Suita, OSAKA, Japan

(Background and Aim) Bacterial infection underlies the pathogenesis of many human diseases, including acute and chronic inflammation. Recently, we found that subclinical pancreatitis was observed in pancreatic tissue surrounding pancreatic cancer. Herein, we investigated a possible role for bacterial infection in the progression of chronic pancreatitis. (Materials and Methods) Pancreatic juice was obtained from patients with pancreatic cancer (n=20) or duodenal cancer/bile duct cancer (n=16) and subjected to PCR using universal primers for the bacterial 16S ribosomal RNA gene. Bacterial species were identified by PCR using bile samples from four pancreatic cancer patients. Immunohistochemical and serologic analyses for *Enterococcus faecalis* infection （antibody against *Enterococcus faecalis* capsular polysaccharide, anti-CPS antibody） were performed on a large cohort of healthy volunteers and patients with chronic pancreatitis or pancreatic cancer and on mice with caerulein-induced chronic pancreatitis. The effect of *E. faecalis* antigens on cytokine secretion by pancreatic cancer cells was also investigated. (Results) We found that 29 of 36 pancreatic juice samples were positive for bacterial DNA. *Enterococcus* and *Enterobacter* species were detected primarily in bile, which is thought to be a pathway for bacterial

infection of the pancreas. *Enterococcus faecalis* was also detected in pancreatic tissue from chronic pancreatitis and pancreatic cancer patients; antibodies to *E. faecalis* capsular polysaccharide were elevated in serum from chronic pancreatitis patients. *Enterococcus*-specific antibodies and pancreatic tissue–associated *E. faecalis* were detected in mice with caerulein-induced chronic pancreatitis. Addition of *Enterococcus* lipoteichoic acid and heat-killed bacteria induced expression of pro-fibrotic cytokines by pancreatic cancer cells in vitro. (Conclusion) Infection with *E. faecalis* may be involved in chronic pancreatitis progression, ultimately leading to development of pancreatic cancer.

106

Analysis of *In vivo* Arginine-glycosylation targets of the NleB/SseK family of effectors reveals discrete effector substrate specificities

Nichollas Scott[1], Joshua Newson[1], Cristina Giogha[2], Jaclyn Pearson[2], Elizabeth Hartland[2]

1. Department of Microbiology and Immunology, University of Melbourne, Melbourne, VICTORIA, Australia

2. Centre for Innate Immunity and Infectious Diseases, Hudson Institute of Medical Research, Clayton, Australia

Effector glycosyltransferases are a growing class of secreted bacterial proteins which subvert normal cellular functions through the modification of host proteins with carbohydrates. Within the enteric pathogens (enteropathogenic *Escherichia coli (EPEC), Salmonella enterica* serovar Typhimurium and *Citrobacter rodentium)* the NleB/SseK family of effector glycosyltransferases mediate the glycosylation of arginine residues. The archetype of this family, NleB1 of EPEC, has been shown to modify multiple death-domain-containing proteins responsible for extrinsic apoptosis, with *N*-acetylglucosamine under ectopic expression/*in vitro* conditions. Using an Arginine-GlcNAcylation affinity proteomic approach we recently showed that under in vivo condition NleB1 preferentially modifies a single protein, Fas-associated death domain protein (FADD), during infection (1). Building on this work we have now investigated the specificity and substrate repertoire of two homologues of NleB1, SseK1 and SseK3, from *Salmonella enterica* serovar Typhimurium during bacterial infections. We demonstrate that in contrast to recent ectopic expression studies during *S.* Typhimurium infection the SseK1 and SseK3 effectors target discrete death domain proteins in the TNF and TRAIL signaling pathways (2). We reconciled the difference between our finding and previous studies by demonstrating that the over expression of the NleB/SseK family of effectors leads to ramped modification of both host and bacterial substrates. Taken together, these findings demonstrate that the NleB/SseK effectors can act far more promiscuously than previously thought when overexpressed, but under endogenous infections each effector antagonise different components of death receptor signalling.

1. Scott, N. E., Giogha, C., Pollock, G. L., Kennedy, C. L., Webb, A. I., Williamson, N. A., Pearson, J. S., and Hartland, E. L. (2017) The bacterial arginine glycosyltransferase effector NleB preferentially modifies Fas-associated death domain protein (FADD). The Journal of biological chemistry 292, 17337-17350
2. Newson, J. P., Scott, N. E., Yeuk Wah Chung, I., Wong Fok Lung, T., Giogha, C., Gan, J., Wang, N., Strugnell, R., Brown, N. F., Cygler, M., Pearson, J. S., and Hartland, E. L. (2019) Salmonella effectors SseK1 and SseK3 target death domain proteins in the TNF and TRAIL signaling pathways. Molecular & cellular proteomics : MCP

107

Systematic chemical synthesis and immunological function of *Campylobacter jejuni* lipid As

Atsushi Shimoyama[1], Sho Nakagawa[1], Yoshifumi Matsuura[1], Koichi Fukase[1]

1. Osaka University, Toyonaka, OSAKA, Japan

A Gram-negative bacteria *Campylobacter jejuni* produces lipooligosaccharide (LOS) as an outer membrane component. *C. jejuni* LOS is composed of core oligosaccharide part and glycolipid part termed “lipid A”. Since *C. jejuni* oligosaccharides have molecular homology with gangliosides in human nervous tissues, *C. jejuni* infection facilitates the induction of auto-antibodies to gangliosides, leading autoimmune diseases[1)]. Although immunostimulation by lipid A proved to be critical to promote antibody production, the role of the *C. jejuni* lipid A in auto-antibody production has not been elucidated because of the synthetic difficulty in *C. jejuni* lipid A.

General lipid As are composed of only glucosamine (GlcN). *C. jejuni* lipid A, however, has structural diversity in the sugar framework consisting in GlcN and 2,3-diaminoglucose (GlcN3N)[2)]. We thus established a diversity oriented synthetic strategy via the key disaccharide intermediates having GlcN-GlcN, GlcN-GlcN3N, GlcN3N-GlcN and GlcN3N-GlcN3N framework respectively, and accomplished the first chemical synthesis of four types *C. jejuni* lipid As.

We then measured the NF-κB activation of *C. jejuni* lipid As in HEK-Blue™ TLR4 cells and found that the slight difference in the sugar frameworks greatly affects the immunological activity of lipid A and GlcN-GlcN3N framework is most important in immunoactivation. The GlcN-GlcN type and GlcN-GlcN3N type lipid As showed agonistic activity (GlcN-GlcN3N type was stronger than GlcN-GlcN type), whereas GlcN3N-GlcN type and GlcN3N-GlcN3N type showed no activity.

On the other hand, *Helicobacter pylori* having the same molecular homology as *C. jejuni* does not induce immunological cross-reaction[3)]. We have synthesized *H. pylori* lipid A and found that it acts as an antagonist against TLR4-MD2 while inducing chronic inflammatory signals[4)]. These results suggest that the immunological cross-reaction is regulated by the differences in lipid A function.

1. N. Yuki, Lancet Infectious Diseases, 2001, 1, 29.
2. A. P. Moran et al., Eur. J. Biochem., 1991, 198, 459.
3. D. A. Sack et al., J. Clin. Microbiol., 1998, 36, 2043.
4. A. Shimoyama et al., Chem Eur J, 2011, 17, 14464.

108

Systems-level analysis of immune development early in life

Petter Brodin[1, 2], Christian Pou[1], Axel Olin[1], Ewa Henckel[3, 2], Yang Chen[1], Tadepally Lakshmikanth[1], Jaromir Mikes[1], Anna Gustafsson[3, 2], Jun Wang[1], Anna Karin Bernhardsson[1], Cheng Zhang[1], Kajsa Bohlin[3, 2]

1. Science for Life Laboratory, Karolinska Institute, Solna, SOLNA, Sweden

2. Department of Newborn Medicine, Karolinska University Hospital, SE-17176, SWEDEN

3. Department of Clinical Science, Intervention and Technology, Karolinska Institutet, SE-14152, SWEDEN

Epidemiological data suggests that early life exposures are key determinants of immune mediated disease later in life. Young children are also susceptible to infections, warranting more analyses of immune system development early in life. Such analyses have mostly been performed in mouse models or human cord blood samples, but these cannot account for the complex environmental exposures influencing human newborns after birth. We have performed a systems-level analysis of newborn immune system development and uncovered drastic developmental changes, triggered by environmental exposures, and following a shared stereotypic pattern. Here I will describe our latest results in our pursuit of understanding how early-life environmental exposures shape human immune systems, how tolerance to colonizing microbes is established and the functional capacity and infectious disease susceptibility is determined in human newborns.

109

Temporal profiles of plasma proteome during childhood development and natural progression of Type 1 Diabetes

Qibin Zhang[1]

1. University of North Carolina at Greensboro, Kannapolis, NC, United States

Human blood plasma proteome reflects physiological changes associated with a child's development as well as natural progression of childhood diseases, such as Type 1 Diabetes (T1D). While age-specific normative values are available for proteins routinely measured in clinical practice, there is paucity of comprehensive longitudinal data regarding changes in human plasma proteome during childhood. We applied TMT-10plex isobaric labeling-based quantitative proteomics to longitudinally profile the plasma proteome in 10 healthy children and 11 T1D patients during their development, each with 9 serial time points from 6 months to 15 years of age. In total, over 2000 protein groups were identified at peptide and protein level false discovery rate of 1% and with at least two razor and unique peptides. The longitudinal expression profiles of these proteins were statistically modeled and their temporal changes were categorized. The patterns and relative abundance of proteins obtained by LC-MS were also verified with ELISA. Oxidative stress related proteins have consistently different dysregulated patterns in T1D group than in age-sex matched healthy controls, even prior to appearance of islet autoantibodies – the earliest sign of islet autoimmunity and pancreatic beta cell stress. Our work represents one of the most comprehensive longitudinal profilings of pediatric plasma proteome to date. The temporal profiles of plasma proteome obtained in this study provide a detailed look of the systemic changes of plasma proteome during the natural progression of T1D and also offer a comprehensive resource and reference for biomarker studies in other childhood diseases. Our results also suggest using strictly age-matched clinical samples in a cross-sectional study in pediatric population.

110

Orthogonal validation of Duchenne Muscular Dystrophy biomarkers using targeted proteomics

Camilla Johansson[1], Helian Vunk[2], Fredrik Edfors[2], Andreas Hober[2], Hanna Tegel[1], Björn Förssträm[2], Mirko Signorelli[3], Annemieke Aartsma-Rus[4], Erik Niks[5], Pietro Spitali[4], Mathias Uhlén[2], Cristina Al-Khalili Szigyarto[1, 2]

1. KTH- Royal Institute of Technology, Stockholm, STOCKHOLM, Sweden

2. Systems biology, SciLifeLab, Solna, Stockholm, Sweden

3. Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

4. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

5. Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands

Validated biomarkers are required to improve diagnosis, monitor disease progression and facilitate development of novel therapies. Protein biomarkers, discovered in multiplexed studies, proved however, to be difficult to implement in clinical use, due to the lack of analytical and clinical validation. In this paper, a novel orthogonal strategy is used to validate already identified serum biomarkers for Duchenne Muscular Dystrophy. Five out of ten disease progression biomarkers, previously identified by immuno-based proteomics methods, were validated using a mass spectrometry-based method. The biomarkers were analyzed by parallel reaction monitoring mass spectrometry assay in 72 longitudinal serum samples from 33 DMD patients and 12 healthy individuals. The serum concentration of carbonic anhydrase III (CA3) was measured between 346.5 and 12.2 fmol/ml in ambulant and non-ambulant DMD patients in contrast to 2.2 fmol/ml in healthy indviduals. Biomarker quantification using the PRM-MS method and an in-house developed sandwich immunoassay had a Pearson's correlation of 0.88, demonstrating the feasibility in

validating biomarkers using this strategy. This results indicate that future, orthogonal validation of DMD biomarkers is possible and reliable, providing means to assess and translate biomarkers to patient bed side.

111

Mass spectrometry workflow for characterization of plasma proteome changes related to ageing

Valentina Siino[1], Giulia Accardi[2], Sonya Vasto[3], Fredrik Levander[1]

1. Department of Immunotechnology, Lund University, Lund, SKANE LAN, Sweden

2. Department of Biomedicine, Neuroscience and Advanced Diagnostics (Bi.N.D.) , University of Palermo, Palermo, Italy

3. Department of Biological Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, Palermo, PA, Italy

Background

With simultaneous increase in life expectancy and age-related disorders, it becomes important to understand the ageing´s process to find early signs of disorders. Blood samples constitute an ideal source for molecular characterization, but the presence of highly abundant proteins and the lack of reproducibility represent challenges in plasma proteomics. However, in the past few years, fast developments of MS-based technologies and sample preparation protocols have revived the interest in plasma proteomics. The study presented here, aims to establish a robust and reproducible workflow for the quantification of plasma proteins. The method was used for the identification of blood biomarkers related to age in a cohort of 150 donors (30 to 100 years old).

Method

Plasma proteins were retrieved from healthy donors. Technical replicates were used for running two different workflows: HILIC (ReSynBiosciences) and S-trap (Protifi). Sample preparation, including digestion, was performed in 2 hours. S-trap and HILIC peptides were separated on a nano LC-system coupled with a QExactive HF-X Orbitrap (Thermo Fisher Scientific). Data were acquired in DDA and DIA mode.

Results

Without any fractionation or depletion, we were able to identify over 400 plasma proteins/sample. Comparing the two sample preparation methods, the most comprehensive results were achieved by using S-trap. Furthermore, by acquiring in DDA and DIA mode over 600 plasma proteins were identified.

Conclusions

From the present study, we concluded that the use of S-trap, coupled with DDA and DIA mode is ideal for the identification of plasma proteins. These results are encouraging, suggesting that a quick and high-throughput methodology can be employed to quantify a wide range of plasma proteins. This method may enable the classification and stratification of patients by their age as well as the identification of specific biomarkers that could predict the development of certain age-related disorders.

112

The sweet separation between bacterial and viral infections by glycopeptide profiling

Esther Willems[1], Jolein Gloerich[1], Anouk Suppers[1], Ronand de Groot[1], Alain J. van Gool[1], Hans J.C.T. Wessels[1], Marien I. de Jonge[1]

1. Radboudumc, Nijmegen, GELDERLAND, Netherlands

Background

Accurate treatment of febrile children is an important challenge faced by healthcare providers worldwide. Current diagnostic tests, based on C-reactive protein or procalcitonin, are unable to unambiguously distinguish life-threatening bacterial infections from self-resolving viral infections. This often results in empirical treatment with broad-spectrum antibiotics. Due to the emergence of antibiotic resistance there is a strong demand for improved diagnostics. Several studies have been conducted to identify better protein biomarkers but thus far failed to enhance diagnostics. The vast majority of proteins are glycosylated, which offers attractive possibilities for biomarker research since aberrant glycomics signatures have been associated with many genetic and acquired human diseases. Here, we applied glycopeptide profiling to identify site-specific glycosylation changes in response to bacterial and/or viral infection.

Methods

Blood plasma samples from 92 febrile pediatric patients, undergoing either a bacterial or viral infection, and 43 samples from healthy individuals were subjected to tryptic digestion and subsequent glycopeptide enrichment. Intact glycopeptides were analyzed by reversed phase LC-MS/MS (maXis plus; Bruker Daltonics). Data pre-processing, integration, and analysis was performed in MATLAB (MathWorks).

Results

We detected 3682 unique glycopeptide features that were consistently detected in at least 75% of the samples from any sample class. A library of glycan- and peptide-moiety identifications from a previous glycopeptide profiling study were mapped onto the feature matrix to infer glycosylation signatures for multiple sites in 37 different proteins, each carrying between 1 to 20 different glycan forms. By means of Partial Least Squares – Discriminant Analysis we obtained a subset of differential glycopeptides, which enabled a clear discrimination between sample classes.

Conclusion

The selected subset of differential glycopeptide features was able to adequately separate bacterial from viral infected patients. Current work is focused on interpretation of the glycosylation changes in response to disease and implementation of these results into novel diagnostics.

113

A comprehensive proteomic analysis of medulloblastoma subtypes reveals mechanistic insights of Group 3 tumors

Manubhai K P[1], Sanjeeva Srivastava[1], Neelam Shirsat[2], Anurag Kumar[1]

1. IIT Bombay, Mumbai, MAHARASHTRA, India

2. 3Advanced Centre for Treatment, Research and Education in Cancer (ACTREC) and Tata Memorial Hospital, Tata Memorial Centre, Kharghar, Navi Mumbai, Mumbai, India, Mumbai, India

Introduction and Objectives

Medulloblastoma is the most common pediatric brain cancer of the central nervous system and it accounts for 20% of all pediatric tumors. Earlier transcriptomics based investigations identified 4 subgroups of medulloblastoma namely, SHH, group 3, group 4 (G4), and Wnt, which is incorporated in WHO revised classification. The current study has majorly focused on identifying proteomic alteration associated with medulloblastoma subtypes and understanding mechanisms of group3(G3) tumors.

Methods

The label-free proteomic samples preparation was accomplished using Urea lysis buffer. The extracted proteins were subjected to enzymatic digestion using trypsin followed by c18 based desalting and subjected to Orbitrap fusion for MS/MS analysis. The total number of proteins obtained from the analysis was 5498 using stringent filtering criteria (1%FDR, >2 peptides). The validation of the significant candidates was performed using Multiple Reaction Monitoring (MRM)

Results and Discussion

The initial analysis was focused on the unsupervised clustering of data to investigate whether the proteomics signatures are falling under the existing transcriptomic classification by WHO and the proteomic clusters were reflecting the same. We were able to identify a panel of the classifier proteins which can differentiate between G3 and G4 tumor types. The G3 tumor exhibited differential proteomic profile in metabolic pathways specifically in glycolysis and TCA cycle. The proteins like POSTN was significantly altered in G3 tumor and it usually induced by reduced oxygen tension. MRM results were reflecting the same trend and large scale validation of key findings are ongoing

Conclusions

This comprehensive label free proteome profiling of medulloblastoma using tissue samples will provide an in-depth and mechanistic insight into proteomic alterations associated with medulloblastoma subtype.However, further validation of identified protein targets on a larger cohort of medulloblastoma patients is needed before anticipating their diagnostic impact.

Keywords

Medulloblastoma**,** Central Nervous System (CNS), World health organization (WHO)**,**

114

Investigating plant derived antimicrobial peptides using “PepSAVI-MS”

Leslie M Hicks[1]

1. University of North Carolina Chapel Hill, Chapel Hill, NC, United States

As current methods for antibiotic drug discovery are being outpaced by the rise of antimicrobial resistance, new methods and innovative technologies are crucial to replenish our dwindling arsenal of antimicrobial therapeutics. While natural products are a well-studied source of biologically active small molecules, peptidyl factors contributing to their medicinal properties remain largely unexplored. To this end, we have developed the PepSAVI-MS (Statistically-guided bioactive peptides prioritized via mass spectrometry) pipeline[1] to identify bioactive peptide targets from complex biological samples. MS/MS techniques are implemented for de novo characterization. To validate this pipeline, we have demonstrated successful detection and identification of a known antimicrobial peptide, cycloviolacin O2 (cyO2), from the botanical species *Viola odorata*. Additionally, we have widened the known antimicrobial spectrum for *V. odorata* cyclotides, including antibacterial activity of cyO2 against *A. baumannii* and novel anticancer activities for cycloviolacins by their cytotoxicity against ovarian, breast and prostate cell lines. The developed platform is highly versatile as it is adaptable to any natural product source of peptides and can test against diverse physiological targets, including bacteria, fungi, viruses, protozoans, and cancer cells for which there is a developed bioassay. As such, we demonstrate extension of this pipeline to fungal and bacterially-sourced AMPs through the identification of the killer toxin KP4 from *Ustilago maydis*[2] and the bacteriocin Bac-21 from *Enterococcus faecalis* harboring pPD1[3]. Additionally, we begin to probe the vast array of botanical natural product sources to prioritize highly active species for downstream analysis.

References

1. "The “PepSAVI-MS” pipeline for natural product bioactive peptide discovery." *Analytical chemistry*2 (2017): 1194-1201.

2. "Fungal Secretome Analysis via PepSAVI-MS: Identification of the Bioactive Peptide KP4 from Ustilago maydis." *JASMS*5 (2018): 859-865.
3. "Exploring bioactive peptides from bacterial secretomes using Pep SAVI-MS: identification and characterization of Bac-21 from Enterococcus faecalis pPD 1." *Microbial biotechnology*5 (2018): 943-951.

115

Mapping proteome-wide targets of protein kinases in plant stress responses

Pengcheng Wang[1]

1. Shanghai Center for Plant Stress Biology, Chinese Academy of Sciences, Songjing, SHANGHAI, China

Protein kinases are major regulatory components in almost all cellular processes in eukaryotic cells. By adding phosphate groups, protein kinases regulate the activity, localization, protein-protein interactions and other profiles of their target proteins. It is known that protein kinases are central components in plant response to environmental stresses such as drought, high salinity, cold and pathogen attack. Although the phosphorylation-dependent signaling plays critical roles in plant stress biology, only a few targets of these protein kinases have been identified. Moreover, how these protein kinases regulate the downstream biological processes and mediate the stress responses are still largely unknown. In this study, we introduce a novel strategy based on isotope-labeled in vitro phosphorylation reactions using in vivo phosphorylated peptides as substrate pools and apply this strategy to identify in putative substrates of nine protein kinases that function in plant abiotic and biotic stress responses. As a result, we identified more than 5,000 putative target sites of osmotic stress-activated SnRK2.4 and SnRK2.6, ABA activated protein kinases SnRK2.6 and Casein-Kinase Like 2 (CKL2), elicitor-activated protein kinase CDPK11 and MPK6, cold-activated protein kinase MPK6, H_2O_2-activated protein kinase OXI1 and MPK6, salt-induced protein kinase SOS1 and MPK6, as well as the low-potassium-activated protein kinase CIPK23. These results provide comprehensive information on the role of these protein kinases in the control of cellular activities and could be valuable resource for further study on the mechanism underlying plant response to environmental stresses.

116

Chitosan remodels Extracellular Matrix Integrity and regulate Stomatal Function leading to Immunity against Wilt disease

Subhra Chakraborty[1]

1. National Institute of Plant Genome Research, New Delhi, DELHI, India

Extracellular matrix (ECM) acts as a physical scaffold preventing recognition and entry of phyto-pathogens while guard cell perceives and integrates signals metabolically. Wilt disease caused by *Fusarium oxysporum* is a major impediment for crop productivity. Chitosan and its acetylated form chitin are known MAMPs implicated in plant defense, the precise mechanism of chitosan triggered immunity (CTI) that leads to resistance against pathogens remains unknown. To understand the role of ECM and CTI against wilt disease, examination of stomata, ECM components, ROS, NO, eATP levels was investigated in wilt resistant and susceptible chickpea (*Cicer arietinum*) genotypes. Multiomics using quantitative ECM proteomics and metabolomics was investigated .

Patho-stress was imposed on untreated and chitosan treated three-week-old chickpea seedlings and tissues were harvested at different post-infection time points. Stomatal mechanics were measured using GFS3000. ECM components were analysed in scanning electron microscopy and raman spectroscopy. Temporal ECM proteome and metabolome was developed using iTRAQ coupled Triple-TOF/MS and GC-MS analyses, respectively. Chitosan-induced immunity related network was built and qRT-PCR analysis was performed to validate omicsdatasets.

Morpho-histological examination revealed stomatal closure, reduced stomatal conductance and transpiration rate in chitosan-treated compared to untreated seedlings upon fusariosis. ECM showed fortification leading to oligosaccharide signalling as documented by increased galactose, pectin and secondary carbohydrates. Multiomics using ECM proteomics and metabolomics identified 325 chitosan-triggered IRPs (CTIRPs), 65 CTIRMs that includes, LYM2, kinases, sugars, organic acids and aminoacids linking ROS production, stomatal movements, root nodule development and architecture. Data provide evidence that ROS, NO and eATP governs CTI. Induction of PR proteins, CAZymes and PAL was observed during CTI.

Chitosan regulate ECM dynamics, stomatal movements and root architecture. ECM signaling and stomatal immunity plays pivotal role in CTI. Multi-omics provides evidence how chitosan imparts immunity against Fusariosis. Increased ROS production, antioxidant enzymes, NO and eATP restrict *Fusarium* attack.

117

Elucidation of tolerant mechanism in soybean treated with plant-derived smoke under flooding stress using omics technique

Setsuko Komatsu[1]

1. Fukui University of Technology, Fukui, Japan

Global climate change influences the magnitude and frequency of hydrological fluctuations and causes an unfavorable environment for plant growth and development. Food shortages are one of the most serious global problems in this century, and it is important to increase food production. Soybean, which is rich in protein and vegetable oil, is cultivated in several climatic zones; however, its growth is markedly decreased due to flooding. On the other hand, smoke is used in traditional farming systems for improving seed germination and seedling vigor. To clarify the mechanism of flooding tolerance in early-stage soybean, plant-derived smoke was applied, which exhibited a flooding tolerant phenotype. The growth of soybean seedlings was suppressed under flooding stress, but it recovered after water removal following treatment with plant-derived smoke. Early-stage soybeans treated with plant-derived smoke under flooding stress were collected for gel-free/label-free proteomic, RNA-sequencing based transcriptomic, and mass-spectrometry based metabolomic analyses. Data sets were analyzed using functional categorization/cluster separation; furthermore, enzyme-activity, immuno-blot, and gene-expression experiments were performed to confirm the data acquired from comprehensive analyses. These results suggest that plant-derived smoke protects newly synthesized proteins via ribosomal protein and enhancing the activities of antioxidative enzymes to remove reactive-oxygen species. Furthermore, an integrated approach of proteomics with computational genetic modification effectiveness analysis was applied to explore flood-tolerant genes in soybean, suggesting that proteins related to energy metabolism might play an essential role to confer flood tolerance in soybeans.

118

Morphological and proteomics analyses of petioles rigidity in sacred Lotus

Pingfang Yang[1]

1. State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, CN

Very little is known about the mechanism controlling petiole rigidity in sacred lotus (*Nelumbo nucifera* Gaertn.). To investigate the mechanism controlling the lotus petiole rigidity, morphological and proteomic analyses were performed. Anatomically, there is a great variation between the petioles of floating and vertical leaves. The number of vascular bundles, ligneous cells and thickness of cell wall were higher in the initial vertical leaf petiole (IVP) compared to the initial floating leaf petiole (IFP). A total of 4855 proteins were quantified through comparative proteomic analysis, among which 421 proteins expressed 1.5 folds higher in IFP and 483 proteins expressed 1.5 folds higher in IVP. Protein function categories indicated hundreds of proteins involved in cell wall biosynthesis, organization and assembly. Functional enrichment analysis for the differentially abundant proteins indicated the enrichment of 105 proteins in 6 different pathways, while 43 out of them were enriched in lignin biosynthesis pathway. In consistent with genes and proteins expressions in lignin biosynthesis, the contents of lignin monomers precursors were significantly different in IFP and IVP. These findings support the involvement of lignin in lotus petioles rigidity.

119

Molecular changes in the course of ice plant C_3 to CAM transition revealed by proteomics and metabolomics

Sixue Chen[1], Qijie Guan[1], Wenwen Kong[1], Mi-Jeong Yoo[1], Hope Hersh[1], Daniel Dufresne[2], Jingkui Tian[3]

1. University of Florida, Gainesville, FL, United States

2. Chemistry, Florida Atlantic University, Miami, FL, USA

3. Zhejiang University, Hangzhou, Zhejiang, P.R. China

Crassulacean acid metabolism (CAM) is a specialized type of photosynthesis: stomata close during the day, enhancing water conservation, and open in the night, allowing CO_2 uptake. Therefore, water use efficiency (WUE) of CAM plants is much higher than C_3 and C_4 plants under comparable growth conditions. *Mesembryanthemum crystallinum* (common ice plant) is a facultative CAM species that can switch from C_3 to CAM under salt or drought stresses. However, the molecular mechanisms underlying the transition from C_3 photosynthesis to CAM remain unknown.

Common ice plant leaves were used to determine the transition period from C3 to CAM photosynthesis after salt or drought stresses. Leaf diurnal changes in stomatal movement, carbon assimilation, vacuolar acidity, gene expression of key CAM enzymes (e.g., phosphoenolpyruvate carboxylase (PEPC) and PEP carboxykinase (PEPCK)), as well as their activities were measured to determine the critical transition time points of CAM initiation from C3. Leaves, stomatal guard cells, and mesophyll cells were collected for proteomics and metabolomics analyses. Bioinformatics and machine learning tools were used to determine molecular switches at different levels of regulation.

We have determined the transition time from C_3 to CAM in *M. crystallinum* under drought and salt stress. Our previous transcriptomics work identified 495 differential transcripts between control and salt-treated samples during the C_3 to CAM transition, including seven CAM-related genes, 18 transcription factors, and 285 known guard cell expressed genes. *PEPC1* and *PPCK1*, which encode key enzymes of the CAM photosynthesis, were up-regulated in guard cells after seven days of salt treatment, indicating that guard cells themselves can transit from C_3 to CAM. Label-free quantitative proteomics and metabolomics data have been acquired, and they will be presented after quality control and statistical/bioinformatic analyses.

This study unravels novel molecular switches and provides important information towards introducing CAM into crops to enhance WUE and yield.

120

Proteome and proteoform diversity exposed and quantified by hybrid mass spectrometry approaches

Albert Heck[1]

1. Utrecht University, Utrecht, NL, Netherlands

Through advances in technologies the role of mass spectrometry in structural and molecular biology is rapidly expanding. Especially hybrid MS approaches combining top-down, middle-down and bottom-up proteomics analysis with native MS, HD exchange MS or XL-MS assist us to provide unprecedented detail on the structural and function heterogeneity present in dynamically evolving proteins and protein assemblies, including complex plasma glycoproteins but also in intact ribosomes and viruses.

In this talk I will describe novel developments in MS instrumentation for native MS and top-down MS, especially new modifications to an Orbitrap based instrument that offers high-sensitivity and mass resolving power, allowing an in-depth detailed analysis of highly glycosylated plasma glycoproteins, protein assemblies up to even whole intact ribosomes and viruses.

I will describe how we use these new mass analysers to characterize in depth pharmaceutical proteins and plasma glycoproteins. Thereby we focus on how the proteoform profiles of targeted plasma glycoproteins respond to perturbations such as vaccination, sepsis and cancer.

I will also describe the latest developments in the lab around cross-linking mass spectrometry where we attempt to perform proteome-wide cross-linking using XlinkX, extending depth by imporved fragmentation schemes and our in-house developed enrichable cross-linker PhoX.

Finally, I will highlight some of the latest developments in using hybrid fragmentation methods including HCD, ECD/ETD and UV/IRMPD laser photodissociation for native and top-down proteomics.

121

Organelle shape and function in the context of viral infections

Ileana M Cristea[1]

1. Princeton University, , United States

Shape and function are tightly linked in biology. At the tissue and cell level, an alteration in shape induces the activation of numerous signaling cascades. This ability to transmit information regarding shape is essential for numerous biological processes, including the sensing of mechanical force during development. At the intracellular level, changes in the shape of cellular organelles are core components of cell cycle, metabolism, apoptosis, and transcriptional regulation. This tight link between organelle shape and function is exploited by viruses during infections as mechanisms acquired to either support virus replication or inhibit host defense responses. Here, we investigate mechanisms underlying alterations in shape and the molecular function of these changes in different biological contexts. For these studies, we integrate microscopy, proteomics, lipidomics, and the development of mathematical modeling and computational platforms for data analysis. We uncover finely-tuned temporal alterations in organelle shape that are used to either activate or inhibit specific organelle functions at different stages of infection. Examples include peroxisome alterations that induce metabolic changes to support virus production, as well as mechanisms controlling lamina integrity at the nuclear periphery to inhibit virus capsid egress. Our studies also point to shape alterations that are connected to protein movements between organelles and changes in protein interactions. We present our efforts to globally characterize the formation and dissociation of protein interactions during the progression of herpesvirus infection. Finally, we report the development of a computational platform that allows the users to visualize protein interactions that are dynamic in space and time, and to integrate information of subcellular localization, functional annotations, and protein abundances.

122

A multi-omics approach to drug target discovery for novel bis-triazine antimalarials

Katherine Ellis[1]**, Ghizal Siddiqui**[1]**, Jonathan Baell**[1]**, Stuart Ralph**[2]**, Darren Creek**[1]

1. Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Vic, Australia

2. Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, VIC, Australia

Malaria causes 445 000 deaths annually and threatens approximately 40% of the world population. The malaria parasite has developed resistance to most approved antimalarials, and there is a critical need to discover new drugs with novel mechanisms of action. We have identified a novel series of bis-triazines with potent antimalarial activity in cell and animal models. The aim of this work was to identify the mechanism of action of these novel compounds.

As these bis-triazines are structurally distinct from other known antimalarials, and no prior mechanistic information was available, the mode of action was explored using a combination of untargeted comparative multi-omics analysis and chemical proteomics.

Incubation of *P. falciparum*-infected red blood cells with a potent bis-triazine compound induced a unique metabolomic profile that differed from other known antimalarials. A dose-dependent accumulation of dimethyl-arginine (DMA) was the most significant unique metabolic perturbation observed in treated cells, and levels of monomethylated arginine and lysine were also increased. Stable-isotope tracing suggested that the DMA accumulation was due to inhibition of demethylase activity or increased degradation of methylated proteins. Gene ontology enrichment analysis of the proteomics data revealed depletion of several

nucleic-acid binding proteins in treated parasites, and peptidomics revealed accumulation of peptides from methylated nuclear proteins.

Chemical proteomics studies using bifunctional bis-triazine analogues bearing photoreactive and 'click chemistry' motifs allowed enrichment of triazine-bound proteins, and LC-MS analysis identified 15 candidate proteins as potential drug targets. Five of these candidates are nuclear proteins, which currently represent the most likely candidates based on the impact on nuclear proteins observed in the multi-omics studies. Fluorescence microscopy has confirmed co-localisation of bis-triazine probes with the parasite nucleus and further work is ongoing to confirm the specific protein target(s) responsible for the action of these promising new antimalarial compounds.

123

Systems immunology reveals factors driving anti-viral CD8 T cell immunity

Nathan P Croft[1], Ting Wu[1], Jing Guan[1], Stewart A Smith[2], Andreas Handel[3], Mohammad Shahbazy[1], Jana Pickering[2], John Sidney[4], Bjoern Peters[4], Pouya Faridi[1], Matthew J Witney[2], Prince Sebastian[2], Inge EA Flesch[2], Sally L Heading[2], Linda M Wakim[5], Xavier YX Sng[1], Alessandro Sette[4], Paul G Thomas[6], Nicole L La Gruta[1, 5], David C Tscharke[2], Anthony W Purcell[1]

1. Infection and Immunity Program, Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC, Australia

2. John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia

3. Department of Epidemiology and Biostatistics, Health Informatics Institute and Center for the Ecology of Infectious Diseases, University of Georgia, Athens, Georgia, 30602, USA

4. La Jolla Institute for Allergy and Immunology, La Jolla, California, USA

5. Department of Microbiology and Immunology, University of Melbourne, The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC, Australia

6. Department of Immunology, St. Jude Children's Research Hospital, Memphis, TN, USA

Background: The magnitude of $CD8^+$ T lymphocyte responses to infection are a function of a multitude of factors that include the available naïve T cell repertoire combined with the context and duration of peptide presentation by major histocompatibility complex (MHC) molecules. Measurement of these factors therefore allows an assessment of their contribution to – and ultimately a chance to model and predict – immunogenicity.

Methodologies: Here, we have used a combination of discovery and targeted mass spectrometry to assess the virus-derived immunopeptidome following different models of virus infection and antigen presentation, combined with measurements of peptide-MHC binding affinity and the elicited CD8 T cell response of each peptide.

Results: These data highlight the diversity in viral peptide-MHC display following infection and the dramatic changes in absolute abundance and kinetics, as well as modulation of the presentation hierarchy in direct- versus cross-presentation pathways. Screening of each peptide for immunogenicity in multiple virus-infected mice revealed a wide range of immunogenicities. Together, these data have helped to train mathematical models that delineate the relative importance of these factors and provide a critical step towards predicting immunogenicity.

Conclusions: This study highlights how high quality quantitative proteomics and peptidomics data are pivotal in unraveling the complex ecosystem of immune responses to viruses and provides the foundation for the rational design of interventional and therapeutic strategies for viruses that remain a serious threat to humankind.

124

A *Mycobacterium tuberculosis* protein atlas

Amir Banaei-Esfahani[1, 2], Julia Feldmann[3], Lukas Folkman[4], Sonia Borrell[3], Andrej Trauner[3], Karsten Borgwardt[4], Ben Collins[1], Sebastien Gagneux[3], Rudolf Aebersold[1, 5]

1. Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

2. PhD Program in Systems Biology, University of Zurich and ETH Zurich, Zurich, Switzerland

3. Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland

4. Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

5. Faculty of Science, University of Zurich, Zurich, Switzerland

Tuberculosis is caused by *Mycobacterium tuberculosis* (Mtb) and claims 1.8 million lives annually. Clinical strains of Mtb reveal diverse phenotypes that are largely determined by the state of the proteome. Therefore, the model strain H37Rv that has been used for most studies in the field does not represent the full genomic and phenotypic diversity of Mtb. Here, we aim to identify master regulators that control the Mtb proteome of genetically diverse strains driving different phenotypes in two conditions.

SWATH-MS profiled the proteome of 70 clinical strains of Mtb cultivated under normal and nitric oxide stress, a major bactericidal agent within macrophages. Moreover, we fully sequenced the genome of the respective strains. Six strains were subjected to transcriptional measurements.

We quantified ~2700 proteins corresponding to ~80% of the expressed genes, reproducibly across the large Mtb sample cohort. To address the aim of the project, we developed two exploratory computational frameworks, a genome-scale transcriptional model and dysregulation analysis for protein complexes, and analyzed the generated dataset using those two pipelines as well as QTL mapping. The data indicated that the basal expression and response of various protein functional groups such as IdeR and DosR regulon significantly differ between Mtb lineages. Twenty-eight transcription factors orchestrating the Mtb transcriptional network between lineages have been identified. Several subnetworks are regulated differently in terms of their stoichiometry across various lineages including the protein association Rv0068-Rv2187. QTL analysis, implemented for the first time in bacteria, revealed that a single mutation in KstR affects the expression of its seven targets involved in cholesterol

metabolism, the main carbon source during Mtb infection. It further provided the first evidence of the existence of protein isoforms in Mtb. We have shown that genomic differences between the clinical isolates determine the state of the proteome and mediate different clinically relevant phenotypes.

125

Spatiotemporal proteotype analysis of vaccinia virus infected cells reveals dynamic host surfaceome repopulation with viral proteins

Fabian Wendt[1, 2], Moona Huttunen[3], Karel Novy[1, 4, 2], Jason Mercer[3], Bernd Wollscheid[1, 2]

1. Institute of Molecular Systems Biology & Dep. of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

2. Swiss Institute of Bioinformatics (SIB), Switzerland

3. Medical Research Council Laboratory for Molecular Cell Biology, University College London, London, United Kingdom

4. current address: Biognosys, Zurich, Switzerland

Viral pathogens, such as vaccinia virus (VACV), hijack host signalling networks in order to survive and propagate themselves. VACV is the prototypic member of the *Poxviridae*, a family of large double-stranded DNA viruses that include variola virus, the causative agent of smallpox. For many pathogens, including VACV, the molecular mechanisms underlying successful survival, but also cellular defence and pathogen clearance, are not well understood. Especially, detailed molecular knowledge about protein composition, abundance and nanoscale organisation within the surfaceome of infected cells is sparse. The knowledge about the involved signaling networks and extracellular protein interactions engaged during infection would enable rational engineering anti-viral strategies. Here, we investigated using spatiotemporal analysis how VACV infection perturbs the host proteotype with specific focus on longitudinal surfaceome changes.

Using our recently developed, automated and miniaturized Cell Surface Capture (autoCSC) technology, we took relative quantitative proteotype surfaceome snapshots of VACV infected Hela CCL2 cells across 24 hours of the infectious life cycle of VACV. AutoCSC, combined with DIA-based mass spectrometric analysis, enabled the sensitive identification of the pool of cell surface-residing glycoproteins in a time-resolved manner. Subsequently, the longitudinal analysis allowed for the profiling of surfaceome proteotype dynamics during the viral life cycle.

We observed moderate changes within the host surfaceome post-infection, but additionally detected a set of viral glycoproteins. These VACV proteins are translated and modified/glycosylated by the host's cellular machinery. Subsequently, they relocalize into the plasma membrane. In order to elucidate the functional role of these viral proteins at the host surfaceome, we investigated their trans-interaction/receptor space applying chemo-proteomic technologies. Together, we found that the cell's acute infection state is characterized by a repopulation of the host surfaceome with viral glycoproteins which supports VACV survival and eventually immune evasion.

126

Salmonella proteomic profiling during infection distinguishes the intracellular environment of host cells

Zezhou Li[1], Xiaoyun Liu[2]

1. College of Chemistry and Molecular Engineering, Peking University, Beijing, China

2. School of Basic Medical Sciences, Peking University, Beijing, China

Essential to bacterial pathogenesis, *Salmonella enterica* serovar Typhimurium (*S*. Typhimurium) has evolved the capacity to quickly sense and adapt to specific intracellular environment within distinct host cells. Here we examined *S*. Typhimurium proteomic remodeling within macrophages, allowing direct comparison with our previous studies in epithelial cells. In addition to many shared features, our data revealed proteomic signatures highly specific to one type of host cells. Notably, intracellular *S*. Typhimurium differentially regulates the two type III secretion systems (T3SSs) far quicker in macrophages than in epithelial cells, so do bacterial flagellar and chemotaxis systems degenerate. Importantly, our comparative analysis uncovered vast induction of bacterial histidine biosynthesis in macrophages but not in epithelial cells. Targeted metabolomic measurements revealed markedly lower histidine levels within macrophages. Intriguingly, further functional studies established that defective histidine biosynthesis (due to a *hisG* mutation) renders the bacterium (strain SL1344) hypersensitive to intracellular shortage of this amino acid. Indeed another *S*. Typhimurium strain 14028s with a fully functional biosynthetic pathway exhibited only minor induction of the *his* operon within infected macrophages. Our work thus reveals novel insight into *S*. Typhimurium adaptation mechanisms within distinct host cells and also provides an elegant paradigm where proteomic profiling of intracellular pathogens is utilized to discriminate specific host environment (e.g., on nutrient availability).

1. Li Z, Liu Y, Fu J, Zhang B, Cheng S, Wu M, Wang Z, Jiang J, Chang C, Liu X. (2019) Salmonella proteomic profiling during infection distinguishes the intracellular environment of host cells. mSystems, 4, e00314-18.
2. Wang Z, Sun J, Xia T, Liu Y, Fu J, Lo YK, Chang C, Yan A, Liu X. (2018) Proteomic delineation of the ArcA regulon in Salmonella Typhimurium during anaerobiosis. Mol. Cell. Proteomics, 17, 1937-1947.
3. Liu Y, Liu Q, Qi L, Ding T, Wang Z, Fu J, Hu M, Li M, Song J, Liu X. (2017) Temporal regulation of a Salmonella Typhimurium virulence factor by the transcriptional regulator YdcR. Mol. Cell. Proteomics, 16, 1683-1693.
4. Liu Y, Yu K, Zhou F, Ding T, Yang Y, Hu M, Liu X. (2017) Quantitative proteomics charts the landscape of Salmonella carbon metabolism within host epithelial cells. J. Proteome Res., 16, 788-797.
5. Liu Y, Zhang Q, Hu M, Yu K, Fu J, Zhou F, Liu X. (2015) Proteomic analyses of intracellular Salmonella enterica serovar Typhimurium reveal extensive bacterial adaptations to infected host epithelial cells. Infect. Immun., 83, 2897-2906.

127

PASS-DIA: A novel data-independent acquisition approach for discovery studies

Akhilesh Pandey[1, 2, 3, 4, 5], Dong-Gi Mun[1, 2], Santosh Renuse[1, 2], Mayank Saraswat[2, 3, 4, 5], Anil Madugundu[1, 2, 3], Savita Udainiya[1, 2], Hui Zhao[2], Raja Sekhar Nirujogi[4, 6], Chan-Hyun Na[6, 7], Nagarajan Kannan[2], Sang-Won Lee[8]

1. Center for Individualized Medicine, Mayo Clinic, Rochester, USA

2. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, USA

3. Center for Molecular Medicine, National Institute of Mental Health and Neurosciences , Bangalore, India

4. Institute of Bioinformatics, International Technology Park, Bangalore, India

5. Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India

6. McKusick-Nathans Institute of Genetic Medicine , Johns Hopkins University School of Medicine, Baltimore, USA

7. Neurology, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, USA

8. Department of Chemistry, Center for Proteogenome Research, Korea University, Seoul, Republic of Korea

Data-independent acquisition (DIA) approach is being increasingly adopted as a promising strategy for identification and quantitation of known peptides across a large set of samples. As most of DIA datasets are acquired with much wider isolation windows than data-dependent acquisition (DDA) experiments, complex MS/MS spectra are generated, which hampers obtaining maximum peptide information through classical protein database search methods. Therefore, analysis of DIA data mainly relies on evidence of existence of peptides from pre-built spectral libraries using a peptide-centric approach. Consequently, one major weakness of this method is that it does not account for peptides which are not included in spectral library, precluding the use of DIA for discovery studies. Here, we present a strategy termed PASS-DIA (Precursor ion And Small Slice-DIA) in which MS/MS spectra are acquired with small isolation windows and MS/MS spectra are interpreted with accurately measured precursor ion masses. This method enables direct application of conventional spectrum-centric analysis pipelines for peptide identification and precursor ion-based quantitation. The performance of PASS-DIA was superior to both DDA and conventional DIA experiments with regard to identification of peptides. Application of PASS-DIA for analysis of samples with post-translationally modified peptides such as phosphorylation and N-glycosylation again revealed its superior performance. Finally, the use of PASS-DIA to characterize a rare proteome of human fallopian tube organoid samples revealed biologically relevant and low abundance proteins. Overall, PASS-DIA is a novel DIA approach for use as a discovery tool which outperforms both conventional DDA and DIA experiments to provide additional protein information. We believe that PASS-DIA method will become an important strategy for discovery type studies when deeper proteome characterization is necessitated.

128

Proteogenomics — connecting cancer genotype with molecular phenotype

Janne Lehtio[1]

1. Science for Life Laboratory, Dep of Oncology and Pathology, KAROLINSKA INSTITUTET, SOLNA, SE

The explosion of genomics data has improved our understanding of cancer greatly in recent years. However, the knowledge of how genomic aberrations affect the functional proteome at the systems level is still very limited. Proteome data represents the combined effect of epigenetic, transcriptional and translational regulation and will therefore provide an important **molecular phenotype data layer** for multi-omics analysis. To allow effective systems biology analysis including proteomics, we have generated tools that take advantage of massive genomics data by incorporating sequence information to the proteomics data-analysis pipeline. This will allow protein level analysis of gene variants as well as detection of novel protein coding regions[1]. To control error rate in variant detection, we have a combined experimental isoelectric point data from peptide fractions (**HiRIEF** LC-MS/MS) and bioinformatics approaches into the proteogenomics workflow (**IPAW**)[2]. A proteogenomics analysis of histologically human tissues using the IPAW pipeline reveals novel coding regions. When applied on breast cancer tumor sample, we could demonstrate in-depth quantitative analysis revealing drug target interesting correlations as well as discovers putative cancer neoantigens[3]. To gain knowledge of the novel proteins, we analyzed the subcellular location of these in human cell line models. For location analysis, we used **SubCellBarcode** based proteome wide location analysis[4]. Further, the single amino acid variant detection pipeline enabled detection of paternal and maternal proteins transferring placenta during pregnancy, suggesting molecular communication between fetus and mother[5].

1. Branca M., Orre L-M., Johansson H.J., Granholm V., Huss M., Pérez-Bercoff Å., Forshed J., Käll L., Lehtiö J. HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nature Methods*, 2014 Jan;11(1):59-62.
2. Zhu Y, Orre LM, Johansson HJ, Huss M, Boekel J, Vesterlund M, Fernandez-Woodbridge A, Branca RMM, Lehtiö J. Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow. *Nature Commun*. 2018 Mar 2;9(1):903.
3. Johansson HJ, Socciarelli F, Vacanti NM, Haugen MH, Zhu Y, Siavelis I, Fernandez A, Aure MR, Sennblad B, Vesterlund M, Branca RM, Orre LM, Huss M, Fredlund E, Beraki E, Garred Ø, Boekel J, Sauer T, Zhao W, Nord S, Höglander EK, Jans CD, Brismar H, Haukaas TH, Bathen TF, Schlichting E, Naume B, OSBREAC, Luders T, Borgen E, Kristensen VN, Russnes HG, Lingjærde OC, Mills GB, Sahlberg KK, Børresen-Dale L, Lehtiö J., Breast cancer quantitative proteome and proteogenomic landscape. *Nature Comm*, 10 Apr, 2019.
4. Orre LM, Vesterlund M, Pan Y, Arslan T, Zhu Y, Fernandez Woodbridge A, Frings O, Fredlund E, Lehtiö J. SubCellBarCode: Proteome-wide mapping of protein localization and relocalization. *Mol Cell*. 2019 Jan 3.
5. Pernemalm M, Sandberg A, Zhu Y, Boekel J, Tamburro D, Schwenk JM, Bjork A, Wahren-Herlenius M, Amark H, Ostenson C.G, Westgren M, Lehtiö J. In-depth human plasma proteome analysis captures tissue proteins and transfer of protein variants across the placenta. *eLife*, Apr. 8., 2019.

129

Implementing ProteoGenome-Driven Oncology and Global Data Sharing

Henry Rodriguez[1]

1. National Cancer Institute, National Institutes of Health, Bethesda, MARYLAND, United States

Giving the right treatments to the right patients at the right time, known as *precision medicine*, has become the new ideal for effective care. As progress has been made to develop precision medicine, genomic information has been an integral part of the effort. It has been used to differentiate various tumor types by gene transcript abundance and facilitate therapeutic drug selection for patients to have remarkable responses - including complete remissions.

However, for most clinical problems, precision strategies remain aspirational. This is due, in large part, to the many layers of regulation between gene transcription and protein expression, and even further to post-translational modifications. These regulatory layers make it exponentially challenging to identify important biological variances for effective treatment from transcriptional differences alone, and becomes a barrier for optimal intervention and patient care. As a result, precision medicine has yet to deliver for the vast majority of cancer patients.

But as precision medicine develops, the emerging field of *proteogenomics* provides an opportunity to generate new cancer insights through the melding of genomics *and* proteomics, allowing a more complete understanding of how somatic genomes activate aberrant signal transduction events that drive cancer pathogenesis. This seminar will discuss how genomics, transcriptomics, and proteomics are being combined in the quest to better understand the etiology of cancer – in basic clinical sample studies and translational research (clinical trials). It will also discuss how some of the world's leading cancer research centers have united around proteogenomics through the U.S. National Cancer Institute's Cancer Moonshot activities. By comprehensively characterizing commonly diagnosed cancers within their respective populations, these research centers aim to develop and coalesce a public cancer proteogenome atlas representative of the diversity of people with cancer worldwide.

130

Proteomic Data Commons: a resource for proteogenomic analysis

Chris R Kinsinger[1], **Ratna R Thangudu**[2], **Erika Kim**[1], **Paul A Rudnick**[3], **Michael Holck**[2], **Michael MacCoss**[4], **Anand Basu**[2]

1. National Cancer Institute, Bethesda, MD, United States

2. ESAC, Rockville, MD, USA

3. Spectragen-Informatics, Seattle, Washington, USA

4. University of Washington, Seattle, Washington, USA

The objective of the Proteomic Data Commons (PDC)[1] is to make cancer-related proteomic datasets accessible to the public. As a domain-specific repository within the Cancer Research Data Commons (CRDC), the vision for the PDC is to provide researchers the ability to find and analyze proteomic data across a variety of tumor types. Currently, the PDC houses data, supported by a large collection of metadata attributes, for ~20 datasets produced by CPTAC and other large-scale cancer research programs, each with cohort sizes >100. Most of the datasets in the PDC also have corresponding genomic data and images available in the Genomic Data Commons and The Cancer Imaging Archive.

The PDC is continuing the trend to replace downloading multiple local copies of data with bringing software and tools to the data in the cloud for analysis. Users may bring their own tools to co-analyze genomic and proteomic data available from a common sample. They can also define cohorts of interest and perform their own analysis of any of the publicly available data or a combination of public and private data. Private data may be stored in a PDC workspace.

One specific type of analysis facilitated through the PDC is proteogenomic analysis. The PDC provides quick access to mapping of peptide identities and quantities on the human genome as well as patient/tumor-specific protein databases containing genomic events such as single nucleotide variants and alternative splicing. It also enables fast, accurate, and convenient proteomic validation of novel genomic alterations through PepQuery.[2] These tasks are enabled through close integration with the Genomic Data Commons and NCI's Cloud Resources.

The presentation will provide an overview of the PDC with specific examples of proteogenomic analyses.

1. [1] The PDC is available at: https://pdc.esacinc.com/pdc/
2. [2] Genome Res. 2019. 29: 485-493

131

The proteogenomic landscape of curable prostate cancer

Ankit Sinha[1], **Vincent Huang**[2], **Julie Livingstone**[2], **Thomas Kislinger**[1, 3], **Paul C Boutros**[2, 1, 7, 4, 5, 6]

1. Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

2. Ontario Institute for Cancer Research, Toronto, Ontario, Canada

3. Princess Margaret Cancer Centre, Toronto, Ontario, Canada

4. Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada

5. Department of Human Genetics, University of California, Los Angeles, California, USA

6. Jonsson Comprehensive Cancer Centre, University of California, Los Angeles, California, USA

7. Department of Urology, University of California, Los Angeles, California, USA

Contributions from large international consortia have extensively characterized the mutational landscape and signaling alterations that drive tumour initiation and progression. For example, DNA sequencing has identified recurrent mutations that drive the aggressiveness of prostate cancers. Surprisingly, the influence of genomic, epigenomic, and transcriptomic dysregulation on the tumour proteome remains poorly understood. Hence, integration along the central dogma may provide more accurate multi-omic biomarkers. To test this hypothesis, we systematically profiled and integrated whole-genome, epigenome, transcriptome and proteome profiles of 76 clinically annotated, localized intermediate-risk prostate cancer tumours [1]. As a result, we discovered that genomic subtypes of prostate cancer converge on five proteomic subtypes, with distinct clinical trajectories. Although ETS fusions are the most common alteration in prostate tumours; however, they influence different genes and pathways at the proteome and transcriptome. Additionally, changes in mRNA abundance explained approximately 10% of the variation in protein abundance. Hence, in our study, prognostic biomarkers that combine genomic or epigenomic features with proteomic features significantly outperform biomarkers comprised of a single data type.

1. Sinha et al., 2019, Cancer Cell 35, 414–427

132

Mapping functional interactions of testis germ cell-specific proteins with proteogenomic assays

Christina Schiza[1], Mirzo Kanoatov[2], Keith Jarvi[2], Andrei Drabovich[3]

1. Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

2. Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada

3. Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB, Canada

Recent studies on tissue specificity revealed ~900 genes with restricted expression in the testis germ cells. These genes encode proteins involved in spermatogenesis and may be essential for reproduction. A fraction of testis-specific genes, however, is functionally redundant and was generated by random duplications to eventually drive evolution of proteins with novel functions. We hypothesized that mapping of interactomes of testis germ cell-specific proteins may reveal their functional role in spermatogenesis and identify genes essential for human reproduction. Since male germ cells cannot not be cultured *in vitro*, we studied germ cell-specific proteins in human clinical samples, such as testicular tissues and spermatozoa.

We first investigated the interactome of TEX101, a germ cell-surface chaperone and a validated biomarker of male infertility [1]. We identified by co-immunoprecipitation-mass spectrometry (co-IP-MS) the physical interactome of human TEX101. Germ cell-surface dipeptidase 3 emerged as a top hit, and TEX101-DPEP3 interaction was extensively validated [2].

We then designed a proteogenomic approach to identify germ cell-surface proteins which were degraded in the absence of TEX101 chaperone. Genotyping of 386 men revealed four men homozygous for rs35033974, a missense variant resulting in the near-complete degradation of the variant G99V TEX101 protein. Differential proteomics with label-free quantification measured 8,046 proteins in spermatozoa and identified germ cell-surface proteins down-regulated in patients homozygous for rs35033974. Significantly reduced levels of LY6K protein were confirmed by targeted proteomics and immunofluorescence. Since fathers homozygous for rs35033974 had biological children, TEX101 could be a nonessential and functionally redundant human gene [3]. Our deep proteome profiling of the human spermatozoa also revealed 46 testis-specific proteins with no prior evidence at protein level.

Our current efforts are devoted to development of co-IP-MS and proteogenomic assays for numerous germ cell-surface proteins in order to map their interactomes, validate their essentiality, and investigate as male infertility biomarkers.

1. Korbakis, D.; Schiza, C.; Brinc, D.; Soosaipillai, A; Karakosta, T.D.; Legare, C.; Sullivan, R.; Mullen, B.; Jarvi, K; Diamandis, E.P.; Drabovich, A.P. (2017) BMC Medicine, 15, 60
2. Schiza, C.; Korbakis, D.; Panteleli, E.; Jarvi, K.; Drabovich, A.P.; Diamandis, E.P. (2018) Molecular & Cellular Proteomics, 17, 2480-2495
3. Schiza, C.; Korbakis, D.; Jarvi, K.; Diamandis, E.P.; Drabovich, A.P. (2019) Molecular & Cellular Proteomics, 18, 338-351

133

A Genetic Dissection Approach to Functional Glycomics

Henrik Clausen[1]

1. Copenhagen Center for Glycomics, Depart. of Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

The structural diversity of the human glycome is vast and poses challenges for exploring biological functions and dissecting specific structure-function relationships. The true functional diversity of the glycome is, however, predicted to be lower when grouping structures with common or repeated features. The human glycome is orchestrated by over 200 distinct glycosyltransferase genes, and our knowledge of the properties of these and their roles in known glycosylation pathways in cells is relatively advanced. Currently some 170 glycosyltransferase genes can be assigned to rather specific roles in biosynthetic steps for the human glycome, although for many of the isoenzyme families our understanding of the unique functions of individual enzymes is limited. Current knowledge of assembly of the human glycome suggests that it is simpler to explore and dissect the glycome by a genetic entry point rather than from a structural entry point. Emergence of nuclease-based gene-editing tools enabling highly specific and facile knockout and site-directed knockin of glycosylation-related genes have led to wide use of the genetic approach to glycomics and new tools for the field. Our group has extensively adopted the genetic approach for dissection

and discovery of biological functions of protein glycosylation, and explored the options for custom design of glycosylation of recombinant glycoprotein therapeutics as well as using cell libraries for display of the glycome. An overview of these efforts will be presented.

134

Negotiating the labyrinth of O-glycopeptide analysis

Katalin F. Medzihradszky[1], Adam Pap[1, 2], Zsuzsanna Darula[1]

1. Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary

2. Doctoral School in Biology, Faculty of Science and Informatics, University of Szeged, Szeged, Hungary

In order to understand the biological role(s) of glycosylation one has to know which proteins are glycosylated and what is the degree of macro- and microheterogeneity. We are interested in mucin-type O-glycosylation, and we are attempting to study site-specific glycan diversity from 'wild-type' sources in a high throughput manner. Currently we are investigating human urine samples.

Human urinary glycopeptides were enriched from tryptic digests by lectin-affinity chromatography using wheat germ agglutinin. The resulting mixtures were subjected to LC/MS/MS analyses using HCD and diagnostic fragment ion-triggered EThcD experiments performed on a Fusion Lumos Tribrid mass spectrometer. Data interpretation was performed in an iterative manner. First, peptides modified with the most common O-glycans were identified from a full database search. Then additional glycoform candidates were lined up using an HCD-data filtering script searching for Y_0 and Y_1 fragments, and suggesting glycan compositions from the mass differences between the unmodified peptide and the precursor ions. Finally the proposed structures were manually validated or discarded using both HCD and EThcD data.

We have found that EThcD performed at minimal CE frequently revealed significant structural details about the glycan structures, even permitted the differentiation of some isomeric oligosaccharides. Our investigation showed that urinary mucin-type structures display a much wider diversity than originally suspected. For example, we identified glycans displaying blood-type antigens, and oligosaccharides featuring sialic acids at different states of O-acetylation, even in disialo units. Permitting so many different structures, plus considering variable covalent glycan modifications, could be counterproductive in 'normal' database searches. Our iterative approach provides a better tool for 'in-depth' data mining, although further software development is necessary.

This work was supported by the following grants: New Szechenyi Plan GOP-1.1.1-11-2012-0452, the Economic Development and Innovation Operative Programmes GINOP-2.3.2-15-2016-00001, and GINOP-2.3.2-15-2016-00020 from the Ministry for National Economy of Hungary.

135

Strategic applications of data dependent and targeted multistage fragmentation in dual modes to expedite mass spectrometry-based sequencing of glycopeptides

Kay-Hooi Khoo[1]

1. Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

While glycomics aims at mapping the entire range of glycan entities expressed at any one patho-physiological stage of a particular cell type, tissue or organism, glycoproteomics ventures a step further in locating their protein site-specific distribution. Although recent advances in mass spectrometry (MS) have enabled meaningful glycoproteomic undertakings, many technical limitations remain unsolved. Among these, ability to efficiently sequence the peptide backbone for de novo identification, delineating multiple N- and O-glycosylation sites on single glycopeptides and deriving more glycan structure information to discriminate isomeric glycoforms, are well acknowledged practical problems to be urgently tackled. We have been exploring all different kinds of experimental workflows for a most productive LC-MS/MS-based glycopeptide analysis, focusing particularly on the multiple glycosylated peptides and those carrying sulfated glycans. We have examined the complementarity of various fragmentation modes and most recently the judicious applications of negative ion mode sequencing to supplement a robust positive mode LC-MS2/MS3 workflow. We found that sulfation and sialylation drastically alter the fragmentation pattern in negative ion mode and the characteristic features identified can be utilized to program the most informative MS3. Moreover, facile elimination of the O-glycan moieties under MS2 affords an easy way to discover additional O-glycosylation on a glycopeptide that would otherwise not obvious. By high sensitivity detection of the characteristic glycan fragmentation ions produced by N- or O-glycans in combination with readily distinguishable elimination of the entire O-glycans in negative ion mode, we demonstrated in this work that unambiguous MS/MS sequencing of intact glycopeptides could be extended to those bearing both N- and O-glycans, or multiple O-glycans. Ability to advance from definitive identification of single to multiple site-specific glycosylation pattern on the same peptide backbones is anticipated to have a significant impact on the level of structural and biological insights one can gain in glycoproteomics applications.

136

Development of cancer biomarker for biliary tract cancer and pancreatic cancer with serum haptoglobin glycan analyses

Miyako Nakano[1], Soya Usui[1], Taiki Sugiyama[1], Shiro Takahashi[1], Eiji Miyoshi[2]

1. Unit of Biotechnology, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Hiroshima, JP

2. Department of Molecular Biochemistry & Clinical Investigation, Graduate School of Medicine, Osaka University, Suita, Osaka, JP

[Background] Biliary tract and pancreas are located in the deep part of the body. Therefore, it is difficult to find out genesis of cancer and is associated with poor prognosis. Biliary tract is pathway of bile, composed of bile ducts and gall bladder. We previously reported that fucosylated glycans of serum haptoglobin (Hpt) were significantly increased in serum samples of pancreatic cancer. In this study, we investigated whether this serum fucosylated Hpt is a cancer biomarker that can distinguish between bile duct cancer and gall bladder cancer and whether or not the increase of fucosylated glycans also occurred on cell membrane proteins in pancreatic cancer tissue.

[Methods] Hpt derived from serum of patients with biliary tract cancers was immunoprecipitated and subjected to SDS-PAGE and then transferred onto PVDF membrane. Cell membrane proteins derived from tumor section and their surrounding non-tumor section of pancreatic tissue was dotted onto PVDF membrane. *N*-Glycans were released from Hpt or cell membrane proteins by PNGaseF, and structural glycan analyses were performed by LC-ESI MS.

[Results] Lewis-fucosylated glycans on Hpt were significantly increased in samples of both gall bladder cancer and bile duct cancer. Core-fucosylated glycans on Hpt were significantly increased in sample of only bile duct cancer. *N*-Glycans containing ABO antigens were significantly decreased in tumor sections of pancreatic tissue, whereas they were observed as major components in tumor sections.

[Conclusion] We speculate that increases in Lewis-fucosylated Hpt are due to binding with hemoglobin in the bloodstream which is increased by genesis of cancer. Increases in core-fucosylated Hpt are due to choking of cancer cells in the bile duct, and bile flowing normally into duodenum flows back into bloodstream. We suggest the reduced activity of fucosyltransferase involving synthesis of H antigen results in decrease of ABO antigens in tumor section.

[Keywords] haptoglobin, cancer-biomarker

137

glyXbox^{CE}: A powerful tool in the glycoanalytical toolbox - improving biologics development and biomarker discovery for personalized diagnostics

Erdmann Rapp[1, 2], René Hennig[2], Samanta Čajić[1], Alexander Behne[2], Robert Kottler[2], Udo Reichl[1, 3]

1. Max-Planck-Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

2. glyXera GmbH, Magdeburg, Germany

3. Otto-von-Guericke University, Magdeburg, Germany

Glycomics is a rapidly emerging field that can be viewed as a complement to other „omics“ approaches. Hence, there is a dramatic increase in the demand for analytical tools and specific databases in glycobiology, respectively, glycobiotechnology. In order to enhance and improve the comparatively small existing glycoanalytical toolbox, automated, sensitive, reliable, high-throughput and high-resolution analysis methods including automated data evaluation are required. Our glycoanalysis approach, based on multiplexed capillary gelelectrophoresis with laser induced fluorescence detection (xCGE-LIF), shows high potential for high-performance analysis of glycoconjugates, as it allows fully automated, highly sensitive, instrument-, lab- and operator-independent "real" high-throughput glycoanalysis. This novel modular glycoanalysis system “glyXbox^{TM}” and its application to different fields with respect to sample preparation, separation and data analysis is presented. First, an optimized modular sample preparation workflow is presented with respect to performance and feasibility regarding high-throughput analytics. Second, parallel sample-measurement is shown to result in massive reduction of the effective run-time per sample. Third, automated data analysis with a newly developed modular software-tool “glyXtool^{CE}” for data processing and data analysis is demonstrated that involves integration of a corresponding oligosaccharide-database. Using this high-performance xCGE-LIF based glycoanalysis system, the generated “normalized” electropherograms of glycomoieties (“fingerprints”) can be evaluated on three levels: (1) “simple” qualitative and quantitative pattern comparison (“glycofingerprinting”), (2) identification of compounds in complex mixtures via database matching (“standard glycoprofiling”), and (3) extended structural analysis using exoglycosidase sequencing in combination with xCGE-LIF based glycoprofiling with repeated xCGE-LIF-based glycoprofiling (“extended glycoprofiling”), including linkage analysis. The broad applicability of the system is demonstrated for different types of glycosamples and different application fields - from biologics for biopharma to body fluids for personalized diagnostics.

1. Hennig, R., et al. (2016). “Towards personalized diagnostics via longitudinal study of the human plasma N-glycome.” Biochim Biophys Acta 1860: 1728-1738.
2. Thiesler, C.T., et al. (2016). “Glycomic Characterization of Induced Pluripotent Stem Cells Derived from a Patient Suffering from Phosphomannomutase 2 Congenital Disorder of Glycosylation (PMM2-CDG).” Mol Cell Proteomics 15: 1435-1452.
3. Hennig, R., et al. (2015). “N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF.” Methods Mol Biol. 1331: 123-143.
4. Huffman, J.E., et al. (2014). “Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research.” Mol Cell Proteomics 13: 1598-1610.

138

Isoform-specific functions of polypeptide GalNAc-transferases probed with glycoengineered human skin organoids and mass spectrometry

Ieva Bagdonaite[1], Emil MH Pallesen[1], Zilu Ye[1], Sergey Y Vakhrushev[1], Irina N Marinova[1], Mathias I Nielsen[1], Signe H Kramer[2], Stine F Pedersen[2], Hiren J Joshi[1], Eric P Bennett[3], Sally Dabelsteen[3], Hans H Wandall[1]

1. Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen N, Denmark

2. Cell Biology and Physiology, Department of Science, University of Copenhagen, Copenhagen N, Denmark

3. Department of Odontology, University of Copenhagen, Copenhagen N, Denmark

Post-translational modifications (PTMs) greatly expand the function and potential for regulation of protein activity, and O-glycosylation is among the most abundant and diverse PTMs. Initiation of O-GalNAc glycosylation is regulated by 20 distinct GalNAc-transferases (GalNAc-Ts), and deficiencies in individual GalNAc-Ts are associated with human disease, causing subtle but distinct phenotypes in model organisms. Here, we generated a set of isogenic keratinocyte cell lines with and without the three dominant and differentially expressed GalNAc-Ts. Through the ability of keratinocytes to form epithelia, we investigated the phenotypic consequences of the loss of individual GalNAc-Ts. We also used isogenic keratinocyte cell lines to probe cellular responses to the ablation of GalNAc-Ts through global transcriptomic, differential glycoproteomic, and differential phosphoproteomic analyses. We demonstrate that loss of individual GalNAc-T isoforms causes distinct epithelial phenotypes through their effect on distinct biological pathways; GalNAc-T1 targets are associated with components of the endomembrane system, GalNAc-T2 targets – with cell-ECM adhesion, and GalNAc-T3 targets – with epithelial differentiation. Thus, GalNAc-T isoforms serve specific roles during human epithelial tissue formation.

1. Radhakrishnan P, Dabelsteen S, et al. Immature truncated O-glycophenotype of cancer directly induces oncogenic features. (2014) Proc Natl Acad Sci USA 111(39):E4066-75.
2. Lavrsen, K, et al. De novo expression of human polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6) in colon adenocarcinoma inhibits the differentiation of colonic epithelium. (2018) The Journal of biological chemistry 293, 1298-1314.
3. Schjoldager KT et al. Deconstruction of O-glycosylation--GalNAc-T isoforms direct distinct subsets of the O-glycoproteome. (2015) EMBO Rep 16: 1713-22.

139

Proteome-Based Diagnostics: The Next Revolution in Medicine and Pathology

Michael H. Roehrl[1]

1. Memorial Sloan Kettering Cancer Center, New York, NEW YORK, United States

The application of deep quantitative proteomics to diagnosis, risk stratification, treatment guidance, and therapeutic resistance surveillance will revolutionize how we practice medicine. It has become clear that nucleic acid-based sequencing approaches alone are insufficient to capture the dynamic functional biochemical aberrations that characterize human disease. Importantly, essentially all current drug targets are proteins, and thus a precision pathology approach to proteomics is urgently needed.

In the first part of my talk, I will show examples of how we apply deep mass spec-based proteomics together with genomics and transcriptomics to characterize gastrointestinal (GI) tumors (such as colon cancer), to develop new prognostic and treatment-predictive biomarkers, and to discover new protein drug targets for these diseases.

Understanding the biology of the metastatic process will be key for advancing effective therapies for lethal cancers that are frequently only diagnosed after metastasizing, such as colonic adenocarcinoma. Availability of matched normal-primary-metastasis sample triples from patients is often rate-limiting in research, and thus unlocking archival FFPE (formalin-fixed paraffin-embedded) samples for deep proteome profiling will open unmatched resources for biomarker discovery. I will describe a robust deep proteome profiling method for FFPE and illustrate how we applied it to proteomic biomarker discovery in primary and matched metastatic colonic adenocarcinomas.

In the second part of my talk, I will introduce our work using proteomics to define the human autoantigen-ome, i.e., defining the totality of human proteins that can become autoantigens and elicit a humoral antibody response. We have developed an approach that uses a patient's tissue samples jointly with the same patient's serum antibodies to catalogue the patient-specific autoantigen-ome in both health and disease. We have applied this approach to patients with autoimmune disorders (such as lupus) but also to cancer patients. Our novel approach promises to become a powerful diagnostic and dynamic monitoring tool for such patients and offers tremendous potential for understanding the response to novel immuno-oncology checkpoint inhibitors.

The development of Proteome-Based Precision Diagnostics and Theranostics in Pathology will revolutionize how diseases are analyzed, classified, treated, and monitored.

140

Enabling Precision Medicine for Alzheimer's Disease through Biofluid-Based Biomarkers

Danni Li[1]

1. University of Minnesota, Minneapolis, MN, United States

Alzheimer's disease (AD) is a public health crisis. Currently, no effective strategies to prevent or slow AD exists. This is largely due to lack of biological definition of AD as well as lack of complete understanding of the mechanisms that contribute to AD pathophysiology. This presentation will highlight how biofluids-based biomarkers have great potential to address these critical issues and to enable precision medicine for AD prevention and treatment.

141

Heterogeneity study of pancreatic tumor tissue: Proteomic characterization of tumor cells obtained by laser micro-dissection (LMD), coring, and bulk sampling techniques.

Michael Schnaubelt[1], Lijun Chen[1], David Clark[1], Yingwei Hu[1], Jianbo Pan[1], Daniel W Chan[1], Hui Zhang[1], Qing Kay Li[1]

1. the Johns Hopkins Medical Institutions, Baltimore, MARYLAND, United States

Background: Tumor tissue represents a mixture of tumor cells, stroma and other cellular components. To minimize potential effect of tumor heterogeneity during proteomic analysis, several approaches such as LMD and coring technique have been used to selectively isolate tumor cells from other cellular components. Although these techniques can provide a relatively pure quantity of tumor specimens, the potential effect of sampling techniques on proteomic analysis is still largely unknown. In this study, we analyzed proteomic signature of tumor cells obtained by different sampling techniques from pancreatic cancer (PCA).

Method: Four pairs of PCA and tumor-matched normal pancreatic tissue, and four non-paired PCA samples were included. Tumor cells were harvested using three techniques (LMD, coring and bulk section). Peptides were separated on a Dionex Ultimate 3000 RSLC nano system (Thermo Scientific). Data was analyzed by the MS-PyCloud proteomics pipeline, using the MS-GF+ search engine to search against a concatenated target-decoy database.

Findings: A median number of 7388, 8202, 8213 proteins were identified in LMD, coring and bulk specimens, respectively. However, differential proteomic profiles were identified among samples. Differentially expressed proteins in tumors vs matched normal from both coring and bulk samples had a similar profile, and the upregulated-tumor proteins were more consistent with TCGA mRNA expression. Protein profiles, particularly the phosphoproteome in LMD samples, revealed a different pattern by clustering analysis. For top 150 upregulated-tumor proteins based on TCGA mRNA expression, we found 25 (17%), 75 (50%), and 51 (33%) proteins were upregulated in LMD, coring, and bulk samples.

Conclusion: Heterogeneity of tumor tissue is an important issue for proteomics research. Tumor cells can be isolated using LMD and coring technique with high purity. However, the data from LMD samples demonstrates a unique proteome profile, and with a fewer protein identification number.

142

Intra- and inter-individual variation in the proteome of high-grade serous ovarian cancer

Maiken M Espersen[1], Srikanth Manda[2], Rohan Shah[2], Steve Williams[2], Natasha Lucas[2], Dylan Xavier[2], Sadia Mahboob[2], Andrew Robinson[2], Peter G Hains[2], Brett Tully[2], Roger Reddel[2], Phil J Robinson[2], Qing Zhong[2], Anna deFazio[1], Rosemary Balleine[2]

1. Centre for Cancer Research, Westmead Institute for Medical Research, The University of Sydney, Westmead, NSW, Australia

2. ProCan®, Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

High-grade serous ovarian cancer (HGSOC) is a disease with persistently poor survival rates. Targeted therapeutics are beginning to show improved outcomes for selected patients. There is an urgent need for predictive biomarkers that can be reliably measured in small tissue samples. The aim of this study was to characterize variation in the HGSOC proteome.

Proteomic profiles from 447 frozen samples, taken from primary (n=11) and matched metastatic (n=10) tumour tissues were generated by DIA on SCIEX 6600 triple TOF instruments (7-49 samples per tissue). Data were analysed with OpenSWATH followed by Diffacto to compile a protein matrix including over 3,000 quantified proteins.

The data were highly structured with closest similarity between samples from the same tissue, followed by similarity between primary and metastatic tissues from the same individual. Around 15% of proteins were detected in all samples, with ribosomal proteins being the top ranked category, consistent with their housekeeping role. There were ~20 proteins absent from all samples in some patients, whilst uniformly present in all samples of others. This pattern of stable intra- / variable inter-individual expression demonstrates features that could form a basis for robust sub-classification.

The HGSOC proteome is highly individualized with marked spatial variation. Global proteomic profiling can facilitate biomarker discovery by screening for discriminative features associated with treatment response and clinical outcome.

**MM Espersen and S Manda contributed equally.*

143

Proteomic Tissue Biomarkers for Early Prediction of Prostate Cancer Progression

Sudhir Srivastava[1], Jacob Kagan[1]

1. Cancer Biomarkers Research Group, National Cancer Institute, National Institutes of Health, GAITHERSBURG, MARYLAND, United States

Although approximately 40% of the screen-detected prostate cancers are indolent, advanced-stage prostate cancer is a lethal disease with 5-year survival rates around 29%. The challenge is to identify biomarkers for early detection of aggressive disease, when the cancer is still organ-confined. We have developed ultra-sensitive, high-pressure, high-resolution separations coupled with intelligent selection and multiplexing-selected reaction monitoring (PRISM-SRM) assays for 52 protein markers selected from existing prostate cancer genomics data sets and known prostate cancer drivers. These PRISM-SRM assays were applied in analysis of organ-confined primary tumors from prostate cancer patients (n = 338) presenting different post-surgery features: distant metastasis, biochemical recurrence (BCR), and no progression after more than ten years of follow-up after radical

prostatectomy. Several prostate differentiation/androgen receptor signaling related proteins (FOLH1, PSA and NCOA) and tumor progression-related proteins (TGFB1, CCND1 and SPARC) had significantly different expression levels between the three groups. Combining the protein biomarker panel with existing clinical based models achieved an area under the receiver operating characteristic curve (AUC) of 0.88 and a negative predictive value (NPV) of 0.91 for predicting distant metastasis. The molecular test of these protein biomarkers therefore provides a complementary tool for early detection of aggressive prostate cancer, as well as better selection of patients with indolent and low-risk cancers for active surveillance.

144

Oxygen-sensitive interactions between glycolytic enzymes and a cancer-testis antigen established signaling scaffold are regulated by lysine acetylation

Soren Naaby Hansen[1], Arabinda Mandal[2], Allan Stensballe[1]

1. Aalborg University Hospital, Nørresundby, DENMARK, Denmark

2. University of Virginia, Virginia, USA

Reactivation of the male gametogenic expression program is tightly associated with the most malignant and metastasis-prone tumours and the emergence of aggressive subclones of tumour cells, which are highly resistant to stress-induced apoptosis. While the cancer-testis antigens (CTAs) CABYR and AKAP3/4 roles during gamatogenesis and their importance for flagellar movement have gradually emerged, their function in cancer cells have remained obscure.
In this study we combine immunoprecipitation (IP), mass spectrometry (MS) and western blot (WB) analysis to unravel their functional roles in therapy resistant lung and ovary adenocarcinoma cells by identifying their interaction partners. CABYR variants were shown to oligomerize and interact with AKAP proteins to generate a HMW signal scaffold structure, which was found to bind several glycolytic enzymes and signal transducers. Forward & reverse IP experiments followed by WB confirmed interactions between CABYR and LDH, ALDO, PFK, TPI-1, GAPDH, ENO-1 and GSK3b.
Transistion from normoxic to hypoxic growth conditions disrupted the associations between glycolytic enzymes and the CABYR-AKAP signaling scaffold in the cancer cells, leading to a 3.2-fold increase in their production and secretion of lactic acid. Hypoxic growth conditions resulted in increased acetylation of lysine residues in both CTAs and triggered deacetylation of lysines in LDH and aldolase. Treatment with resveratrol prevented hypoxia-induced dissociations, suggesting that the regulation of oxygen-sensitive protein interactions within the CABYR-AKAP-Glycolysome complex involve changes in the acetylation of lysines in the engaged proteins.
MS analysis of IPs finally revealed interactions between CABYR and proteins associated with the cancer cells contractile cytoskeleton.
Based on these findings it is tempting to speculate that hypoxia-induced release and subsequent local activation of glycolysomes from cytoskeleton-associated CABYR-AKAP scaffold structures might be instrumental for cancer cells ability to maintain a steady energy supply to their contractile cytoskeleton and thereby sustain their migratory and invasive capability despite encountering severe reductions in environmental oxygen levels.

145

Proteomic dissection of the plant immune system

Hirofumi Nakagami[1]

1. Max Planck Institute for Plant Breeding Research, Cologne, Germany

The infection of plants by pathogens is causing a considerable economic loss to agricultural industries. Recent studies estimated that the global yield loss due to biotic stresses averages over 23 percent of the estimated attainable yield across major cereals. An understanding of the fundamental mechanisms underlying disease resistance processes is expected to provide novel ideas for improving the current situation.
During the last several decades, extensive analyses revealed that plants utilize a two-branched immune system for defence against pathogens. In the first branch, transmembrane pattern recognition receptors (PRR) are used to recognize and respond to slowly evolving pathogen-associated molecular patterns (PAMP). In the second branch, either a direct or an indirect recognition of the pathogen through disease-resistance (R) proteins is used for response to pathogen virulence factors (effectors). The R-gene mediated resistance has been widely used in crop improvement. However, the resistance is pathogen race-specific and, frequently, a rapid breakdown of the resistance has been reported. Consequently, alternative approaches to provide a broad-spectrum, durable resistance in crops are highly desirable. Meanwhile, several studies have implied that manipulation of the PRR-mediated immune system can be a smart strategy to enhance the broad-spectrum resistance.
While extensive genetic screens successfully identified a number of PRRs and components which affect abundance and maturation of PRRs, signal transduction mechanisms that lead to defence responses is thus far limited. This partly stems from limitations of forward genetics caused by lethality and/or genetic redundancy. Accordingly, we have been taking proteomic approaches to understand the basic framework of the PRR-mediated immune system.

146

Quantitative analysis of protein synthesis and degradation rates in plants using progressive labeling with stable isotopes

Lei Li[1], Owen Duncan[1], Karzan J Silah[1], Hui Cao[1], Harvey Millar[1]

1. ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA, Australia

We are using progressive stable isotope labeling in plant systems to provide a bird's-eye view of the activity of the translation machinery and the proteolysis network as they maintain and sculpt the proteome[1]. Using peptide mass spectrometry, the progressive labeling of new peptides and the decrease in the abundance of peptides with natural isotope profiles enable the degradation and synthesis rate of specific root, leaf and seed proteins to be separated and quantified[2]. This allows new insights in selective proteolysis of proteins in vivo for different applications in fundamental and applied plant science. New evidence of changes in turnover rate of specific leaf and root proteins in autophagy mutants will be presented. This shows the selective effect of pathways in autophagy on the fate of organelle types and on biochemical functions in leaves and roots. New data on differential protein synthesis and degradation rates amongst cytosolic 80S ribosome subunits in the model plant Arabidopsis and between storage proteins and enzymes during grain filling in wheat will also be presented. These datasets are revealing the maintenance of protein complexes, the selectivity of autophagy, and allowing the cost of synthesis and degradation of particular cellular structures and processes in plants to be calculated.

1. Nelson CJ and Millar AH (2015) Protein turnover in plant biology. Nature Plants 1, 15017
2. Li L, Nelson CJ, Trösch J, Castleden I, Huang S, Millar AH. (2017) Protein Degradation Rate in Arabidopsis thaliana Leaf Growth and Development. Plant Cell. 29:207-228

148

Na_2CO_3-responsive mechanism in alkaligrass revealed from redox proteomic analysis

Shaojun Dai[1, 2], Ying Li[2], Yongxue Zhang[1, 2]

1. College of Life Sciences, Shanghai Normal University, Shanghai, 中国

2. Alkali Soil Natural Environmental Science Center, Northeast Forestry University, Harbin, Heilongjiang, 中国

Soil alkalization is one of the majors threatens that affect plant production and ecological environment. Saline-alkali stress causes oxidative stress on plants, and seriously affect plant growth and development. Halophyte alkaligrass (*Puccinellia tenuiflora*) is an outstanding pasture with strong tolerance to salt and alkali stresses. In this study, the molecular mechanisms of Na_2CO_3 response in alkaligrass were analyzed using redox proteomics and molecular genetics approaches. Besides, the activities of enzymes in reactive oxygen species (ROS) scavenging system were detected in alkaligrass leaves and roots. The abundances patterns and redox states of Na_2CO_3-responsive proteins indicated that Ca^{2+}-mediated kinase signaling, membrane trafficking, cytoskeletal remodeling, ROS homeostasis, stress and defense, gene expression and protein turnover, carbohydrate metabolism, amino acid metabolism were pivotal to Na_2CO_3 adaptation in alkaligrass.

Moreover, two proteins were proved to be crucial in alkaligrass under the Na_2CO_3 stress. The developmentally regulatory plasma membrane polypeptide protein (DREPP) were involved in the regulation of binding phosphatidylinositol phosphate (PtdInsPs) and regulation of the cytoskeletal structure. Overexpression of *PutDREPP* in wild-type Arabidopsis enhanced the root resistance to saline-alkali stresses. And the expression of *PutDREPP* were increased significantly in response to Na_2CO_3, NaCl, $NaHCO_3$ and H_2O_2 stresses. Besides, cysteine-rich repeat secretory protein 55 (CRRSP55) modulates cell-to-cell trafficking, and it has the conserved active motif C-X8-C-X2-C of RLK. The Cys78 and Cys214 of PutCRRSP55 was oxidized under Na_2CO_3 stress, which would affect PutCRRSP55 protein function. The expression of *PutCRRSP55* were increased significantly in alkaligrass in response to Na_2CO_3, NaCl, and H_2O_2 stresses. This provides new clues for further study on the regulation network of plant saline-alkali stress response, and also provide important information for the breeding of salt-tolerant crops.

149

Global profiling of dehydration-induced mitochondrial dynamics and defense response in rice

Niranjan Chakraborty[1], Dipak Gayen[1], Pragya Barua[1], Nilesh Vikram Lande[1], Subhra Chakraborty[1]

1. National Institute of Plant Genome Research, New Delhi, DELHI, India

Background

Elucidation of molecular basis of dehydration-induced responses aids understanding of plant adaptation and has direct implications towards fortification of sustainable agriculture. Previously, we successfully established dehydration-responsive proteome map of nucleus (1) and cell wall (2) in rice. Next we focused on mitochondria, a dynamic microenvironment for integration of cellular metabolism and signalling towards understanding energy metabolism under water-deficit conditions.

Methodologies

Four-week-old rice seedlings were subjected to progressive dehydration, and integrity and purity of mitochondrial fraction was evaluated. Protein identification was carried out at 1% FDR. Statistical analyses were performed on the data with a one-way analysis of variance (ANOVA) and Duncan's Multiple Range test (DMRT) using GraphPad Prism V 7.00. Gene regulation in the native system was carried out under the control of dehydration-responsive rd-29 promoter.

Findings

A critical screening of mitochondrial dehydration-responsive proteins revealed the presence of a DUF2488 or YCF54 domain containing protein, reported so far in the chloroplasts of photosynthetic eukaryotes. Stress-induced accumulation of YCF54 protein in mitochondria was confirmed by immunoblot analysis. Proteome-scale interactome networks coupled with yeast two-hybrid screening and co-immunoprecipitation analysis identified reducing enzyme, peroxiredoxin, as an interacting partner. To further characterize its function *in vivo*, we generated transgenic rice overexpressing the sensor under the control of rd-29 promoter. The overexpressing seedlings displayed enhanced tolerance to oxidative stress, possibly through peroxiredoxin-

modulated ROS detoxification. Stress-induced overexpression of the sensor helped maintain mitochondrial respiration and ATP production. To confirm that the increased ROS detoxification is associated with mitochondrial redox homeostasis, we overexpressed both the interacting partners in non-photosynthetic eukaryote *Saccharomyces cerevisiae*. Interestingly, co-expression of the interacting partners showed accelerated ROS catabolism.

Concluding statement

The present study ascertained a comprehensive view of effects of dehydration on mitochondria of rice. This study suggests that OsDUF1 plays a key role in mitochondrial redox metabolism, and confers stress tolerance.

1. 1. Choudhary, M. K., Basu, D., Datta, A., Chakraborty, N., & Chakraborty, S. (2009) Dehydration-responsive nuclear proteome of rice (Oryza sativa L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. Mol. Cell. Proteomics. 10.1074: 1579-1598.
2. 2. Pandey, A., Rajamani, U., Verma, J., Subba, P., Chakraborty, N., Datta, A., Chakraborty S., & Chakraborty, N. (2010) Identification of extracellular matrix proteins of rice (Oryza sativa L.) involved in dehydration-responsive network: a proteomic approach. J. Proteome Res. 10.1021: 3443-3464.

147

Systems Biology Analysis of Root Tips: Towards Bioengineering of Rice Root Structure to Enhance Drought Stress Tolerance

Somayeh Abdirad[1], Zahra Ghorbanzadeh[1], Armin Soleimaninia[2], Mohsen Rahiminejad[1], Yunqi Wu[3], Zahra_Sadat Shobbar[1], Mohammad Reza Ghaffari[1], Brian Atwell[3], Mehdi Mirzaei[3], Paul A Haynes[3], Ghasem Hosseini Salekdeh[1]

1. Department fo Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran, Karaj, Iran

2. Department of Biotechnology, University of Tehran, Tehran, Iran

3. Department of Molecular Sciences, Macquarie University, Sydney, Australia

Rice (*Oryza sativa*) is a staple food for nearly half of the world's population. It is susceptible to drought stress due to its shallow rooting relative to other cereal crops. Roots are the primary organs that perceive changes in soil, and hence play a vital role in the response of plants to drought. As an essential strategy to minimize the negative impact of drought stress, the manipulation of root structure towards wider and deeper distribution of roots in the soil may enable plants to avoid drought-induced stress by extracting water (and nutrients) from deep soil layers. The most important part of roots are the root tips, which encompass the root cap, apical meristem and elongation zones, as these regions determine the fate of root length, root diameter and root angle.

In this project, we are aiming to modify the root architecture of a commercial rice variety to transform shallow rooting plants into deeper rooting plants, while retaining commercially desirable characteristics such as grain yield. The project is built on three pillars: phenotyping, gene and protein discovery, and functional analysis. We profiled the proteome, transcriptome and epigenome of three zones of root tips from two different rice genotypes with contrasting root architecture phenotypes; a lowland rice with shallow roots (IR64) and an upland rice with deeper roots (Azucena), grown under control and water deficit stress conditions. Many of the expressed molecules exhibited zone/genotype/water regime-specific differential expression patterns. We also observed substantial differences in expression of isoforms and lncRNAs in different zones. This systems biology approach resulted in a short list of candidate genes for further characterisation. We are examining the function of these root structure associated genes by transferring them into shallow rooted plants, which will be screened for root phenotype, drought tolerance, and nutrient uptake.

150

Wheat pan-proteomics: Unifying data-independent LC-MS proteome measurements across diverse genetic backgrounds for trait prediction

James A. Broadbent[1], Sally Stockwell[1], Keren Byrne[1], Utpal Bose[1], Jessica Hyles[2], Kerrie Ramm[2], Ben Trevaskis[2], Shannon Dillon[2], Michelle Colgrave[1]

1. Agriculture and Food, CSIRO, St Lucia, Queensland, Australia

2. Agriculture and Food, CSIRO, Black Mountain, Australian Capital Territory, Australia

Background

Wheat is a major global commodity. With more than a decade of stalled commercial yield in Australia, new approaches are vital to improve crop productivity. In this regard, the prediction of time-to-flowering is a valuable measurement for optimising farm resource allocation and yield. Pan-transcriptome analysis has shown promise for flowering time prediction across diverse wheat varieties. Herein we demonstrate the utility of pan-proteomics for predicting flowering time in mature plants using sample extracted from plants at the two-leaf stage.

Methodologies

A diverse panel of Australian wheat varieties were grown under short (8h) or long day length (16h) conditions to facilitate variation in flowering time. A total of 632 unique wheat samples were processed and measured by variable window SWATH acquisition along with control measurements. Peptide responses were extracted from raw SWATH data along with peak group false discovery

rate estimates. These data were processed and analysed using a suite of scripts within the R environment. The resulting matrix of pan-wheat measurements was subject to multivariate and machine learning analyses to assess variance and quantify the ability of protein abundances to predict flowering time.

Findings

We show that the major proteome variation can be readily attributed to day length using t-Distributed Stochastic Neighbour Embedding machine learning. We also quantify the ability for pan-proteome measurements from plants at the two-leaf stage to predict the flowering time of the mature plant through application of random forest analysis.

Concluding

Wheat is a substantial source of global nutrition and economic benefit. With the growing population and coincidental requirement for nutrition from cereals projected to increase by 50% over the next two decades, efficiency gains in grain production are required. Herein we demonstrate the ability to predict wheat traits using pan-proteome measurements that can inform on-farm practices aimed at improving crop quality and yield.

151

Single ion mass spectrometry to measure proteoforms and their complexes with complete molecular specificity

Jared Kafader[1], Rafa Melani[1], Kenneth Durbin[1], Bon C Ikwuagwu[1], Bryan Early[1], Ryan Fellers[1], Steven Beu[2], Vlad Zabrouskov[3], Alexander Makarov[3], Joshua Maze[3], Deven Shinholt[3], Ping Yip[4], Danielle Tullman-Ercek[1], Mike Senko[3], Compton Phil[1], Neil Kelleher[1]

1. Northwestern University, Evanston, ILLINOIS, United States

2. S C Beu Consulting, Austin, TX, US

3. Thermo Fisher, San Jose, CA, US

4. Neuroscience, Queen Mary Univ. of London, London

A new Orbitrap-based single ion analysis procedure is shown to be possible by determining the direct charge on numerous measurements of individual protein ions to generate true mass spectra (i.e. in the mass, not *m/z* domain). The deployment of a robust, Orbitrap-based system for high resolution charge detection enables the characterization of highly complicated mixtures of proteoforms and their complexes in both denatured and native modes of operation, revealing information not obtainable by traditional measurement of an ensemble of ions

152

Novel Strategies in Top-Down Proteomics

Ying Ge[1]

1. UW-Madison, Madison, WI, United States

Top-down mass spectrometry (MS)-based proteomics is arguably the most powerful method to comprehensively characterize proteoforms that arise from genetic variations, alternative splicing, and post-translational modifications (PTMs), but myriad challenges remain. We have been developing novel strategies to address the challenges in top-down proteomics in a multi-pronged approach. To address the protein solubility challenge, we have recently identified a photo-cleavable anionic surfactant ("Azo") that can be rapidly degraded upon UV irradiation, for top-down proteomics. Azo is MS-compatible and can effectively solubilize proteins with performance comparable to SDS. Importantly, Azo-aided top-down proteomics enables the solubilization of membrane proteins for comprehensive characterization of PTMs. To address the proteome complexity challenge, we have been developing new chromatography materials and novel strategies for multi-dimensional liquid chromatography (MDLC) to separate intact proteins. We developed novel hydrophobic interaction chromatography (HIC) materials for high-resolution separation of intact proteins under non-denaturing mode and demonstrated the potential of online HIC/MS for top-down proteomics. Given the difficulty in detecting large proteins in top-down MS, importantly, we developed a novel serial size exclusion chromatography (sSEC) strategy for size-based protein separation that can be coupled with online reverse phase chromatography (RPC) and high-resolution MS which enabled the top–down MS analysis of large proteins (>200 kDa). Furthermore, we established a robust top-down LC/MS-based targeted proteomics platform for quantification of protein expression and PTMs concurrently in complex mixtures with high throughput and high reproducibility. To address the proteome dynamic range, we have been developing novel nanomaterials that can bind low abundance proteins and PTMs with high specificity. First, we designed and synthesized novel superparamagnetic nanoparticles (NPs) for capturing phosphoproteins globally out of the human proteome with high specificity. Subsequently, we developed an intergrated top-down phosphoproteomics work flow that coupled NP-based phosphoprotein enrichment by functionalized NPs with online top-down LC/MS/MS to enrich, identify, quantify, and characterize intact phosphoproteins directly from cell lysates and tissue homogenates. To address the challenge in under-developed software, we developed MASH Suite Pro, a comprehensive software tool for top-down proteomics including protein identification, quantitation, and characterization with visual validation and versatile user-friendly interface.

157

Cellular atlas of the transcriptome and proteome

Kathryn Lilley[1]

1. University of Cambridge, Cambridge, CAMBRIDGESHIRE, United Kingdom

The complexity of living organisms does not scale with the predicted number of protein coding genes. Many factors contribute to increasing complexity, including non-coding RNA mediated control mechanisms and post-transcriptional and post-translation processing. The location of protein synthesis also plays a key role in expanding protein functionality, with aberrant spatial translation being a driver in multiple diseases.

In order understand cellular physiology, a thorough understanding of the spatial relationship of the transcriptome, translatome and proteome is required. Several existing methodologies are able to capture the transcriptome and proteome at specific sub-cellular locations. Holistic approaches, however, are required to construct cell-wide models that can give insight into the multi-purposing of components that leads to the expansion of cellular functions.

I will describe new approaches to capture the spatial relationship between RNA and protein on a cell-wide scale. I will discuss methods designed to map the cellular spatial proteome (1) (2) based on physicochemical fractionation of cellular components (LOPIT), that also give insights into the effect of post-translational modification on protein location. I will describe the tools we have developed to robustly capture dynamic re-localization of proteins utilising Bayesian analysis (3) to identify dynamic changes in the spatial and temporal proteome and uncover different types of translocation events. I will also discuss how we have significantly modified these approaches in order to capture the spatial transcriptome (LoRNA– localization of RNA). I will introduce the orthogonal organic phase separation (OOPS) protocol that recovers both RNA and protein from cross-linked RNA-protein complexes in an unbiased manner independent of polyadenylation status of RNA (4). Finally, I will demonstrate that when applied in concert, these approaches reveal the spatial interplay of the proteome and transcriptome on a cell–wide scale producing three over-lapping maps:

1. Protein map (LOPIT)
2. RNA binding protein (RBP) map
3. Total RNA map (LoRNA)

I will describe some unexpected findings from these maps including the RNA binding capacity of many metabolic enzymes and therapeutic targets and the steady state location of mRNA species that code for different protein families.

1. Mulvey et al (2017) Nature Protocols - doi: 10.1038/nprot.2017.026
2. Geladaki et al (2019) Nature Communications - doi:10.1038/s41467-018-08191-w

153

Extending the upper mass range for top-down proteomics with proton-transfer reactions, parallel ion parking, and 21 tesla FT-ICR MS

Lissa C. Anderson[1], **Chad R. Weisbrod**[1], **David Butcher**[1], **Zeljka Popovic**[2], **Alan G. Marshall**[1, 2], **Greg T. Blakney**[1], **Jeffrey Shabanowitz**[3], **Donald F. Hunt**[3, 4], **Christopher L. Hendrickson**[1, 2]

1. National High Magnetic Field Laboratory, Tallahassee, FL, United States

2. Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL, 32306

3. Department of Chemistry, University of Virginia, Charlottesville, VA, 22904

4. Department of Pathology, University of Virginina, Charlottesville, VA, USA

Top-down proteomics is routinely complicated by several analytical challenges including the inherent distribution of protein ion signal among multiple charge states, isotopologues, adducts, modification states, and fragment ion channels. Such challenges are further compounded by poor chromatography, incomplete fragmentation/sequence coverage, slow spectral acquisition rate, and high data complexity. These effects limit proteoform detection and characterization, and grow exponentially worse as mass increases, making routine analysis of proteins larger than ~30 kDa difficult.

We will provide performance benchmarks for integration of proton-transfer reactions (PTR), parallel ion parking (PIP), and 21 T FT-ICR MS for analyses of proteins >30 kDa. Use of PTR coupled with PIP concentrates the majority of ion current from multiple charge states of each precursor proteoform into just a few charge states, which increases sensitivity compared to conventional MS1 spectra. For example, isotopic distributions of patient-derived human serum albumin proteoforms (66-69 kDa) are resolved without signal averaging in an on-line LC-PTR/PIP-MS experiment. Signal-to-noise ratio of the most abundant charge state in the PTR/PIP spectra was improved by a factor of 15 compared to corresponding conventional MS1 spectra. Preliminary results from LC-MS/MS experiments with human whole cell lysate demonstrate significant improvement in the number of proteoforms detected from 1404 (conventional MS1) to 2472 (PTR/PIP). We will also demonstrate the utility of ETD in tandem with PTR for sequence analysis of large proteins. ETD/PTR MS/MS reduces fragment ion overlap, enabling more comprehensive sequence analysis compared to ETD alone. Sequence coverage of recombinant protein AG (50 kDa) was improved from 12% with ETD alone versus 50% with ETD/PTR. These techniques, combined with the high charge capacity, resolving power, and mass measurement accuracy achieved at 21 T, facilitate an enhanced ability to observe, identify, and characterize higher molecular weight proteins in biological samples.

154

Identification of the functional status of proteoforms and their interactomes in blood plasma

Laura Heikaus[1], **Marcus Wurlitzer**[1], **Thomas Renné**[1], **Christoph Krisp**[1], **Hartmut Schlüter**[1]

1. Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Analysis of proteoforms is the next challenging level in proteomics (1). Identification of proteoforms in relationship to their function is demanding. Genes coding proteases usually have different proteoforms, including **inactive proteoforms**, which in many cases are converted by a proteolytic truncation into **active proteoform**. Active proteases are controlled by protease inhibitors. Some of

these inhibitors belonging to the suicide inhibitors are binding covalently to the active proteoform, thereby forming a new **inhibited proteoform** - a fusion protein consisting of the protease and the inhibitor.

We have developed a method by which proteoforms, coded by a single gene, with different functional status can be identified (2). Therefore, the sample containing different proteofoms is separated with respect to the molecular weights of the proteins. The proteins of each of the resulting fractions are digested with trypsin. The tryptic peptides from each of the different fractions are analyzed by LC-MSMS, the peptides identified by a search engine, the relative amounts calculated by the area under the curves (AUCs) of their extracted ion chromatograms (EICs) and these values plotted against the molecular weight of the fractions. Thereby, proteoforms coded by the same gene but with different molecular weight can be identified. We applied this approach for analyzing different functional status of proteoforms of factor XII (FXII) in human blood plasma, including the inactive precursor, the activated proteoform generated by proteolysis, and several proteoforms inhibited by different inhibitors like the C1-inhibitor (C1-Inh). Furthermore this approach is giving access to the identification of substrates and their proteoforms (here: e.g. kallikrein) of the target protease as well as down-stream proteases and substrates (here: e.g. kininogen) of proteolyse products.

In summary, we have developed a tool, by which the functional status of protease proteoforms and their interactomes as well as corresponding protease cascades can be investigated.

1. How many human proteoforms are there? Aebersold R, et al., Kelleher NL, et al., Schlüter H, et al. Nature Chem Biol. 2018 Feb 14;14(3):206-214.
2. Homogenization of tissues via picosecond-infrared laser (PIRL) ablation: Giving a closer view on the in-vivo composition of protein species as compared to mechanical homogenization. Kwiatkowski M, et al., Schlüter H. J Proteomics. 2016 Feb 16;134:193-202.

155

Application of a novel electron capture dissociation (ECD) mass spectrometry to top-down and bottom-up proteoform characterization on a QTOF.

Blaine R Roberts[1], Eugene Kapp[1], Soumya Mukherjee[1], Yury V Vasil'ev[2], Valery G Voinov[2], Joseph S Beckman[2]

1. Florey Institute of Neuroscience and Mental Health, Parkville, VIC, Australia

2. eMSion, Corvallis, Oregon, USA

Bottom-up approaches using collision induced dissociation for characterizing target proteins yield incomplete information, particularly concerning the colocalization of post translational modifications. Here we described the implementation of a device that yields efficient ECD of proteins that can be reversibly retrofitted into Q-ToFs without diminishing performance. The ECD device does not require trapping ions as needed for ETD and thus is compatible linear designed instruments. We demonstrate nearly complete sequence coverage of “native”-folded proteins such as the $5^+$and 6^+charge states of ubiquitin and Cu, Zn superoxide dismutase. Due to the high sequence covnerage (>80%) the localization of non-covalent cofactors like Cu and Zn could be determined from the top-down spectra of SOD. We also benchmarked the molecular weight range accessible for top-down protoemics on our QTOF system. Sequence coverage of 80-95% was obtained for small proteins like ubiquitin, amyloid beta and alpha-synuclein (14 kDa). Sequence coverage was 93% for carbonic anhydrase (29kDa) and similarly for green fluorescent protein (27kDa). Half of the human proteome is smaller than 30kDa making this system a viable option for top-down proteomics. The protein spectra consisted primarily of *c*and *z*ions, though the ECD cell also produced a substantial number of *d*and *w*sidechain fragments. These side-chain fragments allow leucine/isoleucine or lysine/glutamine pairs to be distinguished, facilitating *de novo*sequencing. We then applied this technology to protein extracts from human brain to show that we can conduct top-down protein identification on LC time scales. The simpler fragmentation patterns made possible with the ECD device allows existing mass spectrometers to be able to characterize mid-sized proteins even using fast front-end separations such as ion mobility.

156

High-Resolution Differential Ion Mobility Separations with Orbitrap Mass Spectrometry for Middle-Down Analyses of Histone Proteoforms

Matthew A Baird[1], Pavel V Shliaha[2], Gordon A Anderson[3], Eugene Moskovets[4], Victor Laiko[4], Alexander A Makarov[5], Ole N Jensen[2], Alexandre A Shvartsburg[1]

1. Wichita State University, Wichita, KS, United States

2. University of Southern Denmark, Odense, Denmark

3. GAA Custom Engineering, Kennewick, WA, USA

4. MassTech, Columbia, MD, USA

5. Thermo Scientific, Bremen, Germany

The histone code central to epigenetics comprises a multitude of co-existing proteoforms with distinct PTM marks, including isoforms differing only in the PTM localization but having different biological functions. Characterizing that exceptionally complex mixture of highly similar species over wide abundance range is an astounding proteomic challenge, requiring the utmost from separation and identification methods. As nearly all histone PTMs reside on "tails" of ~50 residues protruding from the nucleosome, their characterization approaches that of complete histone. High-definition differential ion mobility spectrometry (FAIMS) has recently achieved broad resolution of common tail isoforms featuring methylation, acetylation, or phosphorylation. Separated forms were identified using standards or electron transfer dissociation (ETD). Those analyses previously employed ion trap MS platforms rather than Fourier-Transform MS (providing the ultimate mass resolution and accuracy mandatory for confident stoichiometric assignments) because requisite FAIMS resolution depended on gas buffers rich in helium or hydrogen that were incompatible with ultra-high vacuum needed for Orbitrap MS. We report enabling the high-resolution FAIMS/Orbitrap MS/ETD capability employing novel FAIMS devices with unprecedented voltages that reach sufficient resolution utilizing only nitrogen, tandem ion funnel FAIMS/MS interfaces removing light gases before the MS stage, and reconfiguration of Orbitrap

vacuum pumping that permits operation with He inflow. The performance of resulting instruments is demonstrated for exemplary modified tail isoforms and D-amino acid containing peptides. An analogous platform with custom curved (rather than planar) FAIMS devices provides better sensitivity at lower specificity. That version is employed in initial analyses of endogenous histone tails from human HeLa cells with online liquid chromatography in "external convolution" for pre-fractionation into isomer groups. Preliminary data indicate significant separation of multiple isoforms within all targeted groups and identification of numerous proteoforms not found in the benchmark LC-MS data without FAIMS or pre-existing public databases.

159

The Use of LC/MS and Bio-orthogonal Chemistry to Study Protein Dynamics in Cardiac Remodeling

John R Yates[1], Yuanhui Ma[1], David Liem[2], Dominic Ng[2], Peipei Ping[2], Daniel McClatchy[1]

1. Scripps Research, Lajolla, CA, United States

2. Department of Physiology, University of California at Los Angeles, Los Angeles, California, USA

Maladaptive cardiac remodeling (MCR) is a complex dynamic process common to many heart diseases. MCR is characterized as a temporal progression of global adaptive and maladaptive perturbations. The complex nature of this process clouds a comprehensive understanding of MCR, but greater insight into the processes and mechanisms has potential to identify new therapeutic targets. To provide a deeper understanding of this important cardiac process, we applied a new proteomic technique, PALM (Pulse Azidohomoalanine in Mammals), together with multi-dimensional LC to quantitate the newly synthesized protein (NSP) changes during the progression of isoproterenol(ISO) induced MCR in the mouse left ventricle. This analysis revealed a complex combination of adaptive and maladaptive alterations at acute and prolonged time points including the identification of proteins not previously associated with MCR. We also combined the PALM dataset with our published protein turnover rate dataset to identify putative biochemical mechanisms underlying MCR. The novel integration of analyzing newly synthesized proteins together with their protein turnover rates demonstrated that alterations in specific biological pathways (e.g. inflammation and oxidative stress) are produced by differential regulation of protein synthesis and degradation.

McClatchy DB, Ma Y, Liem DA, Ng DCM, Ping P, Yates JR 3rd.,Quantitative temporal analysis of protein dynamics in cardiac remodeling., J Mol Cell Cardiol. 2018 Aug;121:163-172. doi: 10.1016/j.yjmcc.2018.07.126.

160

Integrative analysis of plasma proteomes from prediabetes and diabetes progression: An IMI DIRECT study

Mun-Gwan Hong[1], Ana Viñuela[2], Ragna Häussler[1], Matilda Dale[1], Robert Koivula[3], Anubha Mahajan[3], Juan Fernandez[3], Frederico De Masi[4], Roberto Bizzotto[5], Tim McDonald[6], Petra Musholt[7], Andrea Mari[5], Imre Pavo[8], Tarja Kokkola[9], Oluf Borbye Pedersen[10], Jurek Adamski[11], Mark Walker[12], Emmanouil Dermitzakis[2], Søren Brunak[4], Paul Franks[13], Mark McCarthy[3], Hartmut Ruetten[7], Ewan Pearson[14], Jochen M Schwenk[1]

1. SciLifeLab, KTH Royal Institute of Technology, Stockholm, Sweden

2. Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland

3. Oxford Centre for Diabetes Endocrinology and Metabolism, Oxford University, Oxford, UK

4. Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

5. Institute of Neuroscience, National Research Council, Padova, Italy

6. Blood Sciences, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

7. Diabetes Division, Sanofi-Aventis Deutschland GmbH, Frankfurt, Germany

8. Eli Lilly, Vienna, Austria

9. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finnland

10. The Novo Nordisk Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark

11. German Center for Diabetes Research (DZD), Munich, Germany

12. Institute of Cellular Medicine, Newcastle University, Newcastle, UK

13. Department of Clinical Science, Lund University, Malmö, Sweden

14. Molecular and Clinical Medicine, University of Dundee, Dundee, Scotland

In-depth profiling of proteins in plasma can provide valuable insights on the complex processes related to metabolic diseases such as type 2 diabetes (T2D). Within the framework of the EU IMI project DIRECT (www.direct-diabetes.org) a multi-center cohort was built on 3100 subjects of which 2300 were at risk of developing of developing T2D (HbA_{1c}~ 5.6 - 6.5%) as well as 800 with early T2D (HbA_{1c}> 6.5%) [1].

We used several multiplexed affinity proteomic assays to profile ~600 unique proteins in EDTA plasma collected from 3100 study participants at baseline and 2500 at the 3 years (early T2D) or 4 years (risk for T2D) follow-up. With access to extensive metadata, our initial analysis focused on possible sample-related confounders. This identified several pre-analytical variables and consequently, we applied a linear mixed model that included age, sex, study center and sample collection date for defining proteins associated with any of the >50 quantitative clinical traits.

At baseline, we found > 300 proteins in plasma that were associated with diabetes related traits (adjusted p-value < 0.0001), many of which were prominently associated with BMI, such as leptin. Further, IGFBP1 and IGFBP2 associated to Matsuda; adiponectin to basal insulin secretion rate and fasting HDL; LDL receptor proteins to fasting triglycerides; APOM to fasting cholesterol; or IL8 and MCP-1 to fasting liver AST. Making use of other omics data, we performed pQTL analysis to assess any connection between the protein values in plasma and genetic variants. We observed ~400 cis-pQTLs (q-value < 0.05), such as for APOM, which illustrated that many of the studied protein profiles are affected by a genetic component.

Our integrative, large-scale multi-omics analysis revealed insights about known and novel plasma proteins associated to pre- and early T2D, as well as indicators of progression and treatment response.

1. Koivula RW, et al. Discovery of biomarkers for glycaemic deterioration before and after the onset of type 2 diabetes: an overview of the data from the epidemiological studies within the IMI DIRECT Consortium. bioRxiv 300244; doi: https://doi.org/10.1101/300244

161

Systems proteomics of the intermittent fasting response highlights the importance of hnf4a

Luke Hatchwell[1], Dylan Harney[1], Yen Koay[2], John O'Sullivan[2], Mark Larance[1]

1. Charles Perkins Centre and School of Life and Environmental Sciences, University of Sydney, Camperdown, NSW, Australia

2. Heart Research Institute, The University of Sydney, Sydney, NSW, Australia

Every-other-day-fasting (EODF) is an intermittent fasting regime that improves insulin sensitivity and lifespan in model animals without weight loss. However, the mechanisms mediating these beneficial phenotypes remain to be uncovered. Here, we have employed proteome analysis of mouse liver, a key fasting-responsive organ, to identify protein abundance changes after EODF compared to *ad libitum* fed animals. From >6,000 proteins quantified, more than 250 proteins were significantly altered by EODF. Among the most up-regulated proteins after EODF was acyl-CoA thioesterase 2 (ACOT2), which can accelerate liver fatty-acid oxidation that benefits whole-body metabolism. Surprisingly, alpha1-antitrypsin (SERPINA1) was the most down-regulated protein (>16-fold) after the EODF intervention. SERPINA1 function has previously been linked with lipoprotein particle metabolism in mice and humans. Given that SERPINA1 is among the 10 most abundant proteins in blood plasma and only synthesized in the liver, we performed single-shot plasma proteome analysis to quantify the top 200 most abundant proteins. This showed SERPINA1 was down-regulated ~3-fold in plasma of EODF animals, in addition to other significant protein changes such as increased abundance of apolipoprotein A-IV (APOA4), which was also increased in plasma of humans undergoing EODF and likely provides beneficial lower plasma triglycerides.

To identify how the decrease in liver SERPINA1 protein abundance was mediated, we measured its mRNA abundance and chromatin interactions. We observed a significant decrease in liver SERPINA1 mRNA after EODF and using CHIP-qPCR, and reduced association between HNF4A and the promoter of either SERPINA1, or an unrelated canonical HNF4A-target gene (ABCC6). These data suggest that HNF4A is globally inhibited by EODF, however the HNF4A liver protein abundance was not changed by EODF. Therefore, we hypothesise that HNF4A is EODF-regulated by either post-translational modification and/or changes in protein-protein interactions. Immunoprecipitations of HNF4A from the liver tissue are on-going to characterise any significant differences.

162

From disease to biology: how quantitative proteomics can resolve the molecular diagnosis of mitochondrial disease patients and provide insights into mitochondrial biology

Daniella H. Hock[1], Sumudu S. C. Amarasekera[2, 3], Alison G. Compton[2, 3], Ann E. Frazier[2, 3], Boris Reljić[1], David R. Thorburn[2, 3, 4], David A. Stroud[1]

1. Department of Biochemistry and Molecular Biology, University of Melbourne, Bio21 Molecular Science and Biotechnology Institute, Parkville, VIC, Australia

2. Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC, Australia

3. Department of Paediatrics, University of Melbourne, Parkville, VIC, Australia

4. Victorian Clinical Genetic Services, Royal Children's Hospital, Parkville, VIC, Australia

Mitochondrial diseases are the most common type of inherited metabolic disorders, affecting approximately 1 in 5,000 live births. Most routine diagnoses for mitochondrial disease rely on targeted exome sequencing combined with enzyme activity and brain scans. Due to the complex genetics and phenotypic heterogenicity over 40% of patients remain undiagnosed, often harbouring variants of unknown significance (VUS). Quantitative proteomics is a powerful untargeted approach that quantifies the abundance of thousands of proteins, offering functional data in variant prioritisation. We have applied quantitative proteomics in primary cells from patients with unsolved molecular diagnoses. In the first example, we identified an intronic mutation in a novel mitoribosome disease gene (*MRPL39*), which resulted in destabilisation of the large mitoribosomal subunits. Our cDNA studies showed that the variant produced a stable mRNA missing exon 8, leading to a premature stop codon. In another example, the patient harboured novel variants in both *MT-ATP6* and nuclear encoded *ATAD3A*. Our results clearly showed destabilisation of Complex V but no defect in *ATAD3A* transcript stability, protein abundance or pathways associated with known ATAD3 protein function. The *MT-ATP6* variant has been reclassified from VUS to likely pathogenic, such reclassification allows this variant to be included in prenatal tests. In the final example, the patient had a variant in the mitochondrial alanyl tRNA synthetase (*AARS2*), which resulted in a Complex IV defect. Further analysis of the proteomics results led us to investigate HIGD2A function. Using CRISPR/Cas9 we generated *HIGD2A* knockout HEK293T cells, which presented with a clear Complex IV defect. Characterisation of HIGD2A function revealed that it is specifically involved in assembling the MT-CO3 module of Complex IV. Hence, our studies demonstrate the powerful contribution of quantitative proteomics for the molecular diagnosis of mitochondrial disease patients and potential discovery of novel protein function.

163

Multi-omic profiling of metabolic dysfunction caused by myocardial ischemia / reperfusion (I/R) injury

Alexander W. Rookyard[1, 2], Desmond K. Li[2, 3], Yen Chin Koay[2, 4], John F. O'Sullivan[2, 4], Melanie Y. White[2, 3], Stuart J. Cordwell[1, 2, 3, 5]

1. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. Discipline of Pathology, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

4. Heart Research Institute , Sydney, NSW, Australia

5. Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia

Ischemic heart disease involves the occlusion of blood vessels resulting in a cessation of oxygenated blood flow to the heart. This hypoxia, and the necessary reperfusion to salvage surviving myocytes, induces cellular damage. Notably mitochondrial dysfunction occurs, increasing the production of reactive oxygen and reactive nitrogen species (ROS/RNS). This increase in ROS/RNS overwhelms cellular antioxidant defence and can alter protein structure / function via various protein post-translational modifications (PTMs). One target of ROS/RNS is the most redox active amino acid cysteine (Cys). Cys redox PTMs can be broken down into two classes, those that are biologically reversible (e.g. *S*-glutathionylation) or those considered 'irreversible' (sulfinic and sulfonic acid; Cys-SO_2H/SO_3H). Irreversible Cys redox PTM occur with prolonged exposure to high levels of ROS/RNS and are associated with protein dysfunction and/or degradation. A mass spectrometry technique based on parallel reaction monitoring was employed to detect changes in irreversible Cys modification in a Langendorff model of myocardial ischemia/reperfusion injury (I/R). Due to the low abundance of Cys, and low abundance of Cys PTMs, an enrichment strategy utilising strong cation exchange and hydrophilic interaction chromatography was used to better profile the changes in irreversible Cys PTM. I/R significantly increased the abundance of Cys-SO_2H/SO_3H-modified peptides from proteins involved in the tricarboxylic acid (TCA) cycle. By using a targeted MRM metabolomic workflow we observed concurrent perturbations in the concentration of metabolites involved in the TCA cycle also occurred during I/R. The addition of an aminothiol antioxidant MPG (N-2-mercaptopropionylglycine) in reperfusion improved functional recovery of hearts, ameliorated irreversible modification of Cys, and improved the recovery from TCA cycle metabolic dysfunction induced by ischemia.

166

Hyper-citrullinated library workflow to support demand-driven correct identification of citrullinated residues

Justyna Fert-Bober[1], Vidya Venkatraman[1], Ruining Liu[1], Christie Hunter[2], Rakhi Pandey[1], Jennifer Van Eyk[1]

1. Cedars Sinai Medical Center, Los Angeles, CA, United States

2. SCIEX, Redwood City, CA, USA

Background: Arginine citrullination is emerging as a widespread post-translational modification with diverse biological functions. To date, the citrullinated proteins have been reported only in a limited number of studies and while the modification itself has been linked to several diseases, including rheumatoid arthritis and cancer, its physiological or pathophysiological roles remain largely unclear. The bottleneck is limitations in available methodology to robustly enrich, detect, and localize the citrullinated residues. To overcome these limitations, we generated a mouse hyper-citrullinated spectral library and set up coordinates to confidently identify and validate citrullinated sites.

Methodologies: Generation of large-scale multiorgan mouse hyper-citrullinated spectral library that can be used to enable the robust detection of citrullinated peptides using DIA-MS. The step-by-step library generation workflow is based on creation of the hyper-citrullinated samples, collection of the high-quality fragment ion spectra in data dependent acquisition mode, spectral matching of modified peptides to their non-citrullinated form, along with the delta retention time shift (ΔRT), as a signature for citrullination. The validation steps of citrullinated peptides include detection of neutral loss of isocyanic acid in peptides in CID spectra and Skyline validation.

Findings: Using this workflow, we detect ten-fold increase in citrullinated proteome coverage across six mouse organs compared to the current state-of-the art techniques. Our data reveals that the subcellular distribution of citrullinated proteins is tissue-type dependent and that citrullinated targets are involved in fundamental physiological processes, including metabolic process.

Conclusions: Presented a novel approach for the generation of hyper-citrullinated library and detection of citrullinated proteins with the citrullinated site provides a rich resource of candidates for hypothesis generation and will open new avenues for large-scale investigations of citrullinated proteins in clinical research.

167

In Vitro Profiling of Ser/Thr/Tyr Selectivity of Human Protein Kinome

Naoyuki Sugiyama[1], Kazutaka Yachigo[1], Masato Tsuyuguchi[2], Takayoshi Kinoshita[2], Yasushi Ishihama[1]

1. Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, KYOTO, Japan

2. Graduate School of Science, Osaka Prefecture University, Osaka

Protein phosphorylation mediated by protein kinases play significant roles for cellular signal transduction networks. Human genome sequence analysis has revealed at least 518 genes encode protein kinases. These protein kinases have a phosphoacceptor preference, namely that serine/threonine kinases (STKs) and tyrosine kinases (TKs) specifically phosphorylate Ser/Thr and Tyr, respectively, and most STKs show Ser preference rather than Thr preference. There are some reports describing correlations between phosphoacceptor preference and amino acid sequence in a kinase activation loop[1,2], however it still remains unclear how the Ser/Thr/Tyr (S/T/Y) preferences of kinases arise. Here we investigated the phosphoacceptor preferences of human protein kinome by using *in vitro* kinase assay and phosphoproteomic approach, and extracted rules to determine S/T/Y preference from the kinome sequences.

HeLa cell lysate was dephosphorylated with alkaline phosphatase, and reacted individually with a recombinant human protein kinase. Tryptic digest was isotopically labeled with formaldehyde to distinguish substrates and endogenous phosphosites. Phosphopeptides were enriched with hydroxy acid-modified titanium dioxide chromatography and then analyzed with nanoLC/MS/MS.

We profiled 354 kinds of wild-type human protein kinases by using the *in vitro* kinase assay. Most STKs were strongly directed to Ser as expected, but MAPK, CDK families and some kinases had a relatively low Ser preference. Furthermore, most of STE group and a portion of the kinases belonging to TKL and others groups preferentially phosphorylated Thr. As reported by Chen et al[1], the amino acid residues following the common motif "DFG" in the kinase activation loop were highly correlated with Ser/Thr preference. We also found other amino acids governing S/T/*Y* preference in the kinase activation loop and experimentally validated them by *in vitro* assay with mutant kinases and synthetic peptides.

1. Chen, C. et al. Mol Cell 53, 140-147 (2014)
2. van de Kooij B et al. Elife 8: e44635 (2019)

168

Multi-stage discovery of hypermodified peptides and hundreds of post translational modifications from open modification search

Julie Wertz[1], Seungjin Na[2], Benjamin Pullman[1], Nuno Bandeira[1]

1. University of California, San Diego, La Jolla, CA, United States

2. Computer Science, Hanyang University, Seoul, South Korea

Accurate detection of modifications and hypermodified peptides requires careful exclusion of alternative explanations for the same spectra, as well as aggregation of multiple lines of evidence corroborating what would otherwise be very surprising identifications. Furthermore, automated interpretation of mass offsets detected by open modification search is also required to avoid labor-intensive manual curation to distinguish real modifications from search artifacts.

Maestro stratified search begins by identifying spectra in the smallest search space with the most information (spectral library search), followed by typical database search of unidentified spectra (considering only common modifications) and only afterwards is there an attempt to explain spectra that remain unidentified as possibly containing unexpected modifications. In addition, spectral networks algorithms are used to detect correlated peptide fragmentation patterns which help confirm (or challenge) surprising identifications by correlation to less-surprising identifications of related peptides. Finally, ModDecode builds on these to annotate detected mass offsets with (possibly combinations of) known modifications and thus identify many novel and rare post-translational modifications.

Reanalysis of a cell lines dataset (PXD004452) with over 12 million spectra revealed 510,899 modified peptide variants out of a total of 826,539 unique peptide variants. Surprisingly, the diversity of detected modifications spans over 200 different known modifications, all strongly supported by multiple lines of evidence. These also revealed hypermodified proteins with over 4,200 modified variants (84% of all variants mapped to the same protein), as well as 1,582 proteins with >75% of identifications coming from modified peptides. Interestingly, the distribution of modifications is highly non-uniform along the protein backbone, with modifications tending to cluster around hypermodified protein regions with up to 1,236 variants detected on a single 41 amino acid stretch on Alpha-2-HS-glycoprotein (P02765), as well as concentrating on specific peptides such as a single EIF1A peptide which was detected in over 140 distinct modified variants.

169

HUPO Proteome Project Phosphopeptide Challenge Report

Susan Weintraub[1], Michael Hoopmann[2], Ulrike Kusebauch[2], David Shytenberg[2], Nuno Bandeira[3], Magnus Palmblad[4], Lingje He[5], Bin Xia[5], Stoyan Stoychev[6], Justin Jordaan[6], Robert Moritz[2]

1. UT Health Science Center San Antonio, San Antonio, TX, United States

2. Institute for Systems Biology, Seattle, USA

3. University of California, San Diego, La Jolla, USA

4. Leids Universitair Medisch Centrum, Leiden, NL

5. SynPeptide Co., Ltd, Shanghai, China

6. ReSyn Biosciences, Edenvale, Gauteng, South Africa

The Human Proteome Project (HPP) MS Resource Pillar has established the HPP Phosphopeptide Challenge as a resource to help the community evaluate methods, learn procedures and establish their own workflows by comparing results obtained from a standard set of phosphopeptides and their unphosphorylated counterparts prepared as two samples: (1) a complex set of human phosphopeptides (Ser, Thr and Tyr) singly- and multiply-phosphorylated) and their unphosphorylated counterparts; (2) the peptides spiked into a yeast digest. The samples are valuable for method development, evaluation of phosphopeptide enrichment strategies, sequence analysis by mass spectrometry and bioinformatic interrogation. Together with our partners, SynPeptide Co. Ltd in Shanghai (www.synpeptide.com), and Resyn Biosciences Pty Ltd in South Africa (www.resynbio.com), the HPP MS Resource Pillar provided the comprehensive SynPeptide-HUPO phosphopeptide mixture and an affinity Resyn MagReSyn® purification kit including the magnet separator free to HUPO members.

Participants have contributed to the project by analyzing the sets of phosphopeptides by their favorite method and by use of affinity purification to facilitate identification of the phosphopeptides spiked into a complex mixture. As a result of this collaborative endeavor, multiple purification schemes, analytical protocols and data processing strategies have been submitted and evaluated,

making it possible to determine the approach(es) that provide the highest coverage of phosphopeptides in the MS Resource Pillar phosphopeptide mixture. We will discuss these results and provide resources for HUPO members to benefit from this initiative.

170

The International Cancer Proteogenome Consortium (ICPC): innovation and collaboration inspired by the Cancer Moonshot[SM]

Ana I Robles[1], Henry Rodriguez[1]

1. National Cancer Institute, Bethesda, MD, United States

The successes of the Clinical Proteomic Tumor Analysis Consortium (CPTAC) led NCI to strengthen its commitment to proteogenomic research. The International Cancer Proteogenome Consortium (ICPC), inspired by the spirit of international collaboration and data sharing encouraged by the Cancer Moonshot, started in 2016 and its vision is aligned with the Moonshot goals of accelerating progress in cancer research and its translation to patient care. ICPC encourages international cooperation and investments in proteogenomic cancer research. Aligned with CPTAC, ICPC brings together some of the world's leading cancer and proteogenomic research centers that adopt standard operating procedures for biospecimen collection and harmonized proteogenomic technologies/workflows to characterize commonly diagnosed cancers in their respective populations and address their specific unmet clinical needs. Its mission is to develop a cancer proteogenome atlas representative of the diversity of people with cancer worldwide, while implementing strategies to compare proteomic datasets generated using different platforms. The ICPC currently includes 33 institutions from 13 countries. Each institution has committed to publicly releasing study data (genomics, clinical phenotype, proteomics raw files, proteomic assays, antibodies, medical imaging) to the global research community. Currently, 13 cancer types are under characterization at the various sites. ICPC teams have local scientific and/or outreach activities, including roundtables and training workshops. A study using common internal standard reference materials will be conducted to harmonize workflows from sample preparation to informatics analysis across sites. Three ICPC data sets, comprising gastric cancer, liver cancer, and oral squamous cell carcinoma, are already hosted at the CPTAC Data Portal, a centralized repository for the public dissemination of CPTAC proteomic datasets. In the future, harmonized proteomic data from NCI-supported projects, including CPTAC, ICPC and others, such as Applied Proteogenomics OrganizationaL Learning and Outcomes (APOLLO), will be hosted at the NCI's Proteomic Data Commons (PDC), a cloud-based resource currently launched in Beta-phase.

171

Large Scale Melanoma Cancer Studies at the European Cancer Moonshot Lund Center in Partnership with Five International Hospitals

Johan Malm[1], Jeovanis Gil[1], Lazaro Betancourt[1], Indira Pla[1], Aniel Sanchez[1], Roger Appelqvist[1], Tasso Miliotis[1], Magdalena Kuras[1], Henriette Oskolas[1], Yonghyo Kim[1], Zsolt Horvath[1], Jonatan Eriksson[1], Ethan Berge[1], Elisabeth Burestedt[1], Göran Jönsson[1], Bo Baldetorp[1], Christian Ingvar[1], Håkan Olsson[1], Lotta Lundgren[1], Peter Horvatovich[1], Jimmy Rodriguez Murillo[1], Yutaka Sugihara[1], Charlotte Welinder[1], Elisabet Wieslander[1], Boram Lee[1], Henrik Lindberg[1], Krzysztof Pawlowski[1], HoJeong Kwon[1], Viktoria Doma[1], Jozsef Timar[1], Sarolta Karpati[1], A. Marcell Szasz[1], István Balázs Németh[1], Toshihide Nishimura[1], Garry Corthals[1], Melinda Rezeli[1], Beatrice Knudsen[1], György Marko-Varga[1]

1. Lund University, Lund, SKANE, Sweden

Melanoma of the skin is the sixth most common type of cancer in Europe and accounts for 3.4% of all diagnosed cancers. More alarming is the increase of highly aggressive forms of skin cancer and metastases spread, and the degree of recurrence that occurs in approximately 20% of patients lethally relapsing following treatment. Targeted oncotherapy is one of the standard treatments for progressive stage 3 and 4 melanoma, and (e.g; vemurafenib, dabrafenib, encorafenib) combined with a MEK inhibitor (e.g: trametinib, cobimetinib, binimetinib) can be used to effectively treat patients with BRAFV600E-mutated melanomas.

By combining our data from, e.g. phosphoproteomics and acetylomics, the protein expression profiles of different melanoma stages can provide a solid framework for understanding the biology and progression of the disease. By complementing by proteogenomics, customised protein sequence databases generated from our patient-specific genomic and transcriptomic data aid in interpreting clinical proteomic biomarker data to provide a deeper and more comprehensive molecular characterisation of cellular functions underlying disease progression.

In parallel to a streamlined, patient-centric, clinical proteomic pipeline, mass spectrometry–based imaging studies are conducted at the European Cancer Moonshot Center, characterizing the spatial distribution of drugs and drug metabolites within tissues at single-cell resolution. These developments are an important advancement in studying drug action and efficacy in vivo and will aid in the development of more effective and safer strategies for the treatment of melanoma. The primary research focus of the European Cancer Moonshot Lund Center is to understand the impact that drugs have on cancer at an individualised and personalised level. Simultaneously, the centre increases awareness of the relentless battle against cancer and attracts global interest in the exceptional research performed at the centre.

172

Proteogenomics of colorectal cancer liver metastases: complementing precision oncology with phenotypic data

Bernhard Blank-Landeshammer[1], Vincent R Richard[2], Georgia Mitsa[2], Laxmikanth Kollipara[1], Ingo Feldmann[1], Maud Marques[3], Karen Gambaro[3], Suzan McNamara[3], Alan Spatz[4], René P. Zahedi[2, 4], Albert Sickmann[1], Gerald Batist[4], Christoph H Borchers[2, 4, 5, 6]

1. Leibniz-Institut für Analytische Wissenschaften – ISAS - e.V., Dortmund, Germany

2. Segal Cancer Proteomics Centre, Montreal, QC, Canada

3. Exactis Innovation and Segal Cancer Centre, Montreal, QC, Canada

4. Gerald Bronfman Department of Oncology, Jewish General Hospital, Montreal, Canada

5. UVic Proteomics Centre, Victoria, Canada

6. Dept. of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada

In colorectal cancer (CRC), KRAS mutational status is an important precision oncology marker determining whether patients are eligible for targeted anti-EGFR treatment. Nevertheless, response rates to anti-EGFR treatment are only 20-30%, while certain patients respond despite having KRAS-mutations. We describe a proteogenomic analysis of CRC liver metastases (mCRC), representing an ideal setting for the analysis of therapeutic resistance. Deep proteomic profiling of $KRAS^{wildtype}$ and $KRAS^{G12V}$ mCRC enabled the identification of >9,000 proteins. We detected considerable changes in protein expression, both in line and in discordance with whole exosome sequencing (WES) and RNAseq data, including numerous proteins with a known role in progression and resistance of CRC tumors. Interestingly, ERBB2 was massively upregulated in the $KRAS^{wildtype}$ tumor, despite the absence of mutations in both tumors. We identified protein evidence for a number of predicted somatic mutations, among those $KRAS^{G12V}$. For 8 proteins, we developed targeted parallel reaction monitoring (PRM) assays that enabled the precise and sensitive absolute quantification of both, the mutated and the canonical variants. Using these PRM assays, we phenotyped individual mCRC tumors and paired healthy tissues, by determining the actual mutation rates on the protein level. Our data demonstrates that total KRAS protein expression varies between tumors (0.47–1.01 fmol/µg total protein) and healthy tissues (0.13–0.64 fmol/µg). In some patients, KRAS protein expression is comparable between tumor and control tissue, while in others KRAS is significantly upregulated in the tumor. In most $KRAS^{G12V}$-positive tumors, the G12V-mutation level was 42-100%, while one patient's tumor had only 10% $KRAS^{G12V}$ but 90% $KRAS^{wildtype}$. This might be an example for a potential false-negative exclusion from targeted treatment: Although the patient was ineligible for targeted anti-EGFR treatment based on hotspot sequencing and instead received chemotherapy, a (combined) anti-EGFR therapy might have been a therapeutic option – potentially missed because of lacking phenotypic information.

173

Integrated Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma

David J Clark[1], Saravana M Dhanasekaran[2], Francesca Petralia[3], Jianbo Pan[1], Xiaoyu Song[4], Yingwei Hu[1], Felipe da Veiga Leprevost[2], Boris Reva[3], Tung-Shing M Lih[1], Hui-Yin Chang[2], Weiping Ma[3], Chen Huang[5], Christopher J Ricketts[6], Lijun Chen[1], Azra Krek[3], Yize Li[7], Dmitry Rykunov[3], Qing Kay Li[1], Lin S Chen[8], Umut Ozbek[4], Suhas Vasaikar[9], Shrabanti Chowdhury[3], Yige Wu[7], Seungyeul Yoo[3], Shrabanti Chowdhury[3], Matthew A Wyczalkowski[7], Jiayi Ji[3], Michael Schnaubelt[1], Andy Kong[2], Sunantha Sethuraman[7], Dmitry M Avtonomov[2], Minghui Ao[1], Antonio Colaprico[10], Song Cao[7], Kyung-Cho Cho[1], Selim Kalayci[3], Shiyong Ma[1], Wenke Liu[11], Anna Calinawan[3], Zeynep H Gumus[3], Feng Chen[12], Nathan Edwards[13], Phillip M Pieroazio[14], Xi Steven Chen[15], Christian P Pavlovich[14], A. Ari Hakimi[16], Gabriel Brominski[17], James J Hsieh[18], Andrzej Antczak[17], Tatiana Omelchenko[19], Jan Lubinski[20], Maciej Wiznerowicz[17], W. Marston Linehan[6], Chris R Kinsinger[21], Mathangi Thiagarajan[22], Emily S Boja[21], Mehdi Mesri[21], Tara Hiltke[21], Ana I Robles[21], Henry Rodriguez[21], Jiang Qian[23], David Fenyo[11], Bing Zhang[24], Li Ding[7], Eric Schadt[3], Arul Chinnaiyan[2], Zhen Zhang[1], Gil Omenn[25], Marcin Cieslik[2], Daniel Chan[1], Alexey I Nesvizhskii[2], Pei Wang[3], Hui Zhang[1]

1. Department of Pathology, The Johns Hopkins University, Baltimore, MD, USA

2. Department of Pathology, University of Michigan, Ann Arbor, MI, USA

3. Department of Genetics & Genomic Sciences, Icahn Institute for Data Science and Genomic Technology, New York, NY, USA

4. Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA

5. Lester and Sue Smith Breast Cancer Center, Baylor College of Medicine, Houston, TX, USA

6. Urologic Oncology Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA

7. Washington University School of Medicine, St. Louis, MO, USA

8. Department of Public Health Sciences, University of Chicago, Chicago, IL, USA

9. Department of Translational Molecular Pathology, MD Anderson Cancer Center, Houston, TX, USA

10. Department of Public Health Sciences, University of Miami Miller School of Medicine, Miami, FL, USA

11. Institute for Systems Genetics and Department of Biochemistry & Molecular Pharmacology, New York University School of Medicine, New York, NY, USA

12. Departments of Medicine and Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO, USA

13. Department of Biochemistry & Cellular Biology, Georgetown University, Washington D.C., USA

14. Brady Urological Institute and Department of Urology, The Johns Hopkins University, Baltimore, MD, USA

15. Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, USA

16. Department of Surgery, Urology Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA

17. Poznan University of Medical Sciences, Poznan, Poland

18. Department of Medicine, , Washington University School of Medicine, St. Louis, MO, USA

19. Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA

20. Pomeranian University of Medicine, Szczecin, Poland

21. Office of Cancer Clinical Proteomics Research, National Cancer Institute, Bethesda, MD, USA

22. Frederick National Laboratory for Cancer Research, Frederick, MD, USA

23. Department of Ophthalmology, The Johns Hopkins University, Baltimore, MD, USA

24. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

25. Departments of Computational Medicine & Bioinformatics, Internal Medicine, Human Genetics, and School of Public Health, University of Michigan, Ann Arbor, MI, USA

Background: Clear cell renal cell carcinoma (ccRCC) is the most predominant histology of renal cancer, representing 75% of all cases and accounting for the majority of associated deaths. To gain insight into the impact of genomic alterations on the functional modules that drive ccRCC tumorigenesis, we leveraged comprehensive proteogenomic characterization of 110 treatment-naïve renal cell carcinoma (RCC) and 84 paired-matched normal adjacent tissue (NAT) samples.

Methods: We utilized an integrated proteogenomic approach, performing whole genome sequencing (WGS), whole exome sequencing (WES), DNA methylation profiling for all tumors; RNA-seq, proteomic, and phosphoproteomic characterization was performed for all samples.

Results: WGS analysis revealed arm-level loss of chromosome 3p as a frequent event in ccRCC, with 61% of tumors showing evidence of 3p chromosomal translocation events. Comparative profiling of ccRCC and NATs samples identified pathways associated with immune response, EMT, and glycolysis to be up-regulated in ccRCC, while TCA cycle, fatty acid metabolism, and oxidative phosphorylation were down-regulated. Examination of mRNA-protein correlation revealed a non-linear relationship in cellular processes including Warburg Effect-related metabolism, as well as the tumor-specific trend of higher sample-wise correlation associating with prognostically-defined aggressive features of ccRCC. Analysis of differential phosphosite occupancy between tumors and NAT showed MAPK/ERK signalling and G2/M stalling to be up-regulated across the majority of ccRCC cases. We deconvoluted immune and stromal cell gene signatures in the tumor microenvironment (TME), with consensus clustering of the TME compositions identifying four immune-based subtypes: Inflamed CD8+, Inflamed CD8-, VEGF Immune Desert, and Metabolic Immune Desert. Integrated transcriptomic and proteomic profiling of the ccRCC subtypes revealed unique, discriminatory signalling pathways associated with immune exhaustion, cancer-associated fibroblast-related signalling, angiogenesis, and metabolic activity.

Conclusions: Our results link the functional impact of genomic alterations at the protein level, and provides evidence for rational treatment selection stemming from proteomic, phosphoproteomic, and tumor microenvironment signatures.

174

Patient-derived cancer model for proteogenomics: Report by ICPC JAPAN team

Rei Noguchi[1], Emi Hattori[1], Akihiko Yoshida[2], Akira Kawai[3], Tadashi Kondo[1]

1. Division of Rare Cancer Research, National Cancer Center Research Institute, Tsukiji, Tokyo, Japan

2. Department of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tsukiji, Tokyo, 2974049

3. Division of Musculoskeletal Oncology, National Cancer Center Hospital, Tsukiji, Tokyo, Japan

Patient-derived cancer models play an important role in the basic and translational research. They are indispensable for the functional significances of aberrant genes and proteins, and the response to novel anti-cancer drugs. Although many cell lines, organoids, and xenograftsh have been developed, and deposited in the public biobanking, we need to establish more patient-derived cancer models, considering the diversity and complexity of malignancies. The establishment of cancer models has been required especially for the rare cancers such as sarcomas. They are hardly available, probably because of the rarity of disease, hindering the basic and pre-clinical study in rare cancers. ICPC Japan team conducts the proteogenomics study in sarcomas. Sarcomas are the rare mesenchymal malignant tumors in bone and soft tissues, accounting for less than 1% of all malignancies. Sarcomas compose of more than 50 histological subtypes, which show the different clinical and pathological features with distinct molecular backgrounds. We generate the genome and proteome data of sarcomas, integrate them, trying to understand the molecular backgrounds of etiology, disease developmemt, and resistance against therapy. As a part of cancer proteogenomics, we established patient-derived sarcoma models, to study the functional significances of identified genes and proteins. Presently, we have established more than 40 patient-derived xenografts and 30 cell lines. We investigate the genetic mutations, enzyme activities, and effects of inhibitors in the developed cell lines. We found that the presence of kinase mutations does not guarantee the favorable response to the inhibitors against them, and the significant inhibitory effects were not always based on the genetic mutations. The association of the genetic mutations and the response to inhibitors may be the cancer dependent, and the examinations of kinase activity may solve the apparent discrepancies. In this sense, proteogenomics approach will generate intriguing and useful outcome for basic cancer research and clinical application.

175

Proteogenomic Characterization of Human Gastric Cancer

Jianmin Wu[1], Ludovic Gillet[2], Tiannan Guo[2, 3], Yang Du[1], Sijia Cui[1], Luxi Zhang[1], Huan Yu[1], Lihua Cao[1], Peng Xue[2, 4], Yi Zhu[2, 3], Yang Yang[1], Man Wang[1], Xiaoqing Guan[1], Bin Dong[1], Ying Hu[1], Xiaofang Xing[1], Xiangyu Gao[1], Fei Shan[1], Lei Tang[1], Lianhai Zhang[1], Zhongwu Li[1], Xiangqian Su[1], Aiwen Wu[1], Zhaode Bu[1], Ziyu Li[1], Shuqin Jia[1], Ruedi Aebersold[2, 5], Jiafu Ji[1]

1. Key laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Peking University Cancer Hospital & Institute, Beijing, China

2. Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

3. School of Life Sciences, Westlake University, Hangzhou, Zhejiang Province, China

4. Institue of biophysics, Chinese academy of science, Beijing, China

5. Faculty of Science , University of Zurich, Zurich, Switzerland

Gastric Cancer is the fifth most prevalent malignancy and the third leading cause of cancer death worldwide, of which more than half occurred in East Asia. Although previous genomic studies demonstrated the potential for molecular classifications of gastric

cancer for personalized medicine, e.g. microsatellite Instability (MSI)-high subtype of patients showing resistant to chemotherapy, a large proportion of gastric cancer patients need further molecular investigation to develop tailored therapeutic regimen. We characterized the proteogenomic landscapes of 186 gastric cancer patients (including all Lauren classifications) by integrative analysis of genomic, transcriptomic and proteomic data. Fresh-frozen treatment-naive gastric cancer surgical samples were collected from the biobank of Peking University Cancer Hospital (2010.01-2013.06). Tumor cellularity was examined by H&E staining, and samples with tumor cellularity >=60% were selected for exome sequencing, RNA-seq and reproducible proteomic quantification, by pressure cycling technology (PCT) assisted sample preparation followed by SWATH-MS. The mass spectrometry data were processed for consensus clustering analysis, which defined four proteomic subtypes in our cohort. We find that this molecular classification is significantly correlated with patient outcome and presents distinct molecular characteristics between subtypes. By combining genomic and proteomic features, our analysis elucidates the molecular details of gastric cancer, and identifies novel prognosis markers and therapeutic targets. Correspondence to Jianmin Wu (wujm@bjmu.edu.cn), Ruedi Aebersold (aebersold@imsb.biol.ethz.ch) and Jiafu Ji (jijiafu@hsc.pku.edu.cn).

176

Proteogenomic Analysis of Cancer Point Mutations - A Chromosomal Map

Iulia M. Lazar[1], Arba Karcini[1], Shreya Ahuja[1]

1. Virginia Tech, Blacksburg, UNITED STATES OF AMERICA, United States

The heterogeneous, unstable genome of cancerous cell states evolves over time as a result of an accumulation of mutations and chromosomal aberrations. Alterations in the structure of oncogenes and tumor suppressors lead to abnormal protein expression, gains or loss in protein function, perturbed protein-protein interactions, and ultimately to an abnormal cellular response that supports uncontrolled cell proliferation. High-throughput sequencing of thousands of tissues has revealed a critical need for deciphering the functional impact of the complex mutational landscape. The objective of this work was to address this need by exploring the protein-level profile of somatic mutations in cancer cells by using mass spectrometry (MS) detection.

The study was conducted using ER/HER2+ breast cancer and non-tumorigenic cells, cultured under proliferative and arrest-inducing conditions. The cell extracts were analyzed by nano-LC-MS/MS using LTQ-XL, QExactive and Orbitrap Lumos mass spectrometers. Protein/peptide identifications were performed with the Proteome Discoverer 2.3 and Mascot software packages (FDR=0.5-1 %) and an in-house built database containing ~2.5 million mutated peptide sequences.

The MS analysis of various cell states enabled the detection of >5,000 proteins and 150-250 point mutations in each cell line, reflecting the altered state of a number of oncogenes and tumor suppressors. One surprising finding was that the protein-level mutations did not mirror but rather complement the cDNA and RNA profiles. The sources of this discrepancy were explored, and the impact of this finding on identifying cancer driver genes was assessed. Criteria for selecting mutated peptide sequences are proposed, and chromosomal maps of missense mutations were constructed to evaluate diagnostic or prognostic potential. It was concluded that the combined use of genome- and proteome-level mutation data, complemented with protein abundance measurements, will help assess the cumulative effect of mutational hotspots, reveal susceptibility to cancer development, and guide the choice of effective therapeutic decisions.

179

Autoantibody Profiling of Schizophrenia and related Psychotic Disorders in the Australian National Survey of High Impact Psychosis (SHIP)

K. Oliver Schubert[1, 2], August Jernbom Falk[3], David Just[3, 4], Anna Manberg[3], Cherrie Galletly[1, 2], Peter Nilsson[3]

1. Discipline of Psychiatry, Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia

2. Northern Adelaide Mental Health Service, SA Health, Salisbury, SA, Australia

3. SciLifeLab, KTH Royal Institute of Technology, Stockholm, Sweden

4. Uppsala University, Uppsala, Sweden

Psychotic disorders such as schizophrenia, bipolar disorder, or schizoaffective disorder affect approximately 3% of the population. The pathophysiology of these highly debilitating illnesses is still poorly understood and current treatments fail to adequately alleviate symptoms in at least 30% of patients. There is growing evidence for a role of autoimmunity in psychiatric disorders. In the present study, we aimed to explore the IgG repertoire in a cohort of 473 Australian patients diagnosed with psychotic disorders to reveal potential autoantibody repertoires and their relationship to disease phenotypes and functional outcomes.

In an initial screening experiment, we used an in-house planar protein array containing 42000 protein fragments representing approximately 18000 human proteins on eight patient pools representative of diagnostic and phenotypic groups within SHIP. Based on the results from these arrays, previous findings, and literature mining, we then designed a 384-plex suspension bead array (SBA) for analysis of all individuals in the cohort.

Using the planar protein arrays, we found 181 protein fragments that were defined as reactive in one or several of the sample pools. Out of the 181, 20 fragments were found reactive in at least two of the sample pools. Follow-up of these findings in the entire SHIP cohort showed that reactivity profiles are highly individual, but also revealed novel autoantibody targets in psychotic disorders that are associated with specific phenotypic characteristics. The current study represents that largest investigation of autoimmunity profiles in psychosis undertaken to date. Findings may be of relevance for stratification of diagnosis and treatment in psychiatry.

180

Human brain proteome draft atlas to understand biological function and reference for brain tumour investigations

Sanjeeva Srivastava[1], Deeptarup Biswas[1]

1. Indian Institute of Technology Bombay, Mumbai, MAHARASHTRA, India

Introduction: The human brain is primary organ of CNS. The existing understanding of the neuropathology is hindered by the lack of region-specific brain proteome data. In this study we have made a draft human brain proteome atlas by performing high-resolution mass spectrometry-based proteomics of normal regions of brain and publicly available repositories. Further, brain proteome reference map is used for investigation of gliomas, meningioma and medullblastoma tumors & understanding disease pathobiology.

Materials & Methods: In this study, 13 different regions of the brain and 3 regions of the covering of the brain (meninges) were investigated to provide a region-specific brain proteome map. The data for 12 regions were obtained from publicly available repositories whereas the remaining 4 region data was generated in-house. Further analysis was done in Max Quant and protein list was further matched with HPA dataset. Many region-specific proteins, popular proteins and few missing proteins were identified. Further, normal brain proteome reference map was used for our brain tumour projects on gliomas, meningiomas and medulloblastoma to perform quantitative proteomic-based investigations of these deadly tumours.

Conclusion: We have made a brain proteome atlas and found that despite being part of the same organ, the various regions and sub regions exhibit distinct differences in protein expression. A region-specific proteomic analysis of these differences provided a region and sub-region based platform for better understanding of protein function and pathways. We also found region-specific unique proteins which can be used as region specific marker. On the other hand, we also found few proteins that belongs to PE2 level missing protein category.Further, normal brain proteome reference map was used for our brain tumour projects on gliomas, meningiomas and medulloblastoma, which has revealed grade-wise protein biomarkers and provided new mechanistic insights of these deadly brain tumors.

181

Proteomic and bioenergetic analyses demonstrate synaptic mitochondrial alterations due to age-related pathologic tau accumulation

Andrew J Trease[1], Joseph W George[1], Katy Emanuel[1], Howard S Fox[1], Kelly L Stauch[1]

1. University of Nebraska Medical Center, Omaha, NE, United States

Alzheimer's disease (AD), the most common form of dementia, is histopathologically characterized by the accumulation of extracellular amyloid plaques and intracellular neurofibrillary tangles (NFTs, consisting of hyperphosphorylated and aggregated tau) in the brain, which are accompanied by a dramatic synaptic and neuronal loss. Despite evidence that NFT stage is a better correlate with dementia than plaques, the molecular mechanism(s) responsible for tau-induced central nervous system side effects are elusive, although growing evidence points to disturbances in mitochondrial and synaptic function; processes also affected by normal aging. Here, using SWATH-based proteomics and crosslinking mass spectrometry combined with bioenergetics analyses, we establish a role for pathologic tau in synaptic mitochondrial dysfunction. We discover that in mice expressing human tau, but not mouse tau, AD-relevant pathologic forms of tau associate with synaptic mitochondria and induce a compromised energy status at the synapse as these animals age. This work highlights the relationship between tau pathology and impaired mitochondrial function at the synapse, providing a framework for future studies that will lead to improvements in the treatment of neurodegenerative deficits in AD. The mechanism we describe also has implications beyond AD, as tau pathology is observed in a wide range of disease paradigms.

182

A Proteomic Atlas of Senescence-Associated Secretomes for Aging Biomarker Development

Nathan Basisty[1], Abhijit Kale[1], Okhee Jeon[1], Chisaka Kuehnemann[1], Therese Payne[1], Chirag Rao[1], Anja Holtz[1], Samah Shah[1], Luigi Ferrucci[2], Judy Campisi[1, 3], Birgit Schilling[1]

1. The Buck Institute for Research on Aging, Novato, CALIFORNIA, United States

2. National Institute on Aging, Baltimore, Maryland, United States

3. Lawrence Berkeley Laboratory, Berkeley, California, United States

Introduction: The senescence-associated secretory phenotype (SASP) has recently emerged as both a driver of, and promising therapeutic target for, a multitude of chronic age-related conditions, ranging from neurodegeneration to cancer. The complexity of the SASP has been greatly underappreciated and a small set of factors cannot explain the diverse phenotypes it produces in vivo. Here, we present 'SASP Atlas', a comprehensive proteomic database of SASPs, including a novel exosome SASP phenotype, driven by multiple inducers of senescence in different human cell types. We also propose that SASP proteins are promising biomarkers to assess senescent cell burden in aging and disease.

Methods: Secretomes of senescent cells were characterized by comparing the secreted soluble (sSASP) and exosome (eSASP) proteins of irradiation-, oncogenic RAS-, or HIV drug atazanavir-induced senescent human fibroblasts to non-senescent controls. Secreted proteins and exosomes were obtained from the medium of cells cultured for 24 hours in serum-free conditions. Data-independent acquisitions (DIA) were performed on a TripleTOF 6600.

Results: While 172 proteins increased in all inducers and 53 proteins increased in all inducers and cell types, the SASP of each inducer and cell type were largely distinct, totaling over 1000 unique proteins. CXCL1, MMP1, and STC1 were consistently among the top increased proteins in response to all inducers. Senescent cells, on average, secreted larger exosomes containing a highly distinct set of proteins compared with the soluble secretome, including proteins involved in G-protein and RAS signaling, prostaglandin regulation and the complement system.

Discussion: Together, the data demonstrate that the SASP is a complex and highly diverse set of secretory phenotypes, including a unique exosome SASP. This resource will aid in identifying the proteins that drive senescence-associated phenotypes and provide comprehensive catalogs of potential biomarkers for assessing the burden and origin of senescent cells in vivo.

185

Aligning label-based discovery and global DIA validation proteomics to explore bacterial virulence phenotypes

Lok Man[1, 2], Joel A Cain[1, 2], Stuart J Cordwell[1, 2]

1. School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

The use of proteomics to inform subsequent biological validation studies requires substantial rigor in the analytical approach to ensure that the most important leads are followed. Our laboratory explores virulence determinants including an *N*-linked glycosylation (*pgl*) system and nutrient transporters in the gastrointestinal pathogen, *Campylobacter jejuni*. *C. jejuni* is a Gram negative, spiral and micro-aerophilic bacterium with a sequenced genome containing ~1620 genes. Target identification is based on the response of the proteome to environmental conditions that mimic the host, including bile salts, low iron, mucin availability and growth temperatures. The proteomics workflow includes parallel label-based liquid chromatography / tandem mass spectrometry (LC-MS/MS; minimum 3 biological growth replicates) using TMT and / or iTRAQ labelling, and system-wide validation using data independent analysis (DIA-SWATH-MS; minimum duplicate additional biological replicates). We routinely quantify ~80-90% of the predicted *C. jejuni* NCTC11168 proteome using label-based LC-MS/MS (2 peptides; <1% FDR), and ~65-75% of these proteins can be validated by DIA-SWATH-MS. Here, we will discuss the correlation between large-scale datasets in this biological system and how they facilitate subsequent studies, as well as highlight poorly or non-correlating data. In each case, we show how validated changes in the *C. jejuni* proteome reflect 'functional reality' that can be determined by integrated 'omics approaches including transcriptomics, metabolomics, lipidomics and lipid A analysis combined with molecular genetics and cellular and *in vivo* virulence assays. This approach has been employed to characterise functions associated with the *pgl* system and to identify a nutrient transporter essential for maintaining sulphur import and wild-type human cell virulence phenotypes.

186

An unbiased metaproteomic approach to describe the mucosal microbiome of an HIV-exposed African infant cohort

Jonathan M Blackburn[1], Suereta Fortuin[1], Phumudzo Khomunala[1], Matthys Potgieter[1], Jerome Wendoh[1], Nicola Mulder[1], Heather Jaspan[1]

1. University of Cape Town, Cape Town, WESTERN CAPE, South Africa

Background: Infants born to HIV-infected mothers have an altered cellular immunity despite being HIV-infected themselves. It is known that the gut microbiome is crucial to immune development. Currently, 16S rRNA sequencing, the gold standard for characterising the microbiome, yields limited information on the function of the microbial community, predicted open reading frames might not be expressed under in vivo conditions.

Methods: Here, infant stool samples, both HIV-exposed and unexposed infants collected at birth and at 4-7days were analysed to characterise the metaproteome using mass spectrometry-based proteomics. Briefly, cold stored samples were snap-frozen and ground samples were suspended in organic buffer and precipitated proteinaceous material were subjected to in-solution trypsin digestion. Desalted tryptic peptides were analysed in triplicate on the nLC-MS/MS. An in-house developed pipeline, Metanovo, was used to construct a metaproteomic database using the Universal Reference Database and generated mass spectrometry data. The metaproteomic database was used for protein and organism identification in MaxQuant-suite.

Results: Using our unbiased approach, we identified 3943 protein groups for all 55 samples described. The microbial diversity changed dynamically from birth to 4-7 days after birth and the proportion of human proteins identified decreased as the microbial diversity evolved. Using this approach, we identified virus, parasite, bacteria and human proteins in a single sample. The number of proteins belonging to the Bifidobacteriaceae family were significantly different between HEU and HU infants at birth, 1 to 180 respectively. Furthermore, this bacterial family increased dramatically to 1800 proteins within the first week of birth. Members of the Bifidobacteriaceae family are important microbes that colonize the gut of humans and other animals during the early stages of life.

Conclusion: Here we demonstrated that using mass spectrometry-based proteomics approach we were able to construct a metaproteomic profile of the microbiota composition of infant stool samples.

187

Methylation throughout the proteome: the methyltransferases tell the story

Joshua J Hamey[1], Marc R Wilkins[1]

1. Systems Biology Initiative, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia

Protein methylation is a widespread post-translational modification in the eukaryotic cell. It regulates key cellular processes including transcription, chromatin remodeling, signal transduction and RNA splicing. Yet, there remains much to be discovered about the roles of this modification. Central to answering this question are the methyltransferases that catalyse protein methylation. In recent years there has been incredible progress in the discovery of these enzymes, in particular in yeast and human. We discovered two yeast enzymes, Efm3 and Efm7, as well as two human enzymes, eEF1A-KMT1 and eEF1A-KMT3,

that each methylate specific residues in translation elongation factors [1,2]. To assist the characterisation of these methyltransferases we developed Methyltransferase Motif Analysis by Mass Spectrometry (MT-MAMS) [3]. This technique gives unique insight into the sequence specificity of methyltransferases, providing essential clues to the function and potential therapeutic targeting of these enzymes. Functionally, many of these new enzymes specifically target translation elongation factor 1A (eEF1A), an essential protein involved in protein synthesis, protein degradation and cytoskeletal organization [4]. This has led to the realisation that eEF1A is targeted by more independent methyltransferases than any other protein in eukaryotes. Through SILAC-based proteomics, we found that eEF1A methylation is subtle in function, suggesting it may fine-tune translation. Recently, we have been systematically exploring the interplay between methylation two other prominent modifications, phosphorylation and acetylation. We have already found evidence that methylation events, such as those on eEF1A, can co-occur with phosphorylation, suggesting that interplay between these modifications may be more common than previously appreciated. Due to the substantial number of methyltransferases discovered in the field recently, the complete set of all yeast methyltransferases and substrate proteins has nearly been uncovered. Investigation of this near-complete yeast methylproteome network provides systems-level insights into the function and evolution of this important modification and the enzymes that catalyse it.

1. [1] Hamey, J. J., Winter, D. L., Yagoub, D., Overall, C. M., Hart-Smith, G., & Wilkins, M. R. (2016). Novel N-terminal and lysine methyltransferases that target translation elongation factor 1A in yeast and human. Molecular & Cellular Proteomics, 15(1), 164-176.
2. [2] Hamey, J. J., Wienert, B., Quinlan, K. G., & Wilkins, M. R. (2017). METTL21B Is a Novel Human Lysine Methyltransferase of Translation Elongation Factor 1A: Discovery by CRISPR/Cas9 Knockout. Molecular & Cellular Proteomics, 16(12), 2229-2242.
3. [3] Hamey, J. J., Separovich, R. J., & Wilkins, M. R. (2018). MT-MAMS: protein methyltransferase motif analysis by mass spectrometry. Journal of proteome research, 17(10), 3485-3491.
4. [4] Hamey, J. J., & Wilkins, M. R. (2018). Methylation of Elongation Factor 1A: Where, Who, and Why? Trends in Biochemical Sciences, 43(3), 211-223.

188

Phosphoproteomic study on Staphylococcus *aureus* to identify phosphoproteins involved in virulence

Nadine Prust[1], Clément M. Potel[1], Nina van Sorge[2], Simone Lemeer[1]

1. Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands

2. Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands

Staphylococcus aureus is the leading cause of infections worldwide and infection results in a variety of diseases. As of no surprise, phosphorylation is a major game player in signaling cascades and has been shown to be involved in S. *aureus* virulence. Albeit long neglected, eukaryotic-like serine/threonine kinases have been implicated in these complex signaling cascades. Due to the sub-stoichiometric nature of protein phosphorylation and a lack of suitable analysis tools, the knowledge of these cascades is, however, to date still limited.

Here, were apply an optimized protocol for efficient phosphopeptide enrichment via Fe^{3+}-IMAC followed by LC-MS/MS to get a better understanding of the impact of protein phosphorylation on the complex signaling networks involved in pathogenicity. By profiling a serine/threonine kinase and phosphatase mutant from a methicillin-resistant *S. aureus* mutant library, we generated the most comprehensive phosphoproteome dataset of S. *aureus* to date, aiding a better understanding of signaling in bacteria. With the identification of 3800 class I p-sites we were able to increase the number of identifications by more than 21 times compared to recent literature. In addition, we were able to identify downstream targets of the only known Ser/Thr kinase of the S. *aureus* strain USA300, Stk1. Together this work allowed an extensive analysis of the bacterial phosphoproteome and will help to understand mechanisms involved in virulence and antibiotic resistance.

192

Going Beyond 11: TMTpro 16plex Regents for Higher Quantitative Proteomic Sample Multiplexing

Ryan Bomgarden[1], Sergei Snovida[1], Amarjeet Flora[1], Leigh Foster[1], Rosa Viner[2], Aaron Robitaille[2], Andrew Thompson Andrew Thompson [3], Karsten Kuhn [3], Ian Pike[3], John Rogers[1]

1. Thermo Fisher Scientific, Rockford, IL, United States

2. Thermo Fisher Scientific, San Jose, CA, United States

3. Proteome Sciences, London, WC1H 9BB, UK

Introduction

Thermo Scientific Tandem Mass Tag® (TMT®) Reagents enable concurrent identification and multiplexed quantitation of proteins in different samples using tandem mass spectrometry. The maximum number of samples that can be compared in a single experiment is current limited to eleven by the number of stable isotopes incorporated into the reporter and mass balance region. Here, we describe a new set of mass tags (TMTpro regents) with a proline-based reporter and longer linker which can support higher sample multiplexing for up to sixteen samples.

Methods

Cellular protein digests were labeled with TMT 11plex or TMTpro 16plex reagents using standard methods. Samples were analyzed by LC-MS to assess reporter ion fragmentation, labeling efficiency and peptide/protein identification rates. To assess protein quantitative accuracy and precision, HeLa protein digest was spiked with Pierce 6 protein mix at distinct ratios and labeled with different tags. Samples were analyzed on Orbitrap hybrid and tribrid instruments using MS2 or SPS-MS3 methods. Data analysis was performed using Thermo Scientific Proteome Discoverer™ 2.4.

Results and Discussion

Increasing the number of samples that can be compared in a single LC-MS experiment is highly desired to improve sample throughput, assess multiple experimental conditions and decrease missing values among replicates. We have designed a new tandem mass tag set that can be used to concurrently quantify up to sixteen protein samples. The new tag contains nine stable isotopes and a novel proline-based reporter with the same elemental composition and mass as the TMT reporter ion. Compared to TMT reagents, TMTpro reagents have the same peptide labeling efficiency, but require slightly lower collision energy for reporter ion fragmentation. We also demonstrate that TMTpro 16plex samples have similar peptide/protein identification rates compared to TMT 11plex samples, exhibit a 20-fold dynamic range of quantitation and are compatible with standard enrichment and fractionation workflows.

193

Diapasef: Toward The Ideal Mass Analyzer With Data-Independent Acquisition And Parallel Accumulation – Serial Fragmentation

Andreas-David Brunner[1], Max Frank[2], Annie Ha[2], Eugenia Voytik[1], Stephanie Kaspar-Schoenenfeld[3], Markus Lubeck[3], Heiner Koch[3], Scarlet Koch[3], Pierre-Olivier Schmit[4], Gary Kruppa[5], Oliver Raether[3], Ben C Collins[6], Ruedi Aebersold[6, 7], Hannes Röst[2], Matthias Mann[1, 8]

1. Max-Planck-institute of Biochemistry, Martinsried, Germany

2. Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada

3. Bruker Daltonik GmbH, Bremen, Germany

4. Bruker France, Wissembourg, France

5. Bruker Daltonics Inc, Billerica, MA, USA

6. Inst. f. Molekulare Systembiologie, ETH Zurich, Zurich, Switzerland

7. University of Zurich, Zurich, Switzerland

8. University of Copenhagen, Copenhagen, Denmark

Introduction

Data-independent acquisition (DIA) promises reproducible and accurate protein quantification across large sample cohorts. The mass spectrometer typically cycles through many isolation windows covering a broader *m/z* range of interest. Current methods utilize only about 1-3% of all available ions. In principle, all ions could be utilized by parallel ion storage and sequential release from the TIMS device into a Q-TOF mass analyzer. Here, we asked if the PASEF principle could be transferred to DIA.

Methods

Whole-cell proteomes extracted from a human cancer cell line were analyzed via nanoLC coupled to a prototype timsTOF Pro (Bruker). We adapted the instrument firmware to perform data-independent isolation of multiple precursor windows within a single TIMS separation (100ms). We tested multiple schemes for precursor selection window size and placement in the *m/z*-ion mobility plane. Analysis of the four-dimensional data space has been incorporated into OpenSWATH. For ion mobility-aware targeted data extraction, we used a project-specific library from 48 high-pH reverse-phase peptide fractions acquired with PASEF.

Results

As ion mobility and mass are correlated, a large proportion of the peptide ion current can be covered by scanning diagonal lines in the *m/z*-ion mobility space. We derived multiple diaPASEF acquisition schemes from the density distribution of about 130,000 precursors present in the library. TIMS provides highly precise measurements of collisional cross sections (CCS) with CVs << 1% in technical replicates. After linear alignment, CCS values extracted from the diaPASEF runs deviated < 2% from the library. In triplicate 120min runs of 200ng HeLa digest each, we quantified over 7000 proteins at a 1% FDR. Fragment ion-based quantification was very reproducible with a median CVs of 10% and a pairwise mean Pearson correlation >0.96.

Conclusion

The diaPASEF method captures and utilizes a very large proportion of the available ion current, approaching the ideal mass analyzer.

194

Heat and Beat: A one-pot rapid tissue sample preparation for proteomics in under an hour

Dylan Xavier[1], Clare Loudon[1], Peter G Hains[1], Phil J Robinson[1]

1. ProCan, CMRI, Westmead, NSW, Australia

Clinical proteomics on solid tissue samples from human tumour biopsies present multiple challenges. These include limited sample availability, reducing endogenous protease activity and challenges with efficiently solubilising and digesting proteins for high throughput mass spectrometry while achieving comprehensive proteome coverage. Protocols for cancer biopsy proteomics involves three sample preparation steps: tissue washing, lysis and digestion, followed by peptide clean-up by SPE. Our team have reported improvements in all three steps that greatly reduce processing time, enabling greater automation and higher sample throughput.

Our aim was to develop a universal protocol for tissue sample preparation that is quick and simple while improving protein solubilization, proteome coverage and reproducibility. Sample preparation with small (<2 mg) biopsies pose additional limitations including the risk of sample loss and introducing variability, partly due to multiple reagent addition/removal steps. We now report a one-pot method called heat and beat (HnB).

By identifying limitations and challenges of existing protocols, we developed a new protocol with fewer steps by achieving lysis, reduction, alkylation and digestion in essentially a single step which was preceded by a 7 min sample heat treatment step. This was achieved by several changes to the existing protocol including the combined use of a beadbeater for tissue homogenisation

followed by pressure cycling technology (PCT) using a Barocycler to further lyse and digest the sample. The procedure requires 38 minutes and was tested on 6 tissue types, across three tissue embedding techniques, fresh frozen, OCT embedded (FF-OCT), and FFPE.

Our results show that HnB is an efficient new one-pot procedure. The protocol significantly reduced endogenous protease activity, improves peptide yield by up to 3-fold and delivers a digestion efficiency of 85-90% from each sample source. We obtained an increase in the number of proteins identified of up to two-fold while reducing variability and processing time.

196

Pathway-level analysis of comprehensive proteogenomic and phosphoproteomic data to predict clinical outcomes

Jason McDermott[1], Tao Liu[1], Vlad Petyuk[1], Osama Arsad[1], Karin Rodland[1]

1. Pacific Northwest National Laboratory, Canby, OR, United States

Recent advances in the comprehensive cataloging of genomic landscapes in cancer have led to the widespread use of genomic and transcriptomic analyses to classify tumors and predict a targeted course of therapy. However, the clinical performance of targeted therapies based solely on genomic analysis is disappointing, with only a subset of patients responding and minimal gains in median overall survival. This is not surprising, as drugs generally act on proteins, not genes, and the downstream effects of genetic mutations are often manifested by a cascade of changes in protein interactions within pathways. Integration of proteomic and genomic data is key to elucidating the complex biology determining tumor responses and resistance to cancer therapeutic agents, particularly by improving the ability to identify and characterize alterations in cancer-relevant pathways triggered by genomic alterations. In particular, the analysis of post-translational modifications such as phosphorylation provides pathway insights not obtainable from genomic analysis alone. By correlating proteogenomic and phosphoproteomic measurements from clinical trial patients with their accompanying clinical outcome data we are able to address clinical questions related to predicting drug response, toxicity, and resistance. The ultimate goal is of these studies is to address why a patient predicted to respond to a therapy by genomic data did not, and whether proteomics alone or in combination with genomics is a better predictor of response than genomics alone.

197

Multiplex In-Solution Protein Array (MISPA) for high throughput, quantitative profiling of protein interactions and detection of immune responses to pathogen induced cancers

Joshua Labaer[1], Femina Rauf[1], Yanyang Tang[1], Radwa Ewaisha[1], Capria Rinaldi[1], Yunro Chung[1], Jin Park[1], Karen Anderson[1], Brianne Petritis[2]

1. Arizona State University, Tempe, ARIZONA, United States

2. Raybiotech, Atlanta

Understanding how proteins interact with each other to exert biological function has been essential for elucidating the causes and developing treatments for cancer. Cancer researchers continue to devote prodigious effort towards developing and refining methods for detecting and quantifying protein-protein interactions (PPI). Although numerous methods are available to study PPIs most technologies suffer from high false positive rates, missed biophysical interactions, strong biases towards abundantly expressed genes or detecting protein complexes, low throughput and lack of quantification. Moreover, most methods are not compatible for testing clinical material such as immune responses in patient serum. Therefore, there is a great need for multiplexed, robust technologies for quantitative profiling of protein interactions and early detection of immune responses. We have developed an innovative next-generation, in-solution protein microarray platform MISPA which exploits the extraordinary potential of next generation DNA sequencing (NGS) to address unanswered questions in cancer biology such as identifying unknown protein interactions, and early screening. MISPA is a solution phase functional protein microarray platform where each protein is uniquely barcoded with DNA. The barcoded protein cocktail is then interacted *in* solution with a query protein or clinical sample. We tested the feasibility of MISPA as a robust, multiplexed diagnostic tool for early screening of HPV positive oropharyngeal carcinomas (OPC). The proteomes of 12 different serotypes of HPV (96 antigens) was barcoded and tested against HPV positive OPC serum and control samples**.** MISPA showed distinct enrichment of certain HPV antibodies (HPV16E1, HPV16E2, HPV16E6 and HPV16E7) in the OPC patient samples compared to controls. We also noticed enriched immune response against other virulent HPV strains. The assay showed high reproducibility with an average %CV of 3.4. It also demonstrated an expanded dynamic range (> 10^4) with a 30-600 fold signal over background.

198

Improved Survival Prognostication of Node-Positive Malignant Melanoma Patients applying Shotgun Proteomics Guided by Histopathological Characterization and Genomic data

Lazaro Betancourt[1], Krzysztof Pawłowski[1], Jonatan Eriksson[1], Marcell Szasz[1], Shamik Mitra[1], Indira Pla[1], Charlotte Welinder[1], Henrik Ekedahl[1], Per Broberg[1], Roger Appelqvist[1], Maria Yakovleva[1], Yutaka Sugihara[1], Kenichi Miharada[1], Christian Ingvar[1], Lotta Lundgren[1], Bo Baldetorp[1], Håkan Olsson[1], Melinda Rezeli[1], Elisabet Wieslander[1], Peter Horvatovich[1], Johan Malm[1], Göran Jönsson[1], György Marko-Varga[1]

1. Lund University, Lund, SKANE, Sweden

Maligant Melanoma is one of the most common deadly cancers, and robust biomarkers are still needed, e.g. to predict survival and treatment efficiency.

The European Cancer Moonshot Lund Center has undertaken a metastasis tumor tissue study together with the South Swedish healthcare region, comprising 10 hospitals with a population of 2 Mill.

The Melanoma study undertakes protein expression analysis of one hundred eleven melanoma lymph node metastases that were surgically isolated, following strict guidelines, using high resolution mass spectrometry is coupled with in-depth histopathology analysis, clinical data and genomics profiles. This wide-ranging view of protein expression allowed to identify novel candidate protein markers that improved prediction of survival in melanoma patients. We were able to distinguish long survivors from the short survivor patient group. Some of these prognostic proteins have not been reported in the context of melanoma before, and few of them exhibit unexpected relationship to survival, which likely reflects the limitations of current knowledge on melanoma and shows the potential of Proteomics being integrated within clinical cancer research.

199

Delineating mechanisms that confer resistance to kinase inhibitors in head and neck squamous cell carcinoma and melanoma

Keshava Datta[1], Ankit Jain[2], Krishna Patel[2], Aditi Chatterjee[2], Harsha Gowda[1, 3, 4]

1. QIMR Berghofer, Brisbane, QLD, Australia

2. Institute of Bioinformatics, Bangalore, Karnataka, India

3. Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia

4. School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia

One of the significant barriers for effective anti-cancer therapy is emergence of drug resistance. In almost all cases, acquired resistance and relapse often follow an initial effective response to a targeted therapeutic agent. Understanding molecular mechanisms that confer resistance is vital to develop novel therapies to achieve durable response. Kinase inhibitors are some of the widely used targeted therapeutic agents to treat several cancers. In this study, we investigated mechanisms associated with erlotinib resistance in head and neck squamous cell carcinoma (HNSCC) and vemurafenib/dabrafenib resistance in melanoma. To investigate resistance mechanisms, drug resistance clones of HNSCC cell line and melanoma cell lines were derived by subjecting them to selection pressure under chronic drug treatment. Exome sequencing was carried out to identify mutations and copy number alterations associated with resistance clones. Global proteomic and phosphoproteomic profiling was carried out to identify altered signaling mechanisms associated with drug resistant clones. Our data revealed several genomic alterations and proteomic changes associated with erlotinib resistant cells in HNSCC. Integrated analysis of genomic and proteomic data showed activation of MAP kinase pathway as a potential mechanism of erlotinib resistance in HNSCC. Growth inhibition studies using MAP2K1 inhibitor showed potent activity against erlotinib resistant cells. We observed cMET amplification and pathway activation in BRAF inhibitor resistant melanoma cells. MAP2K1 inhibition might be a useful strategy to overcome erlotinib resistance in HNSCC. Similarly, targeting cMET might be beneficial in overcoming BRAF inhibitor resistance in melanoma. Our study demonstrates advantages of integrating genomic and proteomic data to characterize drug resistance mechanisms in cancers.

3. Crook et al (2019) PLoS Comp. Bio. - doi: 10.1371/journal.pcbi.1006516
4. Queiroz et al (2019) Nature Biotechnology – doi:10.1038/s41587-018-0001-2

158

Reverse Engineering to Identify the Next Therapeutic in Heart Failure with Preserved Ejection Fraction?

Jennifer Van Eyk[1], Peter Kilfoil[1], Ahmed Ibrahim[1], Eduardo Marban[1]

1. Cedars-Sinai, Los Angeles, CA, United States

In Europe and North American, someone is diagnosed with Heart Failure (HF) ~ every 40 seconds. In Australia, 30% of all deaths are due to cardiovascular disease which most often ends in HF. In fact, heart disease is the single leading cause of death in Australia. Unlike traditional HF, where the heart is unable to eject sufficient blood (HF with reduced ejection fraction), where a number of effective pharmacological and medical device interventions are available, there are currently no effective treatments for HF where the heart has preserved ejection fraction (HFpEF). HFpEF has only recently been recognized and it occurs predominately in women. We have recently shown that a cell therapy, using cardiac derived cardiosphere (CDCs), in small and large animal models of HFpEF can stop and reverse some of the HFpEF phenotype. This has led to an ongoing clinical trial (Regress-HFpEF) of CDCs in HFpEF patients. Yet, the cardiac mechanistic changes induced by CDC (or the CDC-derived exosomes) and their active biologics agents are not known. This study aims to reverse engineer the CDC-induced tissue effects with specific exosome biologics to potentially identify the active therapeutic agents and pathways. Methods and Results: Dahl salt-sensitive rats fed high-salt diet, with echocardiographically verified diastolic dysfunction, were randomly assigned to either intracoronary CDCs or placebo. Dahl rats receiving low salt diet served as controls. Total protein quantity, and isoforms and phosphorylation status of left ventricular tissues (n=6/group) were quantified by mass spectrometry. A remarkable ~40% of the proteomic changes induced when HFpEF hearts are exposed to CDCs are alterations in the transcriptional and translational subproteome - with additional changes in protein folding and stability. 5 upstream cellular regulators account for ~45% of transcriptional and translational changes induced by CDCs in the static proteome of the HFpEF

heart (MYC, TGFβ1, TP53, mTOR/AKT, and HNF4A). These same 5 cellular regulatory (plus HTT and SRF) target 75% of newly synthesized proteins (assessed using L-Azidohomoalanine labeling of isolated cardiac myocyte form HFpEF animals) within the first 3 hours of treatment - interestingly, these regulators are not at play in myocytes from control rats. In addition, all 32 phosphorylation sites involved in transcription/translation that were altered with CDC treatment had predicted kinases that are regulated by these same 5 cellular regulators. Finally, detailed analyses of isolated CDC exosomes reveal 31 proteins and 5 miRNAs within the exosome cargo that can be directly linked to these same 5 cellular regulators. Individual miRNAs are able to recapitulate subset of CDC exosome activity.
Conclusion: Five potential cellular regulators that may mediated the widespread CDC-induced reprograming of the HFpEF heart are linked to a number of proteins and miRNAs within the CDC-exosome cargo leading to the possibility of these being new therapies for reversal of HFpEF.

164

Rapid, sensitive and site-specific phosphoproteome profiling of EGFR signaling by data-independent acquisition

Jesper V Olsen[1]

1. University of Copenhagen, Copenhagen, Denmark

Tyrosine phosphorylation regulates complex, multi-layered signaling networks with broad implications in (patho)physiology. To decipher these signaling networks directly in tissue samples we analyzed the EGF-dependent in-vivo phosphotyrosine signaling in lung tissue to a depth of more than one thousand phosphotyrosine sites. A phosphopeptide based proteomic screen identified proteins recruited to the EGF-regulated phosphotyrosine sites in lung tissue lysates. We demonstrate how cancer mutations near phosphotyrosine sites can introduce molecular switches that rewire signaling networks, and reveal how such a molecular switch is responsible for oncogenic properties of an EGFR lung cancer mutant. Introduction of cutting-edge LC-MS/MS instrumentation and data independent acquisition (DIA) enables scalable, rapid and high-throughput analysis of phosphotyrosine site interactions and phosphoproteomes. Our approach enables phospho-signaling investigations in tissues for in-depth mechanistic insights into oncogenic rewiring of signaling networks.

165

Global redox proteome and phosphoproteome analysis reveals novel insights into the Insulin Signaling Network

Sean Humphrey[1], Daniel Fazakerley[1], Pengyi Yang[1], Tim Zhiduan Su[2], James Burchfield[1], David James[1, 3]

1. Charles Perkins Centre, School of Life & Environmental Science, The University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, School of Life & Environmental Science, The University of Sydney, Sydney, NSW, Australia

3. School of Medicine, The University of Sydney, Sydney, NSW, Australia

We have previously used phosphoproteomics to map the insulin signalling network in adipocytes revealing more than 2,000 insulin responsive phosphosites. More recently, we have used similar approaches to map changes in the phosphoproteome across multiple models of insulin resistance revealing major changes in the entire signalling network across all models. Because insulin resistance also involves changes in oxidative stress, here we endeavoured to find the intersection between the redox and phospho-proteomes. Our integrative analysis revealed widespread and complex crosstalk between oxidative stress-induced cysteine oxidation and phosphorylation-based signalling. In particular, we observed that oxidation of key regulatory nodes (Akt, mTOR and AMPK) influences the fidelity of signalling via these nodes rather than absolute activity, providing an underappreciated interplay between these modifications. Mechanistic analysis of the redox regulation of Akt identified that two cysteine residues in the pleckstrin homology domain of Akt (C60 and C77) are likely reversibly oxidised and this regulates recruitment of Akt to PIP3 at the plasma membrane via reconfiguration of the PH domain structure. Overall, these multi-omics datasets provide insight into how redox signalling driven by oxidative stress interacts with protein phosphorylation and should serve as a useful resource for dissecting oxidative stress-induced PTMs and understanding their contribution to a variety of diseases.

177

The use of proteomics to identify blood-based protein changes in childhood that are associated with increased risk for later psychosis

Melanie Foecking[1]

1. Royal College of Surgeons in Ireland, Dublin, DUBLIN, Ireland

The identification of early biological changes associated with psychosis is important as it may provide clues to the underlying pathophysiological mechanisms and early diagnosis at treatment of those at risk of future disorder is associated with improved outcome.

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective general population cohort, and a rich resource of demographic, environmental, and clinical data on the individuals involved. We studied a subsample of the cohort who participated in psychiatric assessment interviews at age 12 and 18, and who provided plasma samples at age 12.

A multi-omics approach was used, including mass spectrometry based proteomics, lipidomics and ELISA.

Our findings implicate the blood complement and coagulation system in psychotic experiences and disease from age 12 to age 18. The complement cascade is a major component of the immune defense against infection, and there is increasing evidence for a role of dysregulated complement in major psychiatric disorders.

These findings demonstrate how this multi-omics approach contributes to a better understanding of the molecular pathways dysregulated in the blood during childhood before the development of different forms of psychosis.

178

Comparative glycoproteomic analyses of cerebrospinal fluids reveal novel molecular players in Alzheimer's disease

Lingjun Li[1], Xiaofang Zhong[1], Zhengwei Chen[1], Qinying Yu[1], Cynthia Carlsson[1], Ozioma Okonkwo[1]

1. University of Wisconsin - Madison - Madison, WI, Madison, WISCONSIN, United States

Alzheimer's disease (AD) is the most common form of dementia in the elderly population. While several peptide and protein biomarkers in cerebrospinal fluid (CSF) have been used for AD diagnosis, an unequivocal diagnosis in the early phases of AD is still lacking. Furthermore, the discovery and establishment of reliable biomarkers capable of monitoring progression and degree of cognitive impairment as well as potential efficacy of therapy remains a major challenge. To address this challenge, we are developing multiplexed isobaric and isotopic tagging strategies to discover, identify and evaluate candidate biomarkers of AD in CSFs obtained from asymptomatic cognitively-healthy middle-aged adults, older cognitively-normal adults, and patients with mild cognitive impairment (MCI) and AD. In the current study, we developed global glycoproteomics approach that combines enhanced N-glycopeptide sequential enrichment by hydrophilic interaction chromatography (HILIC) and boronic acid enrichment with electron transfer and higher-energy collision dissociation (EThcD) for large-scale intact glycopeptide analysis. In total, 3596 intact N-glycopeptides from 676 N-glycosites and 358 N-glycoproteins were identified in CSF. To our knowledge, this was the largest site-specific N-glycoproteome dataset reported for CSF so far. As accumulating evidence suggests that aberrant glycosylation is implicated in AD pathogenesis, we perform large-scale comparative glycoproteomic analysis of CSF samples from control and AD patients, revealing distinct glycosylation patterns and dynamic changes of certain glycoforms. 1519 intact N-glycopeptides mapping to 178 glycosites on 107 glycoproteins were quantified via the 12-plex DiLeu (*N,N*-dimethyl leucine) tagging strategy. 19 N-glycopeptides were dysregulated in the progression of AD. Moreover, our data revealed that 25 aberrant N-glycopeptides in MCI were modified into 54 different glycoforms at the same site of a given protein in AD, termed as disease stage-specific N-glycopeptides. Collectively, our approach enables the elucidation of glycoprotein microheterogeneity and correlation of subtle changes in glycan structural repertoire to disease progression of AD.

183

Microbiome in diseases and health

Daniel Figeys[1]

1. University of Ottawa, Ottawa, ON, Canada

The microbiome is emerging as a prominent player affecting human health, and associated with various diseases. Understanding the roles that the microbiome plays in health and diseases and how it reacts to external factors, such as drugs, requires not only to assess its compositions but also its functional changes. Metaproteomics measures quantitative functional changes in the microbiome and therefore will be an important tool to studies the human microbiome. The microbiome is very complex and new tools and software are needed to elucidate its function. We are developing new metaproteomics tools to study the microbiome and are applying these tools to study host-microbiome interactions and microbiome-drug interactions. Bioinformatics is key to the development of metaproteomics. Our lab introduced in 2017 MetaLab to identify and quantify peptides/proteins from gut microbiome. We will launch version 2.0 of MetaLab and iMetaLab a HUPO. We are also interested in the development of in vitro assays to rapidly assess the effects of compounds on individual microbiome. We will present recent development of an in vitro microbiome assay coupled to metaproteomics to understand the effects of different compounds on individual human microbiomes.

184

Histidine phosphorylation: a new dimension in the phosphoproteome

Clement Potel[1], Miao Hsia Lin[1], Nadine Prust[1], Simone Lemeer[1]

1. Utrecht University, Utrecht, UTRECHT, Netherlands

Histidine phosphorylation is known to play major biological roles in prokaryotes, notably in bacterial cell metabolism or in primitive signal transduction pathways. In recent years compelling evidences that histidine phosphorylation also plays important biological roles in eukaryotes gradually accumulated. Nevertheless, the lack of adequate analytical methods to study the unstable histidine phosphorylation constitutes to this day the biggest hurdle to the large-scale study of histidine phosphorylation and consequentlyto the unraveling of the full extent of its biological roles.

Here we present - to this day - the sole method capable of enriching histidine phosphorylation at the peptide level and providing site-specific localization and quantification. We demonstrate that in the right conditions, IMAC can be used to enrich histidine phosphorylation at the peptide level. In addition, we report the use of the diagnostic phosphohistidine immonium ion as a tool to unambiguously localize histidine phosphorylation. Along the way we identified nucleic acid containing biomolecules as the main contaminants present in phospho-enriched samples after affinity chromatography. We tackled this problem by developing a robust and efficient sample preparation protocol, incorporating nucleic acid enzymatic digestion and protein precipitation. This optimized protocol enabled a 10-fold increase of the number of identified phosphosites in both gram negative as well as gram negative

bacteria. The coverage allowed for a comprehensive analysis of the changes in phosphoproteome in response to antibiotic treatment and in antibiotic resistance. All together these recent methodological advances presented here have led to a better insight into bacterial signaling and show that phosphorylation is far more abundant than anticipated in bacteria. In addition, our technical advances could answer a long-standing question: is histidine phosphorylation also playing a major role in higher organisms?

189

How do we make quantitative proteomics quantitative?

Michael MacCoss[1], Brian Searle[2], Lindsay Pino[1], Deanna Plubell[1], Danielle Faivre[1], Andy Hoofnagle[1], Gennifer Merrihew[1], Richard Johnson[1], Brendan Maclean[1]

1. University of Washington, Seattle, WA, United States

2. Institute for Systems Biology, Seattle, WA, USA

Our goal is to develop high throughput method for sampling peptides with a mass spectrometer that can be used as a quantitative measure of the phenotype. To do this we would like a tandem mass spectrometry (MS/MS) method that can comprehensively sample all peptides in a sample continuously throughout the chromatographic elution. MS/MS acquired using data independent acquisition (DIA) offers significant advantages in terms of selectivity, sensitivity, and dynamic range over a single stage of mass analysis. MS/MS has significant technical advantages over MS1 analysis and we are just now in a situation where we can get high selectivity (<4 m/z), across a majority of the mass range (i.e. 400-1000 m/z), and using a rapid duty cycle (<3 sec). That said, this does not mean that there are not substantial challenges to overcome. For example, we need methods to assess whether the peptide measurements are quantitative versus just qualitative. Additionally, global methods like proteomics struggle significantly with signal calibration -- making it difficult to compare quantitative measurements between batches, labs, and instrument platforms. Given the prevalence of complex proteoforms we need to think carefully about what the desired outcome is of a quantitative proteomics experiment using bottom-up methodologies. Finally, while most labs feel it is important to measure as many proteins and peptides as possible, the complications associated with doing this is non-trivial -- ultimately with an increase in the number of analytes measured increases the multiple testing burden and the number of samples required to have the same statistical power. The talk will present the current state of the art of performing quantitative proteomics using DIA. I will present use cases of what we can do, where we think the limitations are, and what work is being done to improve the methods further.

190

Differential Proteomics and Lipidomics in Niemann-Pick Disease, Type C

Stephanie M. Cologna[1], Melissa R Pergande[1], Koralege C Pathmasiri[1]

1. University of Illinois at Chicago, Chicago, IL, United States

Niemann-Pick Disease, Type C (NPC) is a fatal, genetic, neurodegenerative disorder with no FDA-approved therapy. As a result of the genetic defect, unesterified cholesterol and glycosphingolipids accumulate in the endo-lysosomal system. A hallmark of this disease includes progressive, cerebellar degeneration however the link between lipid accumulation and neuron death is not fully understood. In this work, we sought to determine altered proteins and lipids in NPC with a focus on the cerebellum. Utilizing standard flow chromatography and Jet Stream ESI technology, we identified differential proteins and lipids including those involved in phosphoinositide and fatty acid metabolism. This presentation will include a description of the analytical methodology and examples of altered proteins and lipids found in NPC.

191

Rapid and robust high throughput cancer proteomics across multiple instruments in a single facility

Phillip J Robinson[1], Rebecca Poulos[1], Roger Reddel[1], Qing Zhong[1]

1. Children's Medical Research Institute, Children's Medical Research Institute, The University of Sydney, Wentworthville, NSW, Australia

Analysing the proteome of large numbers of solid tissue cancer samples presents challenges such as small samples, tissue fixation reversal, data management, reproducible performance across multiple instruments over time, and large scale data normalisation and analysis (Tully et al 2019). Acquisition of proteome wide data can be relatively reproducible, and highly qualitative and quantitative across different laboratories using SWATH-MS. Is this achievable on a large scale and in a single facility?

The ACRF Centre for the Proteome of Human Cancer (ProCan®) was established to survey the proteomic landscape of all cancers. A suite of Barocyclers and 6 Sciex 6600 tripleTOF mass spectrometers was established in a single facility to operate 24/7 as a single unit. Around 10,000 proteomes have been collected to date.

In processing such large numbers, we encountered many operational bottlenecks such as tedious tissue washing, preparation and digestion (Tully et al 2019). We have developed entirely new workflows that reduce the end-to-end sample processing time from 2-6 hours to under one hour, for both removal of OCT from fresh frozen samples, or dewaxing and reversal of FFPE embedding. We achieved identical peptide and protein numbers from each. By running technical replicates between different MS instruments instead of on the same instrument we identified and removed technical missingness.

To demonstrate reproducibility, we performed a multi-instrument, longitudinal SWATH-MS assessment. Using six instruments, we acquired 1,560 technical replicates of a single set of experimental samples spanning a dilution series comprising a mixture of three biologically distinct tissues. We acquired data bi-daily for a week, weekly for a month, and monthly for 4 months, in a facility with varying maintenance schedules and minimal instrument down-time to reflect a real-world use case. We developed a new statistical method that can normalise these data with a significant improvement over existing methods. The data reveal strong linearity across the tissue dilution, outstanding reproducibility across time and instrument, and that machine learning can accurately predict the concentration of one tissue mixed within another.

Together this establishes how a single facility can effectively function in true high throughput mode, and also integrate and analyse large proteomic data sets across multiple instruments. This enables the reproducible and high-throughput proteomics required to realise the vision of precision medicine.

1. Tully,B., Balleine,R.L., Hains,P.G., Zhong,Q., Reddel,R.R. & Robinson,P.J. (2019) Addressing the challenges of high-throughput cancer tissue proteomics for clinical application: ProCan®. Proteomics e1900109, PMID:31321850

195

Onco-proteogenomics: toward a more complete understanding of cancer biology

Bing Zhang[1]

1. BAYLOR COLLEGE OF MEDICINE, Houston, TX, United States

Cancer is a disease of the genome, but many processes downstream of the genome can affect the tumor phenotype. Using proteogenomics data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC), I will present examples on how proeogenomic integration can expand our knowledge on cancer genes, prioritize cancer drivers, clarify puzzling genomic observations, and correct misinterpreted gene functions. I will also discuss our recent works on using proteogenomics to identify tumor antigens and understand immune evasion mechanisms.

200

Integrative -Omics Reveals Novel Targets Underlying the Pathomechanisms of Uterine Fibroids and Associated Heavy Menstrual Bleeding

Darragh P O'Brien[1], Kavita S Subramaniam[2], Thomas T Tapmeier[2], Chen-Yi Wang[3], Simon Davis[1], Adam Cribbs[3], Jessica Malzahn[3], Thomas M Zollner[4], Bianca De Leo[4], Maik Oberdorf[4], Joerg Muller[4], Martin Fritsch[4], Martin Philpott[3], Krina T Zondervan[2], Udo Oppermann[3], Adrian Harris[5], Christian M Becker[2], Benedikt M Kessler[1]

1. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

2. Endometriosis CaRe Centre, Nuffield Department of Women's & Reproductive Health, University of Oxford, Oxford, United Kingdom

3. Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

4. Bayer AG, Berlin, Germany

5. Department of Oncology, University of Oxford, Oxford, United Kingdom

Background: Uterine fibroids (UF's) are benign tumours affecting up to 80% of women of reproductive age. UFs produce factors that induce angiogenesis with a highly irregular vasculature. Approximately 30% of fibroid patients suffer severe symptoms including pelvic pain and heavy menstrual bleeding (HMB). Although mutations in MED12 or HMGA2 account for the majority of UF occurrence, the processes by which these lead to UF's and HMB remain poorly understood. Using an array of -Omics technologies, we undertake the first comprehensive integrative biology study of the pathomechanisms and vascularisation underlying UF formation and maintenance.

Methodologies: Fibroid, endometrium, myometrium, pseudocapsule, and healthy control samples (n = 130 tissues, across 45 donors) were obtained from patients undergoing surgery at the John Radcliffe Hospital, Oxford University Hospitals, UK. Gene and protein expression profiles between tissue types were investigated by RNA-Seq and quantitative proteomics. Spatial, cell-type resolved proteomic information was generated across tissues using Laser-Capture Microdissection Mass Spectrometry (LCM-MS) of uterine CD31+ cells. Candidate angiogenic factors were validated on FFPE tissues by immunohistochemistry and tested *in vitro* on human uterine microvascular endothelial cells. Their angiogenic potential was evaluated in proliferation and tube formation assays. Vascularised UF architecture was imaged by light sheet microscopy.

Findings: Integration of uterine tissue transcriptomics (>20,000 genes), proteomics (~5,500 proteins), and LCM-MS (~3,000 proteins), revealed altered biological pathways between different uterine tissues, *e.g*., Wnt signalling, which may serve as novel druggable targets for the treatment of UFs and HMB. Immunohistochemistry of UF vasculature confirmed the presence of candidate angiogenic factors, which were validated using tube length, endothelial mesh, and cell proliferation assays. Cleared whole mice uterus showed a distinct blood vessel pattern upon visualisation with 3D reconstruction imaging.

Concluding Statement: Using our multi -Omics approach, we have identified several exciting avenues to better understand the pathomechanisms of UFs and associated HMB.

201

Cellular Cartography at Molecular Detail: *in Situ* Crosslinking Mass Spectrometry

Juri Rappsilber[1]

1. TU Berlin, Berlin, DEUTSCHLAND (DEU), Germany

Bridging scales from atomic resolution to entire cells is a technical and computational challenge. Over the past decade, crosslinking mass spectrometry (CLMS) has developed into a robust and flexible tool that provides medium-resolution structural information. CLMS data provide a measure of the proximity of amino acid residues and thus offer information on the folds of proteins and the topology of their complexes. Here, we highlight notable successes of this technique as well as common pipelines. Novel CLMS applications, such as in-cell crosslinking, probing conformational changes and tertiary-structure determination in complex mixtures, are now beginning to make contributions to molecular biology and the emerging fields of structural systems biology and interactomics.

202

Amine-selective Di-ortho-phthalaldehyde (DOPA) Cross-linking Captured the Conformational Change Associated with the Unfolding of Ribonuclease A

Meng-Qiu Dong[1]

1. National Institute of Biological Sciences, Beijing, China

Chemical cross-linking of proteins coupled with mass spectrometry analysis (CXMS, aka XL-MS) has become an established method for probing the conformation of a protein and for mapping the interface of protein-protein interactions. Central to this technology are chemical cross-linkers. The most popular cross-linkers such as DSS, BS3, and DSSO are all NHS esters, which react with protein amino groups relatively slowly over 10-60 minutes while in competition with the hydrolysis reaction of NHS esters. Concerned with the imperfections of NHS ester cross-linkers, we developed a new class of amine-selective nonhydrolyzable di-ortho-phthalaldehyde (DOPA) cross-linkers. DOPA cross-linking of proteins takes only 10 seconds under near physiological conditions. Besides, DOPA works at low pH, low temperature, or in the presence of high concentrations of denaturants such as 8 M urea or 6 M guanidine hydrochloride (GdnHCl), making it possible to capture conformational changes of proteins during the unfolding or refolding process over a time scale of minutes. The test results on six model proteins showed that 80% or more of the DOPA cross-links are compatible with protein crystal structures. Using two heterodimeric complexes of weakly interacting subunits, we demonstrated that DOPA2 (spacer arm 17.5 Å, maximal Cα-Cα distance 30.2 Å) fixed weak or transient protein-protein interactions better than DSS. We also demonstrated that DOPA2 indeed captured the structural change associated with the unfolding of ribonuclease A induced by either 8 M urea or 6 M GdnHCl. Therefore, DOPA not only adds to CXMS a class of fast, nonhydrolyzable lysine cross-linkers for analysis of protein-protein interactions, but also expands the application of CXMS to analysis of protein folding/unfolding intermediates.

203

Crosslinking mass spectrometry analysis of the yeast nucleus reveals extensive protein-protein interactions not detected by systematic two-hybrid or affinity purification-mass spectrometry

Tara Bartolec[1], Daniela-Lee Smith[1], Ignatius Pang[1], Gene Hart-Smith[1], Joshua Hamey[1], Marc Wilkins[1]

1. Systems Biology Initiative, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW

Saccharomyces cerevisiae has the most comprehensively characterised protein-protein interaction network, or interactome, of any species. This has been generated through multiple, systematic studies of protein-protein interactions by two-hybrid techniques and through multiple, systematic studies of affinity-purified protein complexes. Despite the quality of this interactome, a question remains as to the extent of interactions that are yet to be detected. These may include interactions that are weak or transient, or those involving proteins that are not amenable to study as two-hybrid fusion proteins. Large scale crosslinking-mass spectrometry (XL-MS) has become possible through the development of cleavable crosslinkers, MS3 methods for peptide analysis and new software. To date, these methods have been applied to analysis of mammalian cells and tissues, to some organelles and to a number of bacterial systems however they have not been applied to yeast. Here we used XL-MS to study intact yeast nuclei, using DSSO, MS3 analysis and XlinkX. Linear peptides identified ~3,300 nuclear-associated proteins, generating the most comprehensive proteome of this yeast organelle to date. A total of ~2,500 crosslinked spectral matches were found, resulting in ~1,350 unique lysine-lysine crosslinks. Of these, 65% were intra-protein crosslinks and 35% were inter-protein crosslinks. Approximately one-third of intralinks mapped to PDB structures, 93% which were less than the <30Å distance constraint. Interestingly, intralinks were found for 437 proteins with no existing structural data. In some cases, such intralinks could refine *ab initio* structural models. Application of stringent score cutoffs to interlinks yielded a high confidence nuclear interactome. Strikingly, almost half of the interactions were not previously detected by two-hybrid or AP-MS techniques. Multiple lines of evidence existed for many such interactions, whether through replicates, literature or ortholog interaction data. We conclude that XL-MS is a powerful means to measure protein-protein interactions that can complement two-hybrid and AP-MS techniques.

204

Complex-centric proteome profiling in one day with SEC-SWATH-MS and short gradient analysis

Ben C Collins[1], Peng Xue[2], Claudia Martelli[2], Isabell Bludau[2], George Rosenberger[3], Charlotte Nicod[2], Moritz Heusel[4], Andrea Fossati[2], Federico Uliana[2], Fabian Frommelt[2], Yansheng Liu[5], Rudolf Aebersold[2]

1. Queen's University Belfast, Belfast, ANTRIM, United Kingdom

2. Department of Biology, ETH Zurich, Zurich, Switzerland

3. Department of Systems Biology, Columbia University, New York, United States

4. Department of Clinical Sciences, Lund University, Lund, Sweden

5. Department of Pharmacology, Yale University School of Medicine, West Haven, United States

Introduction

Cellular functions are rarely attributable to a single molecule but rather to sets of molecules organized into modules such as protein complexes. Methods based on protein correlation profiling (PCP), such as size exclusion chromatography-SWATH mass spectrometry (SEC-SWATH-MS), provide rich information on the state of cellular protein complexes but remain impractical, as each biological sample requires weeks of measurement time.

Objectives

We aimed to establish an SEC-SWATH-MS strategy operating at a rate of 1 biological sample analysed per day while minimizing loss of information.

Materials & Methods

Native protein extracts from biological triplicates of HeLa cells (CCL2 or Kyoto variants) were separated by SEC into 70 fractions and analysed by SWATH-MS. We used a 21-minute gradient on the EvoSep One HPLC system that uses embedded gradients to minimize overhead. Data were analysed using OpenSWATH and Spectronaut and protein complex reorganization was determined using CCprofiler and SECAT.

Results

Over 3 replicates each from 2 conditions we detected 4,065 proteins with SEC elution features of which 2,026 had multiple features indicating that the protein is present in multiple distinct assembly states. Querying the data for CORUM complexes, we saw evidence for 590 complexes (5% FDR). Comparison with a prior study using 2 hour gradients showed we retain the majority of information at the protein (99%) and protein complex (95%) levels with >10-fold increase in speed. Quantitative analysis showed that while many proteins are differentially abundant between HeLa variants, changes in protein complexes are less prevalent consistent with complex level-buffering effects. Preliminary data using the newly developed diaPASEF method for this analysis will also be presented.

Conclusion

We have optimized SEC-SWATH-MS to analyse 1 biological sample per day. The dramatic increase in throughput enables the characterization of proteome complex organization with minimal information loss, extended scope, and broad applicability.

205

Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS)

Eneko Villanueva[1], Rayner ML Queiroz[1], Tom Smith[1], Maria Marti-Solano[2], Mie Monti[1], Mariavittoria Pizzinga[3, 4], Dan-Mircea Mirea[1], Manasa Ramakrishna[4], Robert F Harvey[3], Veronica Dezi[4], Gavin H Thomas[5], Anne E Willis[4], Kathryn S Lilley[1]

1. Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, Cambridge, UK

2. MRC Laboratory of Molecular Biology, Cambridge, UK

3. MRC Toxicology Unit, University of Cambridge, Leicester

4. MRC Toxicology Unit, University of Cambridge, Leicester, UK

5. Department of Biology, University of York, York, UK

Existing high-throughput methods to identify RNA-binding proteins (RBPs) are based on capture of polyadenylated RNAs and cannot recover proteins that interact with nonadenylated RNAs, including long noncoding RNA, pre-mRNAs and bacterial RNAs. We present orthogonal organic phase separation (OOPS), which does not require molecular tagging or capture of poly-adenylated RNA, and apply it to recover cross-linked protein–RNA and free protein, or protein-bound RNA and free RNA, in an unbiased way. We validated OOPS in HEK293, U2OS and MCF10A human cell lines, and show that 96% of proteins recovered were bound to RNA. We show that all long RNAs can be cross-linked to proteins, and recovered 1,838 RBPs, including 926 putative novel RBPs. OOPS is approximately 100-fold more efficient than existing methods and can enable analyses of dynamic RNA–protein interactions. We also characterize dynamic changes in RNA–protein interactions in mammalian cells following nocodazole arrest, and present a bacterial RNA-interactome for Escherichia coli. OOPS is compatible with downstream proteomics and RNA sequencing, and can be applied in any organism.

206

High-throughput protein-protein interaction profiling for clinical applications

Claudia Martelli[1], Peng Xue[1, 2], Charlotte Nicod[1], Charlotte Seng[1, 3], Evan Williams[1], Isabell Bludau[1], George Rosenberger[4], Andrea Fossati[1, 5], Chen Li[1, 6], Fabian Frommelt[1], Ben C Collins[1], Ruedi Aebersold[7]

1. Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, SWITZERLAND, Switzerland

2. Institute of Biophysics , Chinese Academy of Sciences, Beijing, China

3. Interfaculty Institute of Biochemistry, Eberhard Karls University of Tübingen, Tübingen, Germany

4. Department of Systems Biology, Columbia University, New York, USA

5. Institute of Molecular Health Sciences, ETH Zurich, Zurich, Switzerland

6. Monash Biomedicine Discovery Institute, Monash University, Melbourne, Australia

7. Faculty of Science, University of Zurich, Zurich, Switzerland

Proteins control and catalyze most of the biochemical functions of cells through their activities and intricate interactions. Generally, neither the abundance nor the interactions of proteins are precisely predictable from e.g. genomic or transcriptomic information. Defining a comprehensive picture of the abundance and subunit composition of protein complexes for a given physiological/pathological state is therefore fundamental for basic and translational research, opening new possibilities for the definition of new clinical biomarkers and/or therapeutic targets.

Our group published an experimental/computational pipeline for the complex-centric analysis of protein complexes. The methods consists of the combination of size exclusion chromatographic (SEC) fractionation of native complexes, the mass spectrometric analysis of sequential fractions by nanoLC- DIA/SWATH MS analysis and a computational framework for the analysis of the resulting data [1].

Here we present further improvements of this pipeline for exploring protein complexes in clinically relevant samples, at unprecedented speed, accuracy and proteome coverage. The Evosep One chromatography system [2] allows to characterize with high reproducibility and robustness ~3500 proteins for 21 min-long gradient run with DIA/SWATH acquisition and ~4000 proteins with diaPASEF mode [3]. This new workflow consents to quantify and determine the stoichiometry of hundreds of complexes in a sample within one day of data acquisition.

We applied the pipeline to different biological sample types such as mouse livers, human plasma, and monocityc cells where proteins are extracted in native conditions for preserving their tridimensional structure and interactions.

We are able to identify across the SEC elution profile ~ 5000 proteins for THP-1 cells, ~ 3400 proteins for mouse liver tissues, and ~700 proteins for human plasma across the SEC elution profile. From these analyses we isolated more than ~200 protein complex from CORUM database for liver tissues and ~300 for THP-1 cells, and we could predict 300 protein-protein interactions for human plasma.

1. Heusel M, et al.; Molecular Systems Biology (2019) 15, e8438.
2. Bache N., et al., Mol Cell Proteomics. (2018) 17: 2284-2296.
3. Meier F., et al., bioRxiv (2019) https://doi.org/10.1101/656207

207

A Multi-Omic Investigation into the Molecular Signatures of Preeclampsia and Gestational Diabetes Mellitus

Melanie T Odenkirk[1], Kristin E Burnum-Johnson[2], Brandie D Taylor[3], Kelly G Stratton[2], Marina A Gritsenko[2], Lisa M Bramer[2], Bobbie-Jo Webb-Robertson[2], Jennifer E Kyle[2], Karl Weitz[2], David Reif[1], Denis Fourches[1], Erin S Baker[1]

1. North Carolina State University, Raleigh, NORTH CAROLINA, United States

2. Pacific Northwest National Laboratory, Richland, WA

3. Temple University, Philadelphia, PA

Pregnancy disorders such as preeclampsia (PE) and gestational diabetes mellitus (GDM) are diagnosed in approximately 5-8% of pregnancies in the United States and substantially higher in under-developed countries. When left untreated PE and GDM greatly increase the risk of neonatal death and long-term disability, as well as maternal death. Post-pregnancy, these diseases have also been linked to an increased risk of cardiovascular and metabolic diseases for the mother. Unfortunately, both PE and GDM screenings are unavailable until 20-28 weeks of pregnancy and are not routinely tested unless the mother meets certain at-risk criteria or specific symptoms. To understand the molecular signatures of each disease and promote earlier diagnosis, plasma samples were acquired from 185 pregnant patients (92 control, 48 PE, 45 GDM) with various phenotypes on the day of delivery. Lipidomic and proteomic analyses were performed on the samples using LC-MS/MS and LC-IMS-MS/MS to elucidate specific molecular changes between mothers with and without these disorders. Proteomic data revealed statistically significant proteins with known functions that varied among diagnoses. Lipidomic data also revealed overall up- and down-regulation of lipid classes specific for GDM and PE patients, as well as variations based on specific fatty acid backbones. When assessed together, the omic analyses illustrated altered mechanisms for each disorder and targets for earlier diagnosis.

208

The role of lipids in plant stress biology

Ute Roessner[1]

1. School of BioSciences, The University of Melbourne, Melbourne, VIC, Australia

Lipids occur throughout the living world and are important molecules found in microbes, higher plants, animals and in all cell types. The main biological functions of lipids include storing energy, signalling, and acting as structural components of cell membranes. Lipids are a large and vastly varied group of organic compounds built from a limited number of building blocks

resulting in an enormous chemical structural diversity. This diversity is necessary to provide the many roles different lipids provide in cell development and metabolism. The function of many lipid species is yet to be established in mammalian systems, and in plants we have even less understanding of the roles and importance of the more diverse set of plant lipids in development, growth and stress response.

The field of lipidomics, which aims to analyse the entire lipidome of an organism, is now well established and has provided substantial breakthroughs in biomedical research. In our research programs, these methodologies are now being deployed to investigate lipidome changes of plants following exposure to environmental stress or genetic alterations. We have already shown that the plant lipidome is significantly more complex compared to mammals, is highly defined by tissue type and is highly responsive to abiotic stress, including salinity and temperature stress. I will present our efforts in the development of comprehensive plant lipidomics methods using orthogonal and imaging mass spectrometry approaches and their application to understand plant responses to abiotic stress.

209

A 'Systems-omics' Strategy to Uncover the Role of Brain Tissue Derived Exosomal Lipids in Alzheimer's Disease

Huaqi (Kate) Su[1, 2], Kevin J Barnham[1, 2], Laura J Vella[2], Gavin E Reid[1]

1. The University of Melbourne, Parkville, VIC, Australia

2. Florey Institute of Neuroscience and Mental Health, Melbourne, VIC, Australia

Lipid dysregulation is known to be associated with Alzheimer's disease (AD) pathogenesis. Exosomes (40-100 nm vesicular bodies) play key roles in effecting phenotypic changes upon fusion or uptake by a recipient cell. However, while exosomal membrane lipids and bioactive lipid cargo are expected to play pivotal roles in exosome secretion, fusion and uptake, and in target cell functional response, detailed characterization of the lipidomes of exosomes secreted from brain tissue of AD patients, and exploration of their functional roles, has not yet been explored. Here, using workflows recently developed for extracellular vesicle secretion, isolation and characterization (Vella, *et al. J. Extracell. Vesicles*. 2017, 6, 1348885) and quantitative mass spectrometry based lipidome analysis (Rustam and Reid. Anal. Chem., 2018, 90, 374–397.), the lipidome compositions of exosomes released by post-mortem frontal cortex brain tissue from a series of AD patients versus healthy controls, and their parental tissues, were acquired. Proteomic analysis of these samples was also performed to enable a 'systems-omics' evaluation of the key lipid metabolism pathways associated with AD pathogenesis. From this study, approx. 400 lipids were identified and quantified in each sample. Enrichment and remodelling of multiple exosome lipid classes and subclasses were observed, including glycerophospholipids and sphingolipids . Many of these changes have previously been reported to play key roles in the regulation of physiological processes relevant to AD pathogenesis. These results establish a foundation for future investigations of brain-derived exosome lipids as potential biomarkers for AD diagnosis and for the development of novel therapeutic agents acting through relevant lipid pathways for the treatment of this disorder.

213

Proximity dependent sensors define a role for HOPS in macropinocytosis-dependent control of cell growth

Anne-Claude Gingras[2, 1], Geoffrey Hesketh[2], Christopher D Go[2, 1], James DR Knight[2]

1. Molecular Genetics, University of Toronto, Toronto, ON, Canada

2. Mount Sinai Hospital, Toronto, ON, Canada

The mechanistic Target of Rapamycin Complex 1 (mTORC1) couples nutrient sufficiency to cell growth, and is an important therapeutic target, notably in cancers. Amino acids can be up-taken in free form from transporters, and detected by different protein complexes that control the nucleotide-binding status of small GTPases of the Rag families. Nucleotide-bound Rag proteins recruit mTORC1 to the surface of the lysosome to mediate its full activation. A less well-characterized pathway of amino acid uptake is through macropinocytosis, in which extracellular proteins are endocytosed and delivered to the lysosome where they are degraded to amino acids, which also results in mTORC1 activation. Macropinocytosis is importantly enhanced by oncogenic KRAS mutations, and is thought to contribute to cancer survival in conditions of low free nutrient supply. Leveraging a proximity-dependent map of a human cell (Go, Knight et al., in prep), we generated proximity-dependent biotinylation "sensors" to study the recruitment of mTORC1 components to the surface of the endolysosomes. This revealed an unsuspected role for the HOPS trafficking complex in the activation of mTORC1 specifically in macropinocytic contexts. We also defined different roles for HOPS and the previously characterized GATOR2 complex in the mechanisms of mTORC1 activation downstream of specific nutrient acquisition pathways. Together, our results help understanding the mechanisms of growth regulation by macropinocytosis, and may offer new therapeutic avenues for cancers in which this pathway is upregulated.

214

Venoms to Drugs

Paul Alewood[1]

1. Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, Australia

Many organisms including snakes, spiders, scorpions, cone snails, anemones and some mammalian species have evolved venom as either a defence mechanism or a weapon for prey capture1. These venoms typically contain a complex cocktail of bioactive disulfide-bond rich polypeptides called toxins that target a wide range of receptors including enzymes, ion channels, GPCRs and transporters. Of interest to drug designers is their high potency and selectivity for ion channels and receptors combined with their resistance to many proteases.

Of particular interest are the venom peptides (conotoxins) from the Conidae2,3, most with small polypeptide chains containing 10-40 amino acids that are highly constrained by one to five disulfide bridges and are structurally well defined. Their high potency and exquisite selectivity has led to two drug candidates4,5 from our laboratories.

In this presentation I will outline our program of discovery using a venomics strategy3, describe the amazing diversity of molecular structures being discovered as well as regioselective chemistry6,7 that has facilitated the determination of structure-function relationships. This has led to the availability of potent molecules with novel function8 and exceptional stability when exposed to reducing environments and in plasma. Together, these results underpin the development of more stable and potent peptide mimetics suitable for new drug therapies, and highlight the application of this technology more broadly to disulfide-bonded peptides and proteins.

1. 1. Sébastien Dutertre, Ai-Hua Jin, Irina Vetter, Brett Hamilton, Kartik Sunagar, Vincent Lavergne, Valentin Dutertre, Bryan Fry, Agostinho Antunes, Paul F. Alewood and Richard J. Lewis. Nature Communications 5:3521, 2014.
2. 2. Akondi KB, Muttenthaler M, Dutertre S, Kaas Q, Craik DJ, Lewis RJ, Alewood PF (2014). Chemical Reviews 114 (11) 5815.
3. 3. Vincent Lavergne, Ivon Harliwong, Alun Jones, David Miller, Ryan J Taft, Paul F Alewood. Proceedings of the National Academy of Sciences (USA). 112 (29) E3782-E3791, 2015.
4. 4. RL Lewis, D Adams, I.Sharpe, M Loughnan, T Bond, L Thomas, A. Jones, J Matheson, R Drinkwater, K Nielsen, DJ Craik and PF Alewood (2000). J Biol Chem, 275(45) 35335.
5. 5. I Sharpe, J Gehrmann, M Loughnan, L Thomas, D Adams A Atkins, DJ Craik, D Adams PF Alewood and RJ Lewis (2001). Nature Neuroscience, 4(9) 902.
6. 6. M. Muttenthaler, S. T. Nevin, A. A. Grishin, S. T. Ngo, P. T. Choy, N. L. Daly, S-H. Hu, C. J. Armishaw, C. I. A. Wang, R. J. Lewis, J. L. Martin, P. G. Noakes, D. J. Craik, D. J. Adams, P. F. Alewood. Journal of the American Chemistry Society, 132 (10) 3514-3522, 2010.
7. 7. Aline Dantas de Araujo, Mehdi Mobli, Stuart M. Brierley, Joel Castro, Andrea M. Harrington, Irina Vetter, Zoltan Dekan, Markus Muttenthaler, Jingjing Wan, Richard J. Lewis, Glenn F. King and Paul F. Alewood. Nature Communications 5, 3165, 2014.
8. 8. Jin AH, Dekan Z, Smout MJ, Wilson D, Dutertre S, Vetter I, Lewis RJ, Loukas A, Daly NL, Alewood PF. Angew Chem Int Ed. 56(47):14973-14976, 2017.

219

Standardizing the performance of SWATH/DIA-MS software tools for label-free quantification using public repository spectral ion libraries

Robert Moritz[1]

1. Institute For Systems Biology, Seattle, WA, United States

Spectral ion libraries are routinely used to analyze data derived from SWATH/DIA type mass spectrometry acquisition methods. For confidence in protein identification and quantitation using SWATH/DIA data, the characteristics of reference ion library is important when processing SWATH/DIA data for quantitative results. Different library construction methods can affect the performance of SWATH/DIA analysis that incorporate the expected multitude of non-standardized SWATH/DIA data acquisition method components such as chromatography flow rates, different mass spectrometry instruments, spectral quality and transferability, and SWATH/DIA analysis software tools. For a consistent assessment of reference ion libraries for SWATH/DIA analysis, we developed DIA-Lib QC, a library assessment tool built into the SWATHAtlas website (www.swathatlas.org) to calculate the metrics of library correctness and its completeness, which provides some level of interrogation for confidence in the database used for library-based SWATH/DIA analysis. With our benchmarking studies using hybrid ground-truth datasets, we evaluated the characteristics and performance of different SWATH/DIA-MS tools based on accuracy of identification, the precision of quantitation, quantitative sensitivity, and relative performance of these applications. Our analysis highlights many inconsistencies in available libraries and their performance for reliable quantification performance, and robustness in label-free quantitative proteomics.

220

Development of clinically applicable lung cancer proteome biomarkers for in vitro diagnostics-multivariate index assay

Jung-Mo Ahn[1], Hye-Jin Sung[1, 2], Yong-In Kim[1], Je-Yoel Cho[1]

1. Department of Biochemistry, BK21 and Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, South Korea

2. Division of Research & Development, ProtanBio, Seoul, South Korea

Without any effective tool for screening and early diagnosis, lung cancer shows the highest mortality in cancer related death. Here, we show the development of lung cancer proteome biomarkers and in vitro diagnostics based on multivariate index assay

(IVD-MIA). By LC-ms/ms–based proteomics approaches combined with glycoprotein enrichment, low molecular weight protein enrichment technologies in the sera of the cancer patients, and secretome analysis from primary cultured lung cancer and normal tissues, we have discovered various potential lung cancer protein biomarkers. TMT or iTRAQ-based quantitative and label-free proteomics combined with fucosylated glycoprotein enrichment approaches also revealed that not only the amount of the glycoprotein biomarkers but also their fucosylation levels and patterns can serve as diagnostic and prognostic serological markers for lung cancers. Fucosylated protein biomarkers were validated by lectin-hybrid ELISA and immunoassays. We also validated the biomarkers by multiple reaction monitoring (MRM) in the sera of the patients. Functional analysis also revealed that biomarker PON1 promotes ROS deregulation protecting the mitochondria from dysregulation. SAA and QSOX1 promote lung cancer metastasis by immunomodulating macrophages. A logistic regression model based on SID-MRM assay of SERPINA4 and PON1 serum protein markers with age factor revealed AUC 0.92 for differential diagnosis between lung cancer and other lung diseases. We have developed various pairs of biomarker-specific monoclonal antibodies and used these antibodies which can be used for lateral flow assay and microfluidics assay. Using three selected potential protein biomarkers and age, a deep learning algorithm showed AUC 0.90 in the sera test of 1,700 lung cancer patients and normal control. Our future development of IVD-MIA will further improve the detection and diagnosis of lung cancers

.

225

Unraveling protein targets of bioactive small molecules using label-free chemical proteomics

HoJeong Kwon[1]

1. Yonsei University, Seoul, South Korea

Identifying protein targets of bioactive small molecules and deciphering the specific mechanisms-of-action at the molecular level are crucial steps in the development of drugs to treat human diseases. We have developed target protein identification methods including conventional affinity chromatography using labeled small molecules as well as recent target identification method with label-free small molecules such as Drug Affinity Responsive Target Stability (DARTS) in combination with LC-MS/MS analysis to identify the direct binding proteins of small molecules. The direct interaction between small molecule and the target protein is validated via bio-physical, and bio-informatics methods. Moreover, biological relevancy of this “small molecule-target” interaction is validated through genetic modulation study and facilitates structure based better drug development. In this presentation, our recent studies on target identification of bioactive small molecules for exploring new mechanism studies and translational applications will be presented by introducing our case studies of protein target identification and validation of bioactive small molecules from natural products and clinical drugs.

231

Big Data And Health And World Wide Omics Profiling

Michael Snyder[1]

1. Stanford University, Stanford, CA, United States

Recent technological advances as well as longitudinal monitoring not only have the potential to improve the treatment of disease (Precision Medicine) but also empower people to stay healthy (Precision Health). We have performed to studies to examine human health and diversity. First, we have profile 109 participants using multiomics technologies (genomics, immunomics, transcriptomics, proteomics, metabolomics, microbiomics and wearables etc) for up to eight years and made 49 major health discoveries. Second, to understand human diversity from around the world, we have begun to profile more than 300 people who attended the HUPO meetings over the past three years using multiomics technologies. The results of these studies will be presented.

210

MS-based tissue profiling for assistance on neurosurgery operations of brain cancer

Igor Popov[1], Anatoly Sorokin[1], Daniil Ivanov[1], Vsevolod Shurkhay[2], Vasily Eliferov[1], Evgeny Zhvansky[1], Stanislav Pekov[1], Alexander Potapov[2], Evgeny Nikolaev[3]

1. Mowcow Institute of Physics and Technology, Dolgoprudny, MOSCOW REGION, Russia

2. N.N. Burdenko National Medical Research Neurosurgery Center, Moscow, Russia

3. Skolkovo Institute of Science and Technology, Moscow, Russia

The implementation of ambient mass spectrometry profiling techniques for assistance on surgery in precise identification of pathological modifications of tissues is recognized as an essential step towards the arise of prospective medicine. However, transfer the mass spectrometry techniques to a clinic, and their integration with clinical protocols and regulations require particular investigations. In this work, the author's experience of the development of the approaches to applying mass spectrometry profiling in neurosurgery was summarized.

The research was conducted on the basis of N.N. Burdenko National Medical Research Neurosurgery Center in accordance with local and international ethics guidelines. Samples of brain tumor tissues resected during elective surgery were analyzed using custom designed Inline Cartridge Extraction system followed by ESI to register lipid profiles. More than 300 samples from more than 150 patients were analyzed in 2018-2019.

Each resected tissue sample was divided into three parts - one of them analyzed immediately on the low-resolution instrument in the clinic, second was examined by the pathologist while the third one was frozen and transferred to the remote laboratory for

an investigation with the high-resolution mass spectrometer. All data were accumulated in a database that allows to compare and analyze results of low-resolution lipid profiling of recently resected sample, high-resolution lipid profiles of frozen samples investigation with histological annotations and the results of ms/ms analysis of conventionally extracted lipids and metabolites. Despite the variations in spectra which originates in biological variability and artificial alterations (e.g., time passing between the moment of the resection and tissue analysis) sustainable combinations of peaks are still present in spectra. Combinations of these stable characteristic peaks is shown to be suitable for reliable automated tissue classification. Performed investigations demonstrate the possibility to apply ambient mass spectrometry to neurosurgical operations using high-resolution data as well as low-resolution ones.

211

Comparative targeted and high-throughput metabolomics workflow of small-volume plasma samples

Stephan Klatt[1], Brunda Nijagal[2], Berin Boughton[3], Christopher Fowler[1], AIBL research group[4], Blaine Roberts[1, 5]

1. The Florey Institute of Neuroscience and Mental Health, Melbourne/Parkville, VIC, Australia

2. Metabolomics Australia, Bio21 Institute, University of Melbourne, Parkville, Melbourne, Australia

3. Metabolomics Australia, School of Biosciences, The University of Melbourne, Parkville, Victoria 3052, Australia

4. https://aibl.csiro.au/about/aibl-research-team, Melbourne

5. Cooperative Research Centre for Mental Health, Parkville, Victoria 3052, Australia

Here, we present a comparative targeted and high-throughput metabolomics workflow of small-volume plasma samples. In detail, we have extracted and analysed 3.5 ul of human plasma from 20 individuals, with the plasma derived from venipuncture and finger-prick blood. Finger-prick derived plasma was collected on specific blood cards (Noviplex™ Plasma Prep Cards). When air-dried, card-applied plasma is stable and cards can be mailed to the next analytical laboratory, making them highly advantageous in rural areas compared to the standard venipuncture blood draw. Plasma samples were extracted with methanol, AQC-derivatised and analysed on a 6495 QQQ LC/MS instrument (Agilent Technologies) in dMRM mode. The dMRM method contains 80 metabolite targets, whereof most play important roles in neurodegenerative diseases like Alzheimer's disease and Parkinson's disease. Included metabolites are 20 amino acids, six hormones, eight metabolites of the kynurenine pathway, three polyamines and many more. In total, 75 metabolites were successfully detected in the 3.5 ul of starting material, with concentration differences between venipuncture and blood card-derived plasma. Method development was further adapted to a 96-well plate format, enabling sample processing in high-throughput and biomarker validation for the study of neurodegenerative diseases.

212

Identifying predictive network of lipids and proteins among individuals at high and low risk of coronary artery disease using iOmicsPASS

Hiromi W.L. Koh[1], Sock Hwee Tan[1, 2], Mark Y. Chan[1, 2], Arthur Mark Richards[1, 2, 3], Hyung Won Choi[1]

1. Cardiovascular Research Institute, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

2. National University Heart Center, Singapore

3. Christchurch Heart Institute, University of Otago Christchurch, Christchurch Hospital, Christchurch, New Zealand

Background

In our previous work, we developed a tool called iOmicsPASS which searches for subnetworks of molecular interactions within and between -omics data types that can predict different phenotypic groups. The tool incorporates biological networks for the measured molecules in pursuit of the predictive interactions of each phenotypic group, and we showed that this approach is highly efficient in the integration of proteome and transcriptome data for cancer subtyping. Nonetheless, the prior information of biological network may not always be available between certain types of molecules, for instance between lipids and proteins. Hence, a more data-driven approach to predict unknown interactions would be desirable.

Methods

We predict the molecular interactions between lipids and proteins by considering the cross-covariance matrix between the two types of –omics data. By applying a shrinkage method, we de-noise the covariance matrix to obtain a sparse matrix which highlights the most strongly correlated molecule pairs, rendering a set of predicted interactions between and within the lipidome and proteome. Then, we applied iOmicsPASS, using the derived lipid-protein interactions in the blood plasma, to identify a set of predictive molecular interactions which can best differentiate individuals of high and low risk of coronary artery disease (CAD). We also explored the biological pathways which were up- and down-regulated in the individuals at a higher risk of CAD.

Results

We illustrate iOmicsPASS using proteomics and lipidomics data from a Singapore Chinese cohort. Individuals were classified into high and low-risk groups based on Framingham risk scores and other cardiovascular risk factors, and lipid-protein and protein-protein interactions predictive of CAD risk groups were identified by iOmicsPASS analysis. We also re-categorize subjects based on their measured plaque volume index in their coronary arteries to find lipid and protein signatures which could potentially highlight early coronary plaque burden in asymptomatic individuals.

215

Unveiling protein targets involved in haptenation during allergic contact dermatitis via high resolution mass spectrometry

Wint Wint Phoo[1, 3, 2], Puck Wee Chan[4, 2], James Chun Yip Chan[4, 2], Eric Chun Yuon Chan[5, 2], Radoslaw Sobota[1, 3, 2]

1. Singapore National Laboratory for Mass Spectrometry, Singapore

*2. Innovations in Food and Chemical Safety Program (IFCS), Biomedical Research Council (BMRC), Agency for Science, Technology and Research (A*STAR), Singapore*

*3. Functional Proteomics Laboratory, Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), Singapore*

*4. Skin Research Institute of Singapore (SRIS), Agency for Science, Technology and Research (A*STAR), Singapore*

5. Department of Pharmacy , National University of Singapore, Singapore

Allergic contact dermatitis commonly known as skin sensitization affects 15-20% of the general population. According to the adverse outcome pathway, following the exposure of a sensitizing compound to the skin, proteins are covalently modified by the sensitizer. The haptenated proteins are degraded and the peptide-hapten adducts are presented as antigens. Hence, the subsequent exposure with the same sensitizer results in epidermal inflammation. There is an urgent industry need for *in vitro* methods to access the sensitizing potential of chemical compounds. In the study, we employ high resolution mass spectrometry combined with cellular thermal shift assay (CETSA-MS) as well as product ion monitoring method to discover proteins targets involved in haptenation using a set of chemicals with known sensitizing potentials. Using CETSA-MS, we have identified a list of proteins that are commonly targeted by extreme sensitizers. In complement to our CETSA-MS approach, we develop the product ion monitoring method for identifying the peptide-adducts in peptide mixtures. These MS based methods provide novel *in vitro* approaches to determine the sensitizing potential of unknown chemical compounds as well as to gain insights into pathways involved in allergic contact dermatitis.

216

Reverse chemical proteomics identifies unanticipated human target of anti-malarial drug, Artesunate, for drug repositioning

Sung Min Cho[1], Michael P. Gotsbacher[2], Peter Karuso[2], Ho Jeong Kwon[1]

1. Chemical Genomics Global Research Laboratory, Department of Biotechnology, College of Life Science & Biotechnology, Yonsei University, Seoul, Korea

2. Department of Molecular Sciences, Macquarie University, Sydney, Australia

Drug repositioning saves the time and cost of drug development. It is a promising approach reducing risk of side effects because it starts from already approved drugs. In this study, reverse chemical proteomics enables access to new targets and pathways of existing drugs targeting human diseases. Artesunate (ART) is used as a case compound in this study. ART is the most potent and safe antimalarial drug which is one of the derivatives of the Artemisinin, the Nobel Prize winning compound in 2015. The unbiased interrogation of several human cDNA libraries, displayed on bacteriophage T7, identified a high affinity human target of ART; the Bcl-2 antagonist of cell death promoter (BAD). ART inhibits the phosphorylation of BAD, promoting the formation of the BAD/Bcl-xL complex and the subsequent intrinsic apoptosis. Direct interaction of ART with BAD was validated by DARTS and ICC experiments using label-free ART and fluorescence-ART, respectively. In addition, ART induces caspase-dependent apoptosis by triggering release of cytochrome c from mitochondria to cytosol in cancer cells. Collectively, this case study of reverse chemical proteomics demonstrates the potential of this method for new target identification of known clinical drug in drug repositioning.

217

Proteome profiling of multiple sclerosis cerebrospinal fluid by data independent acquisition reveals changes in proteins involved in adaptive immunity

David R. Spiciarich[1], Christopher T. Harp[1], Ann E. Herman[1], W. Rodney Mathews[1], Veronica G. Anania[1]

1. OMNI Biomarker Department, Genentech, Inc., South San Francisco, CA, United States

Multiple sclerosis (MS) is a chronic autoimmune, inflammatory neurological disease of the central nervous system (CNS) resulting in damage to myelin and myelinated axons in the CNS and can lead to severe disability. The most common form, relapsing-remitting MS (RRMS), is marked by acute onset of neurologic symptoms (relapses), followed by periods of remission. Another form, primary progressive MS (PPMS) is a severe manifestation of the disease characterized by unabated worsening symptoms without remission. In order to identify candidate biomarkers of disease activity, or for use in evaluation of therapeutic response across different manifestations of the disease, an in-depth proteome analysis of CSF from two independent cohorts of RRMS, PPMS, and control patients with both healthy and other neurologic diseases (OND) was conducted. Mass spectrometry-based proteomics was employed using a data independent acquisition (DIA) workflow with a CSF-specific spectral library. Results from two independent cohorts were compared and proteins with differential CSF expression patterns across disease subtypes were identified for further characterization. This DIA approach allowed for the identification of greater than 1,500 proteins quantified across the samples. Statistical testing allowed for the identification of protein groups significantly changed in PPMS, RRMS or PPMS samples compared to RRMS samples. Pathway analysis revealed changes between RRMS and PPMS in pathways that included B cell and macrophage biology. This extensive study gives a comprehensive analysis of the RRMS and PPMS proteome observed from CSF that will be used to inform clinical biomarker strategies focused on evaluating therapeutic mechanism of action. Herein we provide a comprehensive profile of the MS CSF proteome and identify biomarker candidates for further investigation.

218

Thermostable potassium channel-inhibiting neurotoxins in processed scorpion medicinal material revealed by proteomic analysis: implications of its pharmaceutical basis in traditional Chinese medicine

Fan Yang[1], Danya Wang[2, 1], Yongjia Tong[2, 1], Chenghu Qin[1], Liuting Yang[2, 1], Fei Yu[2, 1], Xin Huang[1], Shuang Liu[1], Zhijian Cao[1, 3], Lin Guo[2, 1], Wenxin Li[1, 3], Yingliang Wu[1, 3], Xiaolu Zhao[2, 1]

1. State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei, China

2. Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan, Hubei, China

3. Center for BioDrug Research, Wuhan University, Wuhan, Hubei, China

The neurotoxins of venomous scorpion act on ion channels. Whether these neurotoxins are retained in processed *Buthus martensii* Karsch scorpions used in traditional Chinese medicine materials is unknown. Here, we performed a comprehensive mass spectrometry-based proteomic characterization of functionally active toxins in the processed medicinal scorpion material. Our proteomic analysis combining HCD and ETD techniques revealed 22 full-length and 44 truncated thermostable potassium channel-modulatory toxins that preserved six conserved cysteine residues capable of forming the three disulfide bonds necessary for toxicity. Additionally, a broad spectrum of degraded toxin fragments was found, indicating their relative thermal instability which enabled toxicity reduction. Furthermore, the suppression of IL-2 production in Jurkat cells and the reduced delayed-type hypersensitivity (DTH) response demonstrated that the extracts have immunoregulatory activity both *in vitro* and *in vivo*. Our work describes the first "map" of functionally active scorpion toxins in processed scorpion medicinal material, which is helpful to unveil the pharmaceutical basis of the processed scorpion medicinal material in traditional Chinese medicine.

22

Protein paucimannosylation is an enriched *N*-glycosylation signature of human cancers

Sayantani Chatterjee[1], Ling Y. Lee[2], Rebeca Kawahara[3], Jodie L. Abrahams[4], Barbara Adamczyk[5], Merrina Anugraham[6], Christopher Ashwood[7], Zeynep Sumer-Bayraktar[8], Matthew T. Briggs[9], Jenny H.L. Chik[10], Arun Everest-Dass[4], Sarah Förster[11], Hannes Hinneburg[1], Katia R.M. Leite[12], Ian Loke[13], Uwe Möginger[14], Edward S.X. Moh[1], Miyako Nakano[15], Saulo Recuero[12], Manveen K. Sethi[16], Miguel Srougi[12], Kathrin Stavenhagen[17], Vignesh Venkatakrishnan[18], Katherine Wongtrakul-Kish[19], Simone Diestel[20], Peter Hoffmann[9], Niclas G. Karlsson[5], Daniel Kolarich[4], Mark P. Molloy[21], Michael H. Muders[11], Martin K. Oehler[22], Nicolle H. Packer[1], Giuseppe Palmisano[3], Morten Thaysen-Andersen[1]

1. Department of Molecular Sciences and Biomolecular Discovery and Design Research Centre (BDDRC), , Macquarie University, Sydney, New South Wales, Australia

2. ISGlobal, Barcelona Centre for International Health Research (CRESIB), Hospital Clínic -Universitat de Barcelona, Barcelona, Spain

3. Department of Parasitology, Institute of Biomedical Sciences, University of Sau Paulo, Sau Paulo, Brazil

4. Institute for Glycomics, Griffith University, Gold Coast, Queensland, Australia

5. Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy,, University of Gothenburg,, Gothenburg, Sweden

6. Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

7. Department of Biochemistry, Medical College of Wisconsin, Milwaukee, USA

8. School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

9. Future Industries Institute, Mawson Lakes Campus, University of South Australia, Adelaide, South Australia, Australia

10. International Collaboration On Repair Discoveries (ICORD), Vancouver Coastal Health Research Institute and Department of Pathology and Laboratory Medicine, The University of British Columbia, Vancouver, Canada

11. Rudolf-Becker-Laboratory, Institute of Pathology, University Hospital Bonn, Bonn, Germany

12. Laboratório de Investigação Médica da Disciplina de Urologia (LIM55), Faculdade de Medicina da FMUSP, Universidade de Sao Paulo, Sao Paulo, Brazil

13. Department of Biological Sciences, National University of Singapore, Singapore

14. Institute for Biochemistry and Molecular Biology, University of Southern Denmark, Denmark

15. Graduate School of Advanced Sciences of Matter, Hiroshima University, Hiroshima, Japan

16. Center for Biomedical Mass Spectrometry, Department of Biochemistry, Boston University School of Medicine , Boston University, Boston, MA, USA

17. Beth Israel Deaconess Medical Center, Department of Surgery and Harvard Medical School Center for Glycoscience, Harvard Medical School, Boston, MA, USA

18. Department of Rheumatology & Inflammation Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

*19. Bioprocessing Technology Institute (BTI), A*STAR, Singapore*

20. Institute of Nutrition and Food Sciences, University of Bonn, Bonn, Germany

21. Sydney Medical School, Bowel Cancer and Biomarker Laboratory, Kolling Institute, The University of Sydney, Sydney, New South Wales, Australia

22. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

We recently discovered an immune-related class of glycoproteins decorated with paucimannosidic *N*-glycans (PMGs) [$Man(M)_{1-3}GlcNAc_2Fuc(F)_{0-1}$] expressed by innate immune cells [Thaysen-Andersen *et al*., J Biol Chem, 290(14):8789, 2015]. Herein, we extend these findings by investigating a possible association between protein paucimannosylation and cancer. To this end, the distribution of PMGs released from proteins expressed by a wide range of human cancer cells and tissues and matching non-cancerous specimens were determined by re-interrogating a large collection of published and unpublished *N*-glycomics datasets acquired over a decade within our laboratories using a uniform porous graphitised carbon liquid chromatography tandem mass spectrometry method. The total level of PMGs within the entire *N*-glycome and the relative distribution of the individual PMG

species (M1, M1F, M2, M2F, M3, and M3F) were accurately determined using this well-established quantitative glycomics platform. In total, *N*-glycome profiling data from 34 different cancer cell lines, and 133 tissue samples spanning 11 cancer types with matching non-cancerous specimens, were (re-) curated from 467 datasets. Although the total PMG levels varied dramatically across and within the investigated cancer types (1.0%-50.2%), particularly the α1,6-fucosylated bi- and tri-mannosylated *N*-glycans (M2F and M3F) were consistently prominent features of most of the cancer cell lines. Analyses of paired (tumour/non-tumour) and stage-stratified tissues demonstrated that PMGs are significantly enriched in tumours from several cancer types including non-melanoma skin cancer ($p = 0.0145$), liver ($p = 0.0033$) and colorectal ($p = 0.0017$) cancers and are increased with prostate cancer and chronic lymphocytic leukaemia progression (both $p < 0.02$). Based on these correlation-type observations we conclude that protein paucimannosylation represents a significant, but non-uniform, glycoepitope of human cancers. These findings advance our understanding of the glyco-features expressed by human cancers and encourage further cause-effect type studies of paucimannosylation in cancer tumorigenesis and metastasis.

221

Multidimension lc-ms/ms analysis of csf samples in the biofind cohort for biomarker discovery in Parkinson's disease

Jarrod Marto[1]

1. Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, United States

BACKGROUND:

Parkinson's disease (PD) is the second most common neurodegenerative disorder, afflicting more than 4 million people worldwide. Diagnosis is typically based on clinical symptoms related to impaired motor function, at which point most dopaminergic neurons have been lost. PD symptoms vary across individuals, and a prodromal phase may precede clinical symptoms by several years. Hence there is an urgent need for robust PD biomarkers to facilitate early detection and all aspects of disease management.

METHODS:

We utilized fully automated, online 3D peptide fractionation with TMT isotope labels for data-dependent quantification of protein, protein phosphorylation, and protein glycosylation across BioFIND CSF samples (BioFIND cohort), comprising 118 PD patients and 88 healthy control (HC) subjects. We used targeted mass spectrometry and biochemical assays to validate candidates in an independent cohort of PD/HC CSF samples from the Harvard Neurodiscovery Center (HNDC cohort).

RESULTS:

We identified over 6,000 proteins across CSF samples in the BioFIND cohort. Requiring zero missing values, we quantified over 1,200 proteins in all BioFIND CSF samples. Regression analysis revealed more than 60 proteins whose abundance in the CSF was associated with PD (*Padj* < 0.05). We used biochemical and targeted mass spectrometry assays to validate more than 15 candidates, including several not previously associated with PD, in the HNDC cohort comprising 40 PD and 40 HC CSF samples.

CONCLUSION:

We generated the most detailed view of the CSF proteome to date. The robustness of our approach provided quantification of more than 1,200 proteins (with no missing values) across all samples in our discovery cohort, including glycosylated and phosphorylated proteins. Our validation assays nominate several proteins as candidates for further interrogation across longitudinal and other PD patient cohorts. Our study further provides a roadmap for generating comprehensive CSF spectral libraries for use in new, data independent quantitative proteomic assays.

222

Human complement factor B: a new pancreatic cancer biomarker with multiple functions

Min Jung Lee[1], Keun Na[2], Heon Shin[3], Chang Moo Kang[4], Soo-Youn Lee[5], Hoguen Kim[6], Hye Jin Choi [7], Sumi Bae[8], Sunghwa Son[9], YOUNG-KI PAIK[1]

1. Yonsei Proteome Research Center, Yonsei University, SEOUL, --, Korea

2. Yonsei Proteome Research Center, Yonsei University, SEOUL, --, Korea

3. Yonsei Proteome Research Center, Yonsei University, SEOUL, --, Korea

4. Hepatobiliary and Pancratic Surgery, Yonsei University College of Medicine, SEOUL, --, Korea

5. Laboratory Medicine and Genetics, Sungkyunkwan University, SEOUL, --, Korea

6. Pathology, Yonsei University College of Medicine, SEOUL, --, Korea

7. Internal Medicine, Yonsei University College of Medicine, SEOUL, --, Korea

8. JW Bioscience Corp, SEOUL, --, Korea

9. JW Holdings Corp, SEOUL, --, Korea

Pancreatic cancer (PC) is the fourth leading cause of death worldwide. Although the serum carbohydrate antigen (CA)19-9 level is a widely used diagnostic marker of pancreatic cancer, this parameter is not sufficiently sensitive or accurate and cannot be used to screen small, resectable early-stage cancers. Accordingly, screening strategies based on novel serum markers that are more specific and sensitive for pancreatic cancer are urgently needed. Our data showed that CFB, either alone or in combination with CA19-9 (termed ComB-CAN) exhibited good diagnostic ability for early-stage (I-II) pancreatic cancer, with ≥90% sensitivity and 98% specificity. ComB-CAN was much more sensitive than CA 19-9 for discriminating early-stage PC from pancreatitis and pancreatic tumors. To study the potential role of CFB in pancreatic cancer cell (PANC1) carcinogenesis, we used short hairpin (sh) RNAs to generate stable MOCK (shControl) and CFB-knockdown (shCFB) PANC1 cell lines. We also analyzed the transcriptome to identify CFB-related genes. We identified 59 and 86 genes exhibiting a >4-fold increase or decrease, respectively, in CFB-knockdown cells relative to MOCK cells. These genes were subjected to gene ontogeny classification based on an ExDEGA search analysis. RNA-seq data analysis demonstrated the downregulation of proteins involved in cell migration, angiogenesis and ECM in CFB-knockdown cells. These data indicate that an increased CFB level is closely related to pancreatic

cancer cell carcinogenesis and supports multiple roles for CFB: oncogenic factor and an early serologic PC diagnostic marker. (This study was supported by grants from the Korean Ministry of Health and Welfare: HI13C2098, and HI16C0257 to Y.-K. Paik)

223

The eutopic endometrium proteome in endometriosis reveals candidate markers and molecular mechanisms of physiopathology

Loren Mear[1], Emmanuelle Com[1], Laetitia Guillot[1], Blandine Guevel[1], Regis Lavigne[1], Khadija Fathallah[2], Arnaud Fauconnier[2], Francois Vialard[3], Charles Pineau[1]

1. PROTIM, Inserm U1085 - Irset, Rennes, France

2. Department of Obstetrics and Gynecology , CHI de Poissy, Saint Germain en Laye, France

3. EA7404-GIG, UFR des Sciences de la Santé Simone Veil, EA7404-GIG, Montigny le Bretonneux, France

Endometriosis is a common chronic gynecological disease that affects up to 10% of women of reproductive age. It is characterized by the presence of endometrium outside the uterine cavity, causing various symptoms, such as infertility and chronic pain. The gold standard for its diagnosis is still laparoscopy and the biopsy of endometriotic lesions due to the absence of non-invasive diagnostic tools. Here, we aimed to compare the eutopic endometrium from women with or without endometriosis to identify proteins that may play a role in the etiology of endometriosis and thus be considered as potential biomarker candidates, calling for further targeted studies.

Eutopic endometrium was collected from patients with endometriosis and women without endometriosis during a laparoscopy surgery between days 19 and 24 of their menstrual cycle (mid-secretory phase). Total proteins from tissues were extracted and digested with trypsin before LC-MS-MS analysis in a data-dependent acquisition mode on a TimsTOF Pro instrument. Data were evaluated using bioinformatics tools and integrated to previously published data.

Among the 5,301 proteins identified, 543 were differentially expressed. This global differential signature allowed us to separate directly the controls from the endometriosis samples. Interestingly differential proteins were enriched in two specific KEGG pathways: focal adhesion and PI3K/AKT signaling. Integration of our data with a yet unpublished large-scale proteomics dataset allowed us to highlight 11 proteins that share the same trend of dysregulation in eutopic endometrium, regardless of the phase of the menstrual cycle, whereas 16 proteins were identified as being overexpressed in the serum of women with endometriosis.

Our results constitute a first milestone towards the identification of potential promising endometrial and serological diagnostic biomarkers. They provide new insights into the mechanisms underlying endometriosis and its etiology. Our results need further confirmation on a larger sample cohort. Data are available via ProteomeXchange with the identifier PXD012981.

224

Phenotyping of multiple biofluids for liquid biomarkers for diagnostics and personalized medicine

Thomas BG Poulsen[1], Tue B Bennike[1], Svend Birkelund[1], Claus H Nielsen[2], Per T Sangild[3], Sten Rasmussen[4], Lars Arendt-Nielsen[1], Allan Stensballe[5]

1. Aalborg University, Aalborg, JYLLAND, Denmark

2. Danish National Hospital, Department of Inflammation, Copenhagen, Denmark

3. Copenhagen University, Copenhagen, Denmark

4. Aalborg university hospital, aalborg, denmark

5. AALBORG UNIVERSITET, Aalborg, DENMA, Denmark

Background & Significance

Inflammatory and autoimmune diseases include multifactual pathomechanisms and systemic responses. Advances in high-throughput molecular technologies have increased investigations into the utility of transcriptomic, proteomic and metabolomic approaches as diagnostic tools for precision medicine (1). Although more commonly diagnosed in adults, all of above diseases can manifest in childhood as early as infancy (2). Thus, Intrauterine bacterial infection predisposes to preterm birth and is associated with dysregulated development of several organs, including gut, lungs and brain. Basic blood tests include inflammatory markers and autoantibodies, however, now analysis speed and robustness allow more readily clinical insight biofluids. We present recent concepts and studies investigating early life inflammatory diseases as well as low grade inflammatory diseases in different biofluids from plasma to CSF accessing causalities leading to inflammation and pain.

Methodologies

Plasma, CSF, synovial fluid and urine samples were investigated in multiple pathologies (Arthritis, sepsis etc.), before and after treatment in patients (biologics; intraarticular gold) or in neonates as well as pig as model animal for neonate conditions. Plasma and urine may be available from the HUPO hPOP project. Deep proteome, PTM and EV profiling were accomplished using quantitative proteomics approaches using quantitative mass spectrometry-based analysis by DIA/PASEF followed by deep datamining (3,4). PTM profiling were evaluated by 4D CCS based feature finding.

Results and Conclusion

Discovery of biomarkers and/or inflammatory signatures through integration of multi-omic data has potential to stratify patients or neonates with sepsis for improved treatment and prognosis. Firstly, our data using next generation proteomics approaches alleviates many pitfalls of missing values and poor proteome coverage including unbiased PTM profiling without enrichment strategies. Next, investigating neonates and molecular ontogeny of newborns revealed distinct immune immunomodulatory and inflammatory profiles readily correlated by cross-species pig studies. Low grade inflammation and treatment regimens are more readily addressed comparing multiple biofluids.

1. 1. Ng S, Strunk T, Jiang P, Muk T, Sangild PT, Currie A. Precision Medicine for Neonatal Sepsis. Front Mol Biosci 2018;5:70.
2. 2. Lee AH, Shannon CP, Amenyogbe N, Bennike TB, Diray-Arce J, Idoko OT, et al. Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. Nat Commun 2019 10(1):1092.
3. Meier F, Beck S, Grassl N, Lubeck M, Park MA, Raether O, et al. Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. J Proteome Res. 2015 12):5378–87.
4. Claridge B, Kastaniegaard K, Stensballe A, Greening DW. Post-translational and transcriptional dynamics – regulating extracellular vesicle biology. Expert Rev Proteomics 2019 Jan 2 16(1):17–31.

226

An Array of 60,000 Antibodies for Proteome-Scale Antibody Generation and Target Discovery

Sheng-ce Tao[1], Zhaohui Wang[2], Yang Li[1], Bing Hou[2], Xun Meng[2]

1. Shanghai Jiao Tong University, Minghang District, SHANGHAI, China

2. The College of Life Sciences, Northwest University, Xi'an, Shanxi , China

Antibodies are essential for elucidating the roles of genes decoded by genome sequencing. However, affordable technology for proteome-scale antibody generation does not exist. To address this, we developed the Proteome Epitope Tag Antibody Library (PETAL) and its array. PETAL consists of 62,208 mAbs against 15,199 peptides from diverse proteomes. PETAL harbors binders for a great multitude of proteins in nature due to antibody multispecificity, an intrinsic feature of an antibody. Distinctive combinations of 10,000-20,000 mAbs were found to target specific proteomes by array screening. Phenotype-specific mAb-target pairs were discovered for maize and zebrafish samples. Immunofluorescence and flow cytometry mAbs for human membrane proteins and ChIP-seq mAbs for transcription factors were identified from respective proteome-binding PETAL mAbs. Differential screening of cell surface proteomes of tumor and normal tissues discovered internalizing tumor antigens for antibody-drug conjugates. By discovering high affinity mAbs at a fraction of current time and cost, PETAL enables proteome-scale antibody generation and target discovery.

227

Proteome-wide detection of cysteine nitrosylation targets and motifs using bioorthogonal cleavable-linker-based enrichment and switch technique (Cys-BOOST)

Ruzanna Mnatsakanyan[1], Stavroula Markoutsa[1], Kim Walbrunn[1], Andreas Roos[1, 2], Steven H. L. Verhelst[1, 3], René P. Zahedi[4, 1, 5]

1. Leibniz-Institut für Analytische Wissenschaften – ISAS - e.V., Dortmund, Germany

2. Department of Neuropediatrics, University Hospital Essen, Essen, Germany

3. Laboratory of Chemical Biology, KU Leuven, Leuven, Belgium

4. Segal Cancer Proteomics Centre, Montreal, QC, Canada

5. Gerald Bronfman Department of Oncology, Jewish General Hospital, Montreal, Canada

Cysteine modifications emerge as important players in cellular signaling and homeostasis. Here, we present a chemical proteomics strategy for quantitative analysis of reversibly modified Cysteines using bioorthogonal cleavable-linker and switch technique (Cys-BOOST). Compared to iodoTMT for total Cysteine analysis, Cys-BOOST shows a threefold higher sensitivity and considerably higher specificity and precision.

Using iodoTMT-based enrichment we quantified 9,966 Cys peptides and 3,446 background peptides (74% specificity), compared to 25,019 Cys peptides and only 581 background peptides for Cys-BOOST (98% specificity). We assessed the technical reproducibility of both workflows through individual processing of technical replicates and obtained relative standard deviations of 36 % (iodoTMT) vs 9 % (Cys-BOOST). We compared the average scaled TMT reporter intensities (sum of all TMT reporter intensities of Cys containing PSMs/number of Cys containing PSMs) of all Cys containing PSMs quantified by either Cys-BOOST or iodoTMT. The average scaled TMT intensity observed for Cys-BOOST was around 4 times higher, despite the higher number of Cys peptides detected with Cys-BOOST which may come along with the identification of many low abundant peptides. These results indicate a considerably higher recovery, which comes along with more precise quantification.

Analyzing S-nitrosylation (SNO) in S-nitrosoglutathione (GSNO)-treated and non-treated HeLa extracts Cys-BOOST identifies 8,304 SNO sites on 3,632 proteins covering a wide dynamic range of the proteome. Consensus motifs of SNO sites with differential GSNO reactivity confirm the relevance of both acid-base catalysis and local hydrophobicity for NO targeting to particular Cysteines. Applying Cys-BOOST to SH-SY5Y cells, we identify 2,151 SNO sites under basal conditions and reveal significantly changed SNO levels as response to early nitrosative stress, involving neuro(axono)genesis, glutamatergic synaptic transmission, protein folding/translation, and DNA replication. Our work suggests SNO as a global regulator of protein function akin to phosphorylation and ubiquitination.

1. Proteome-wide detection of S-nitrosylation targets and motifs using bioorthogonal cleavable-linker-based enrichment and switch technique. Mnatsakanyan R, Markoutsa S, Walbrunn K, Roos A, Verhelst SHL, Zahedi RP. Nat Commun. 2019 May 16;10(1):2195. doi: 10.1038/s41467-019-10182-4.

228

System-wide identification of enzyme substrates by thermal analysis

Amir Ata Saei[1], Christian Beusch[1], Juan Astorga Wells[1], Pierre Sabatier[1], Alexey Chernobrovkin[1], Sergey Rodin[1], Katja Näreoja[1], Ann-Gerd Thorsell[1], Tobias Karlberg[1], Qing Cheng[1], Massimiliano Gaetani[1], Ákos Végvári[1], Elias S.J. Arnér[1], Herwig Schüler[1], Roman A. Zubarev[1]

1. Karolinska Institutet, Stockholm, STOCKHOLM, Sweden

Despite the immense importance of enzyme-substrate reactions, there is a lack of generic and unbiased tools for identifying the molecular components participating in these reactions on a cellular level. Here we developed a universal method called System-wide Identification of Enzyme Substrates by Thermal Analysis (SIESTA). The approach assumes that enzymatic post-translational modification of substrate proteins changes their thermal stability. SIESTA successfully identified several known and novel substrate candidates for selenoprotein thioredoxin reductase 1, as well as poly-(ADP-ribose) polymerase-10 and protein kinase B (AKT1) systems. A number of putative substrates for each enzyme system were confirmed by targeted mass spectrometry and functional assays. Wider application of SIESTA can enhance our understanding of the role of enzymes in homeostasis and disease, open new opportunities in investigating the effect of PTMs on protein stability, and facilitate drug discovery.

229

Mass Spectrometry Imaging of RTKi with protein target in human lung cancer tumor xenograft mouse

Tae young Kim[1], Yonghyo Kim[2], Yutaka Sugihara[2], Melinda Rezeli[2], Marcell Szasz[3], György Marko-Varga[2], HoJeong Kwon[1]

1. Yonsei university, Yonsei-ro Seodaemungu, SEOUL, South Korea

2. Lund University, Lund

3. National Koranyi Institute, Budapest

Receptor tyrosine kinase inhibitor (RTKi) has been widely used to treat NSCLC (non-small-cell lung carcinoma) patients. Currently, RTKi has been known to acquire resistance in the patients treated leading to low efficacy. Many studies have aimed to unveil the cause of resistance by exploring the molecular mechanism *in vitro*. However, low efficacy of RTKi in tumor growth *in vivo* have not been fully addressed so far. Here, we propose the novel evidence of RTKi's low efficacy through analysis of its distribution image using MALDI MSI (matrix-assisted laser desorption ionization imaging) and IHC (Immunohistochemistry) of a protein target of RTKi. We found that H1299, human lung cancer cells, show weak growth inhibition upon RTKi treatment leading to low efficacy in H1299 tumor xenograft model mouse. From the model mouse, we identified that RTKi was distributed on RTKi-treated tumor, liver and kidney tissue co-localizing with its target protein. Additionally, RTKi localized higher in liver and kidney than tumor tissue from the analysis of quantification analysis. Collectively, these results demonstrated that low efficacy of RTKi in H1299 tumor xenograft model comes in part from its nonspecific distribution in a number of organs.

230

An enhanced boosting to amplify signal with isobaric labeling (eBASIL) approach toward comprehensive quantitative single-cell proteomics analysis

Chia-Feng Tsai[1], Rui Zhao[1], Maowei Dou[1], Ronald J Moore[1], Karin D Rodland[1], Richard D Smith[1], Tujin Shi[1], Ying Zhu[1], Tao Liu[1]

1. Pacific Northwest National Laboratory, Richland, WASHINGTON, United States

Mass spectrometry (MS)-based proteomics has tremendous potential to overcome limitations of popular flow cytometry and mass cytometry methods and achieve antibody-independent, comprehensive, and quantitative proteomics analysis of single cells. Such potential has yet to be realized, however, due mainly to ineffective sample processing as well as sample losses prior to MS analysis. To tackle this issue, we recently developed several approaches to significantly improve sample handling and processing, multiplexing, and fractionation, which are being integrated into a robust platform for ultra-sensitive quantitative single-cell proteomics analysis. Individual (i.e. single) cells are first sorted to a nanoPOTS (nanodroplet processing in one pot for trace samples) chip for nearly lossless sample processing, including protein extraction, reduction and alkylation, digestion, and isobaric labeling using 11-plexed tandem mass tags (TMT) reagents, all of which take place in 200-nanoliter droplets and are controlled using a picoliter-resolution robotic system. The resulting samples are then subjected to capillary solid-phase extraction C18 cleanup and nanoscale fractionation and concatenation using a nanoFAC (nanoflow fractionation and automated concatenation) system. An enhanced boosting to amplify signal with isobaric labeling (eBASIL) approach is used for greatly increasing the peptide signal from the single cells. We carefully evaluated and optimized MS data acquisition parameters and potential bias introduced by isotopic impurity of the TMT reagents. Using the optimized conditions, a boosting-to-sample ratio of 1000 can be used to provide ~1500 protein identifications (with no fractionation) without compromising the ability to robustly quantify the proteins in three different acute myeloid leukemia cells. We anticipate the proteome coverage can be further increased with the use of nanoFAC (in progress). We believe this novel integrated platform will enable broad application of single-cell proteomics analysis in biological and biomedical research. It also has the potential to be adapted for phosphoproteome analysis of very small numbers of cells.

232

Large scale crosslinking mass spectrometry of the yeast nucleus reveals insights into data analysis considerations for protein-protein interaction discovery

Tara K Bartolec[1], Daniela-Lee Smith[1], Joshua J Hamey[1], Marc Wilkins[1]

1. Systems Biology Initiative, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Randwick, NSW, Australia

Proteome scale crosslinking mass spectrometry (XL-MS) couples the use of MS-cleavable cross linkers with specialised mass-spectrometry (MS) methods to generate high-throughput and rich protein-protein interaction (PPI) networks from complex samples. *Saccharomyces cerevisiae* is an extensively studied model organism for eukaryotic cell biology. Efforts to systematically perform yeast-2-hybrid and affinity pulldowns of all yeast proteins have generated a comprehensive resource of experimentally derived evidence for PPIs. This makes it an ideal organism to investigate the quality of large scale XL-MS derived PPIs. Here we present the first large-scale yeast nuclear XL-MS dataset. Intact nuclei from actively dividing wild-type yeast were isolated and then crosslinked using the cleavable crosslinker DSSO. Following digestion with Trypsin/LysC, offline strong cation exchange was employed to enrich and fractionate crosslinked peptides. Resulting fractions were analysed on two-hour LC-gradients coupled to tandem MS methods using sequential CID-ETD-MS2 fragmentation and CID-MS3. Two independent biological replicates were collected and analysed, with data processed using Proteome Discoverer 2.3 and inbuilt XlinkX 2.0. A highly interconnected dataset was generated containing over 2,000 unique lysine-lysine linkages at 1% FDR. Over 700 were interlinks, which represent PPIs - three quarters involved nuclear proteins, indicating a nuclear-enriched dataset. 32% of interlinks were previously observed physical interactions, with a further 25% representing predicted interactions from STRING or genetic interaction screens. We observed a distinct and systematic skew towards lower XlinkX scores for interlinks without existing evidence, and show that these are disproportionately excluded when optimised and more stringent XlinkX settings are employed. However, this increased stringency did not affect the quality of intralinks when mapping to protein structures. This dataset directly contributes to ongoing discussion on the complex issue of FDR in large scale XL-MS studies, and represents a significant foundation for quantitative and dynamic XL-MS studies of the yeast nuclear interactome.

233

Global High Resolution Mapping of Nucleotide Base Excision Repair Mechanism

Mohamed Elzek[1], Eneko Villaneuva[1], Tom Smith[1], Rayner Queiroz[1], Anne Willis[2], Kathryn Lilley[1]

1. Biochemistry, University of Cambridge, Cambridge, Cambridgeshire, United Kingdom

2. MRC- Toxicology Unit, University of Cambridge, Cambridge, Cambridgeshire, United Kingdom

Background

In response to genotoxic stress, intricate protein regulatory mechanisms play a key determinant role sensing DNA damage, recruiting DNA repair factors and orchestrating cell-cycle response. To date, an assessment of protein re-localisation upon DNA damage has not been performed with in-depth subcellular resolution. Such a study is important to identify novel re-localisation events. Hyperplexed Localisation of Organelle Proteins by Isotope Labelling (hyperLOPIT) is a leading technology for in-depth interrogations of the proteome subcellular distribution. However, despite the highest sub-cellular resolution of any spatial proteomics method, the hyperLOPIT protocol requires high starting cell numbers that are not always achievable[1]. Here, we present a one-step subcellular fractionation procedure, termed “dynamic LOPIT” (“dLOPIT”), a highly reproducible method that economises on sample processing time and starting material requirements, while maintaining comparable organelle resolution and protein identifications. This approach has paved the way for robust spatiotemporal system-wide proteome analysis of Nucleotide Excision Repair (NER) mechanism.

Methodologies

NER response was triggered in the non-transformed breast epithelial cell line MCF10A by mid-range ultraviolet (302 nm UV-B). Spatial proteome profiling was then performed by combining:

1. Organelle separation by density ultracentrifugation
2. Sample multiplexing by in-vitro covalent labelling with amine-reactive tandem mass tag (TMT) reagents
3. SPS-MS3 acquisition on the Orbitrap Fusion Lumos Tribrid mass spectrometer
4. Multivariate statistical analysis within the pRoloc and MSnbase packages[2,3].

Prediction of the steady state location of proteins of unknown location was achieved by applying feed forward neural networks trained with highly curated organelle markers.

Findings

The dLOPIT approach provides an unprecedented high resolution map of proteins re-localisation across 14 discrete organellar and sub-organellar compartments. Analysis of these data gives insight into cellular response underlying UV mediated nucleotide base excision repair mechanism.

Concluding statement

dLOPIT provides a robust, efficient and scalable alternative to hyperLOPIT protocol to study the spatiotemporal dynamics of DNA damage response.

1. 1. Christoforou, A. et al. A draft map of the mouse pluripotent stem cell spatial proteome. Nat. Commun. 7, 9992 (2016). 2. Gatto, L., Breckels, L. M., Wieczorek, S., Burger, T. & Lilley, K. S. Mass-spectrometry-based spatial proteomics data analysis using pRoloc and pRolocdata. Bioinformatics 30, 1322–1324 (2014). 3. Breckels, L. M., Mulvey, C. M., Lilley, K. S. & Gatto, L. A Bioconductor workflow for processing and analysing spatial proteomics data. F1000Research 5, 2926 (2016).

234

Modulation of phosphorylation and cysteine modifications upon early T-cell activation

Komal K K Mandal[1], Martin R Larsen[1]

1. Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

T-cell activation result in activation of a coordinated protein based signaling network. The response is extremely specific and fast and consists of complex interplays of feedback and feedforward loops, parallel pathways and PTM crosstalk. Proteins orchestrating these events must be fast and reversible, which is best achieved by using PTM's. Phosphorylation is well established to drive the T-cell signaling, however other PTMs are very likely to contribute significantly. In this study, we have uncovered an additional layer of T-cell signaling through reversible cysteine (Cys) modifications integrated with protein phosphorylation.

Jurkat cells were stimulated for 0S, 30S, 60S and 5min using 1:1 mixture of pre-crosslinked anti-CD3 and anti-CD28. The stimulation was quenched using boiling SDS to 1% and the solution was boiled for 5 min at 95°C followed by probe sonication. The free cysteines were blocked with CysPAT and the proteins digested with trypsin after precipitation. Tryptic peptides were labeled using TMT11/16plex. The peptides containing reversible cys modifications were isolated sequentially using various reduction reagents and finally the CysPAT labeled cysteines and phosphopeptides were isolated. Isolated peptides were analyzed by LC-MS/MS and further subjected to bioinformatic analysis.

A large coverage of the T-cell signaling pathway was recorded and multiple known and new phosphorylation sites were mapped to the short T-cell activation. Specifically, we have identified the two active site cysteines in the PTPRC regulated significantly after stimulation. PTPRC, upon T-cell activation recruits and dephosphorylates SKAP1, FYN, LYN and CD receptors and probably also other members of the signaling pathway. Numerous other reversible Cys peptides were found to be significantly modulated after stimulation.

We have successfully used TMT multiplexing in combination with a comprehensive workflow for targeting phosphorylation and reversible cysteine modification to study short time T-cell activation. The crosstalk between the PTMs could be crucial for the mechanism of T-cell activation.

235

Proximity-dependent approach for identifying putative endogenous substrates of protein kinases

Tomoya Niinae[1], Koshi Imami[1], Chia-Feng Tsai[2], Naoyuki Sugiyama[1], Yasushi Ishihama[1]

1. Kyoto University, Kyoto, Japan

2. Pacific Northwest National Laboratory, Richland, United States

Background

MS-based phosphoproteomics allows identifying more than 10,000 phosphorylated sites in a single experiment. Despite enormous phosphosite information has been accumulated in public repositories, protein kinase-substrate relationships remain largely unknown. This is because it is difficult to detect kinase-substrate complexes by conventional methods such as co-immunoprecipitation due to the transient interactions between them. Recently, proximity-dependent biotin identification (BioID) has been developed, in which transient protein-protein interactions as well as stable ones can be captured. In this study, we developed a novel strategies for identifying kinase endogenous substrates with BioID, in combination with phosphoproteome analysis of kinase-perturbed cells and the sequence motifs to exclude indirect substrates.

Methods

We performed BioID experiment for CK2. At first, BirA fused CK2 was cloned and transfected into HEK293T cells. After treating biotin to the cell cultures for 24 h, biotinylated proteins were extracted from cells, purified with streptavidin beads, digested into tryptic peptides and analyzed by nanoLC/MS/MS. For phosphoproteome profiling, HEK293T cells were treated with CX-4945. Proteins were extracted from cells and were digested by trypsin. Phosphopeptides were enriched with titania chromatography, labeled with TMT reagents and analyzed by nanoLC/MS/MS.

Results

At first, in the BioID experiment, 574 CK2 interactors were identified. Then, we analyzed phosphoproteome of kinase-perturbed cells. As a result, 1,959 phosphosites were quantified from CX-4945-treated cells. Among them, 262 phosphosites were down-regulated by CX-4945 and 76 out of 262 phosphosites were located on 59 CK2 interactors. Finally, we calculated Primary Sequence Preference (PSP) scores to estimate how likely the substrate candidates are phosphorylated by the target kinases. By integrating these three filters described above, 41 putative CK2 substrates were selected. We also applied this approach to tyrosine kinases.

Conclusions

We developed a novel strategy for identifying kinase endogenous substrates, and applied this strategy to tyrosine kinases.

25

Allelic association with ankylosing spondylitis fails to correlate with HLA-B27 homodimer formation

Terry CC Lim Kam Sian[1], Saranjah Indumathy[1], Hanim Halim[1], Anja Greule[1], Max J Cryle[1], Paul Bowness[2], Jamie Rossjohn[1], Stephanie Gras[1], Anthony W Purcell[1], Ralf B Schittenhelm[1]

1. Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC, Australia

2. Department of Orthopaedics Rheumatology and Musculoskeletal Science, University of Oxford, Oxford , UK

Expression of HLA-B27 is strongly associated with predisposition towards ankylosing spondylitis (AS) and other spondyloarthropathies. However, the exact involvement of HLA-B27 in disease initiation and progression remains unclear. The 'homodimer theory', which proposes that HLA-B27 heavy chains aberrantly form homodimers, is a central hypothesis that attempts to explain the role of HLA-B27 in disease pathogenesis.

Here we examined the ability of the 8 most prevalent HLA-B27 allotypes (HLA-B*27:02 - HLA-B*27:09) to form homodimers. We observed *in cellulo* that HLA-B*27:03 – a disease-associated HLA-B27 subtype – showed a significantly reduced ability to form homodimers in comparison to all other allotypes, including the non-disease-associated/protective allotypes HLA-B*27:06 and HLA-B*27:09. We used X-ray crystallography in combination with site-directed mutagenesis to unravel the molecular and structural mechanisms in HLA-B*27:03 that are responsible for its compromised ability to form homodimers. We show that the polymorphism at position-59, which differentiates HLA-B*27:03 from all other HLA-B27 allotypes, is responsible for its "compromised" ability to form homodimers. Indeed, Histidine-59 in HLA-B*27:03 leads to a series of local conformational changes that act in concert to reduce the accessibility of the nearby cysteine-67, an essential amino acid residue for the formation of HLA-B27 homodimers.

The ability of both protective and disease-associated HLA-B27 allotypes to form homodimers, and the failure of HLA-B*27:03 to form homodimers challenges the role of HLA-B27 homodimers in AS pathoetiology and rather implicates other features and mechanisms intrinsic to HLA-B27 to be pivotal for disease pathogenesis.

26

Development of an autoantigen microarray for the screening of novel autoantibodies in psoriatic arthritis.

Tianfu Wu[1], Yulin Yuan[1]

1. University of Houston, Houston, TEXAS, United States

Objective: The autoimmune etiology in psoriasis is not very clear, we aim to identify autoantigens and autoantibodies in psoriasis, which may shed light on the molecular and cellular basis of the pathogenesis of psoriasis and psoriatic arthritis.

Methods: In this study, we developed an autoantigen array system harboring a variety of antigens including typical autoantigens in rheumatic diseases as well as skin antigens, inflammatory mediators and putative autoantigens in psoriasis. Sera from psoriasis patients (N = 73) were used to interrogate antigens on the array. Individual ELISA was also used in validation studies.

Results: We found several serum autoantibodies were elevated in psoriasis patients compared to healthy controls; particularly, IgG autoantibodies against two novel antigens, LL37 and ADAMTSL5, were significantly increased in the psoriasis patients compared to healthy controls, $P < 0.001$, respectively. Importantly, Psoriasis Area and Severity Index (PASI) was found to be correlated with serum levels of IgG autoantibodies against LL37 ($r = 0.45$, $P < 0.01$) and ADAMTSL5 ($r = 0.48$, $P < 0.01$). Both autoantibodies also reflected disease progression in longitudinally collected samples from psoriasis patients. Importantly, we found both anti-ADAMTSL5 and anti-LL-37 autoantibodies were significantly elevated in psoriatic arthritis (PsA, N = 22) compared to Non-PsA (N = 32), suggesting that these molecules may be involved in the pathogenesis of psoriatic arthritis.

Conclusion: Our findings suggest that these autoantibodies may be useful biomarkers and indicative of therapeutic targets of psoriasis and psoriatic arthritis.

27

HLA-B27 immunopeptidome: Understanding the connection between *Salmonella typhimurium* infection and ankylosing spondylitis

Terry C.C Lim Kam Sian[1], Pouya Faridi[1], Ralf B Schittenhelm[1, 2], Anthony W Purcell[1]

1. Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

2. Monash Biomedical Proteomics Facility, Monash Biomedicine Discovery Institute, Monash University, Clayton, Australia

The association of the Human Leukocyte Antigen (HLA) B27 with ankylosing spondylitis (AS) is one of the strongest associations that has ever been described between an autoimmune disease and a HLA allele. One of the leading hypotheses is the arthritogenic peptide theory, which assumes that molecular mimicry between a foreign- and a self-peptide leads to breakdown of the immune tolerance and ensuing autoimmunity. A major support for this theory comes from the observation that patients often develop AS following gastrointestinal bacterial infection. To identify putative arthritogenic peptides, we analysed both the linear and spliced HLA-B27 immunopeptidome before and after infection with *Salmonella typhimurium*, which represents one of the bacterial strains associated with AS.

Methods: A combination of high-resolution mass spectrometry, *de novo* sequencing and cutting-edge bioinformatics was used to identify the linear and spliced peptides presented on the 8 most common HLA-B27 allotypes (HLA-B*27:02-HLA-B*27:09) before and after infection with *S.typhimurium*.

Results: We identified ~13500 unique HLA-B27 peptides of which ~4500 have never been reported previously. A careful comparison of these peptides showed that bacterial infection did not impact on the HLA-B27 immunopeptidome with regards to the peptide length preference or the consensus binding motif. A comparative analysis against bacteria that have been reported to be associated with AS, revealed 460 novel putative arthritogenic peptides. Interestingly, <1% of the linear peptides identified were *Salmonella*-derived. In contrast, *Salmonella* spliced peptides represented a higher proportion of the immunopeptidome of ~4% (ratio of cis:trans of 4:1) and were derived from a different subset of source proteins than the linear peptides.

Conclusion: We describe 460 novel putative arthritogenic peptides and further show that spliced peptides can increase the breath of the immunopeptidome. However, further analyses such as peptide binding studies, immunological assays and structural investigations are required to confirm immunogenicity and to validate these peptides.

28

ThyroProt: a robust protein classification system for thyroid disease based on deep learning of proteome data

Yaoting Sun[1, 2], Sangeeta Mantoo[3], Zelin Zang[4], Zhicheng Wu[1, 2], Wei Liu[1, 2, 5], Yi Zhu[1, 2], Lirong Chen[6], Xiaodong Teng[7], Yongfu Zhao[8], Tony Kiat-Hon Lim[3], Bhuvaneswari Hariraman[9], Serene Yeow[10], Syed Muhammad Fahmy bin Syed Abdillah[3], Sze Sing Lee[10], Xue Cai[1, 2], Guangzhi Wang[8], Yi He[11], Xu Zheng[12], Tiansheng Zhu[1, 2], Qiushi Zhang[1, 2], Weigang Ge[1, 2], Huanhuan Gao[1, 2], Guan Ruan[1, 2], Yulong Kong[1, 5], Linyan Wang[13], Xi Zheng[14], Kailun Xu[14], Shu Zheng[14], Ziqing Li[4], Oi Lian Kon[10], N. Gopalakrishna Iyer[9, 10], Tiannan Guo[1, 2]

1. Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou

2. Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou

3. Department of Pathology, Singapore General Hospital, Singapore

4. School of Engineer, Westlake University, Hangzhou

5. Department of Clinical Pharmacology, College of Pharmacy, Dalian Medical University, Dalian

6. Department of Pathology, The Second Affiliated Hospital of College of Medicine, Zhejiang University, Hangzhou

7. Department of Pathology, the First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou

8. Department of General Surgery, The Second Hospital of Dalian Medical University, Dalian

9. Division of Surgical Oncology, National Cancer Centre Singapore, Singapore

10. Division of Medical Sciences, National Cancer Centre Singapore, Singapore

11. Department of Urology, The Second Hospital of Dalian Medical University, Dalian

12. Liaoning Laboratory of Cancer Genetics and Epigenetics and Department of Cell Biology, College of Basic Medical Sciences, Dalian Medical University, Dalian

13. Department of Ophthalmology, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou

14. Cancer Institute, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou

Thyroid nodules are a common endocrine disease affecting approximately 50% of the global population. Fifteen to 30% cannot be confidently diagnosed as benign or malignant nodules by cytopathology before surgery, which may lead to overtreatment. Although molecular diagnostic tests for thyroid nodules have been developed, they are either rule-out or rule-in tests due to low predictive values and other limitations, such as RNA degradation. In this retrospective, blinded and multi-center study, we aimed to characterize benign and malignant thyroid tissues by using pressure cycling technology (PCT) coupled with data independent acquisition (DIA) mass spectrometry. In total, 1,984 formalin-fixed paraffin-embedded (FFPE) punches of five different histological types of thyroid nodules from 826 patients from four clinical centers were analyzed using the PCT-DIA method. In the discovery phase of the study, a classifier model (ThyroProt) to distinguish benign and malignant thyroid nodules based on a panel of 14 promising protein biomarker candidates was established from a patient cohort of 579 patients and 1,793 DIA data files by deep learning. In the blinded validation phase of the study, we tested ThyroProt on an independent cohort of 247 thyroid nodules with 494 DIA data files. ThyroProt correctly identified the benign or malignant status of 230 of 248 thyroid nodules with an accuracy of 92%, showing considerable superiority compared to published accuracies based on DNA and/or RNA analysis, such as 65% by Afirma gene expression classifier, 78% by genome sequencing classifier, and 84% by ThyroSeqV3, respectively. ThyroProt achieved high sensitivity, specificity, positive predictive value and negative predictive value of 90%, 94%, 94%, and 90%, respectively. More importantly, ThyroProt is applicable to FFPE tissues which are the most abundant and robust samples in clinical practice. In summary, ThyroProt is superior both in ruling-in and ruling-out cancer in thyroid nodules compared to extant diagnostic molecular tests.

31

Proteome analysis of *Streptococcus suis* under stress conditions and in host-pathogen interaction

Denise Dittmar[1], Christine Weiße[2], Beata Jakobczak[3], Petra Hildebrandt[1], Ulrike Mäder[1], Manuela Gesell Salazar[1], Stephan Michalik[1], Volker Florian[3], Christoph Georg Baums[2], Uwe Völker[1]

1. Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany

2. Institute of Bacteriology and Mycology, College of Veterinary Medicine, University of Leipzig, Leipzig, Germany

3. Animal Health , IDT Biologika GmbH, Dessau-Rosslau, Germany

Background. *Streptococcus suis* is a commensal of pigs, but can also cause invasive infections. It is responsible for high economic losses in swine farming worldwide. As an emerging zoonotic agent *S. suis* is able to induce meningitis, arthritis and septicemia. The species comprises 35 serotypes of which serotype 2 is the most prevalent serotype [1], followed by serotypes 9, 7 and 3 in Europe [2]. The mechanisms enabling the switch from commensal to an invasive pathogen are only partially resolved. We performed comparative profiling of *S. suis* proteome patterns and immunoproteome screening of *S. suis* antigens to explore adaptation to host niches and to identify new potential virulence/fitness factors and vaccine candidates.

Methods. *S. suis* proteomes were analyzed for serotypes 2, 9 and 7 from different growth phases using a data-independent acquisition (DIA) mass spectrometry workflow [3]. In addition, *S. suis* specific antibody profiles were recorded using recombinantly expressed proteins in a multiplexed suspension bead array.

Results. Initially an in-house spectral library was generated using strain specific genome sequences and *S. suis* protein extracts from samples grown in different media, at different temperatures and after nutrient or iron limitation. Using a DIA-workflow and an optimized peptide preparation routinely around 1100 proteins were quantitatively profiled from low bacteria numbers (10^6 – 10^7) from rich medium, *ex vivo* samples (pig CSF and plasma) or after recovery from infection by cell sorting. Complementing this proteomics profiling the antibody response of infected pigs against a diverse range of *S. suis* antigens was recorded by bead-based immunoproteomics. This revealed differentially abundant proteins between these conditions and immunogenic proteins, which might be involved in adaptation to the different specific host niches and constitute potential new virulence factors and vaccine candidates.

Conclusion. The identified proteins provide a promising basis to evaluate conserved immunogenic antigens for multicomponent vaccines.

1. [1] Rungelrath V et al., Virulence 2018
2. [2] Fittipaldi N et al., Future Microbiol. 2012
3. [3] Michalik S et al., Scientific Reports 2017

32

Leveraging of extensive inter-species homologies to study plasma proteomes of bovids using data-independent acquisition

Pawel Sadowski[1], Zainab Noor[2], Selvam Paramasivan[3], Priya Ghodasara[3], Saul Chemonges[3], Nana Satake[3], Rajesh Gupta[1], Steven Kopp[3], Paul Mills[3], Shoba Ranganathan[2]

1. Queensland University of Technology, Brisbane, QLD, Australia

2. Department of Molecular Sciences, Macquarie University, Sydney, NSW, A

3. The School of Veterinary Science, The University of Queensland, Brisban, QLD, Australia

Blood plasma is an attractive source of proteins for biomarker studies in humans and animals. Being a composite of multiple organ specific proteomes, it is paramount in assessing individual's state of health. Yet, it is one of the most challenging samples for modern mass spectrometry. Recent studies have shown that novel data acquisition strategies, such as SWATH-MS, enable quantitation of hundreds of proteins in non-depleted human plasma. In contrast to human studies however, the SWATH-MS based plasma proteomics in animals is virtually non-existent. Studying animal plasma comes with unique difficulties, such as a lack of protein databases among other challenges. We provide the summary of our 3-year work that resulted in the development of extensive plasma peptide spectral repositories for various domestic animal species, mostly bovids (sheep, cattle), as well as workflows that allow us to reproducibly quantify ~300 proteins in non-depleted plasma using SWATH-MS. We also present a strategy that leverages on the extensive homology between different bovid species to increase the depth of spectral libraries and therefore, increase the number of proteins that can be quantified in plasma. Finally, we present a data analysis pipeline that combines iSwathX, Skyline and in-house scripts to quantify similarly high number of proteins in plasma of non-domestic/exotic ungulate species, for which genome information is not available, and for which building extensive peptide spectral repositories would be unfeasible.

To our knowledge, this is the largest repository of animal plasma proteome as well as the largest number of proteins quantified directly in non-depleted animal plasma. It is also the first example of cross-species utilization of the libraries. Our work has direct implication in wildlife studies where limited access to samples prevents establishing relevant proteomics resources.

33

Proteomic profiling of crocodile spermatozoa refutes the tenet that posttesticular maturation is restricted to mammals

David A Skerrett-Byrne[1, 2], Stephen D Johnston[3], Matthew D Dun[5, 4], Brett Nixon[1, 2]

1. Priority Research Centre for Reproductive Science, School of Environmental and Life Sciences, Faculty of Science, The University of Newcastle, Callaghan, NSW, Australia

2. Pregnancy & Reproduction Research Program, Hunter Medical Research Institute, Newcastle, New South Wales, Australia

3. School of Agriculture and Food Science, The University of Queensland, Gatton, Queensland, Australia

4. Priority Research Centre for Cancer Research, Innovation and Translation, School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, University of Newcastle, Callaghan, New South Wales, Australia

5. Cancer Research Program, Hunter Medical Research Institute, Newcastle, New South Wales, Australia

Competition to achieve paternity has contributed to the development of a multitude of elaborate male reproductive strategies. In one of the most well-studied examples, the spermatozoa of all mammalian species must undergo a series of physiological changes, termed capacitation, in the female reproductive tract prior to realising their potential to fertilise an ovum. However, the evolutionary origin and adaptive advantage afforded by capacitation remains obscure. Here, we report the use of comparative and quantitative proteomics to explore the biological significance of capacitation in an ancient reptilian species, the Australian saltwater crocodile (*Crocodylus porosus*). Our data reveal that exposure of crocodile spermatozoa to capacitation stimuli elicits a cascade of physiological responses that are analogous to those implicated in the functional activation of their mammalian counterparts. Indeed, among a total of 1,119 proteins identified in this study, we detected 126 that were differentially phosphorylated (± ≥1.2 fold-change) in capacitated versus non-capacitated crocodile spermatozoa. Notably, this subset of phosphorylated proteins shared substantial evolutionary overlap with those documented in mammalian spermatozoa, and included key elements of signal transduction, metabolic and cellular remodelling pathways. Unlike mammalian sperm, however, we noted a distinct bias for differential phosphorylation of serine (as opposed to tyrosine) residues, with this amino acid featuring as the target for ~80% of all changes detected in capacitated spermatozoa. Overall, these results indicate that the phenomenon

of sperm capacitation is unlikely to be restricted to mammals and provide a framework for understanding the molecular changes in sperm physiology necessary for fertilisation.

1. Nixon B, Johnston SD, Skerrett-Byrne DA, Anderson AL, Stanger SJ, Bromfield EG, et al. Proteomic profiling of crocodile spermatozoa refutes the tenet that post-testicular maturation is restricted to mammals. Molecular & Cellular Proteomics. 2018:mcp.RA118.000904.

34

Proteomic characterisation of Ancient Egyptian skin, bones and textiles

Dylan Multari[1], Prathiba Ravishankar[1], Constance Lord[2], James Fraser[2], Ronika K. Power[3], Paul A. Haynes[1]

1. Department of Molecular Sciences, Macquarie university, North Ryde, NSW 2109, Australia

2. Nicholson Museum, University of Sydney, Camperdown, NSW 2006, Australia

3. Department of Ancient History, Macquarie University, North Ryde, NSW 2109, Australia

We were among the first to publish data concerning identification of proteins from 4200-year-old ancient Egyptian skin fragments, sampled from the collection of the Egyptian Museum in Turin [1]. Following on from this work, we are currently undertaking an analysis of ancient Egyptian skin, bone and textile samples, as part of the University of Sydney interdisciplinary mummy project. Sample materials have been made available from the coffin and remains of Mer-Neith-it-es, a 26th dynasty (664-525 BCE) mummy excavated from Saqqara, Egypt. The remains are heavily fragmented, likely due to frequent incidences of grave robbing and the transportation of the coffin by sea to Sydney in 1859.

We will present results of proteomic analysis of samples from two different types of textiles found with the remains, one of which contains finer strands than the other, which may be indicative of different source materials. One of the aims of our study was to develop a non-invasive technique for sampling ancient organic remains, using readily available equipment and materials. We set out to examine whether dermatology grade skin sampling strip tape could be successfully applied to the analysis of such materials. Successful development of a nondestructive sampling method would be a big step forward in bioarchaeological proteomics, because it would enable access to a much wider range of ancient materials housed in Museum collections.

Preliminary experiments have shown that we can identify intracellular protein components on the surface of skull fragments, which strongly suggests that they are indeed ancient remains rather than modern contamination. This presentation will include detailed results of proteomic analysis of skull and bone fragments using this novel nondestructive sampling approach.

[1] Jones J. *et al*. Philosophical Transactions of the Royal Society part A. 2016 Oct 28;374(2079).

37

Proteomic Cartography and Biomarkers at the Single-Cell Level: Interrogation of Premalignant and Early Stage Lesions

Jacob Kagan[1], Sudhir Srivastava[1]

1. Cancer Biomarkers Research Group, National Cancer Institute, National Institutes of Health, GAITHERSBURG, MARYLAND, United States

Currently the gold standard for diagnosis and classification of cancer is based on histopathological examination under the microscope of Hematoxylin and Eosin (H&E) stained slides from cancer biopsies and resected tumors. More targeted examination is based on immunohistochemical examination of a limited number of markers or examination of a limited number of targeted gene expression by *in situ* hybridization. Although such assays are very valuable to determine the grade and stage of the tumor, they provide limited information on the molecular and cellular content of a tumorigenic lesion, the physiological state of the cells within the lesion, and tumor heterogeneity. However, in many instances the molecular information is pivotal for proper diagnosis and prognosis. Single-cell proteomics is emerging as another powerful approach for phenotypic characterization of individual cells' type and their physiological state, protein quantitative measurements, and the detection of cancer related post-translational modifications (PTM), that cannot be predicted by genomic/transcriptomic analyses. The National Cancer Institute convened a Think-Tank meeting on the use of single cell proteomics in biomarkers discovery, lesion cartography, disease detection and classification. The Speaker will present the recommendations of this meeting for the benefit of HUPO community and discuss the future potential project for HUPO to consider

38

Integration of transcriptomics and antibody-based proteomics for spatial localization of cell type-specific expression patterns

Feria Hikmet Noraddin[1], Charles Pineau[2], Per Oksvold[3], Linn Fagerberg[3], Mathias Uhlen[3], Cecilia Lindskog[1]

1. Department of Immunology, Genetics and Pathology, Human Protein Atlas, Uppsala University, Uppsala, Sweden

2. Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) , Protim, Rennes, RENNES, France

3. Science for Life Laboratory, School of engineering Sciences in Chemistry, Biotechnology and health, , KTH - Royal Institute of Technology, Stockholm, STOCKHOLM, Sweden

The Tissue Atlas, generated by the Human Protein Atlas project, focuses on integrated omics for spatial localization of all human proteins down to the single cell level[1]. Genome-wide mRNA expression data is used for categorization of all human genes based on expression level and tissue distribution, combined with standardized immunohistochemistry for studying the protein localization in the context of neighboring cells. Recent advances of the Tissue Atlas include in-depth characterization of cell type-specific expression patterns, with main emphasis on the testis-specific proteome.

In the 2019 update of the public database www.proteinatlas.org, mRNA expression data from three different sources is merged and normalized for comprehensive categorization of all human genes in 37 different organs and tissues. The analysis shows that testis has by far the highest number of tissue-specific genes, however, many of the corresponding proteins lack a known function. Recently, the *in situ* expression of >500 proteins was characterized in eight different testicular cell types, allowing us to identify six distinct clusters of expression at different stages of spermatogenesis[2]. The analysis included numerous proteins previously classified as missing proteins (MPs). In a continued effort, we focus on single cell evaluation and spatial localization of >3,000 additional proteins with a cell type-specific expression in testis. By using multiplex immunofluorescence, the overlap between proteins with well-known testis functions and previously uncharacterized proteins can be determined, highlighting important targets for testis specific research.

Knowledge of the architecture of every human cell aids in identification of proteins that may accelerate research in molecular medicine, and contributes to further knowledge of underlying disease mechanisms. The Human Protein Atlas constitutes a comprehensive knowledge resource for gaining biological insight on human proteins. The publicly available datasets and high-resolution images allow for in-depth analysis of cell type-specific expression patterns and further exploration of MPs.

1. Uhlén M et al. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419. doi: 10.1126/science.1260419.
2. Pineau C et al. Cell type-specific expression of testis elevated genes based on transcriptomics and antibody-based proteomics. Submitted.

39

A Offset Mass Triggered Data Acquisition Approach to Single Cell Proteomics Experiments

Atticus McCoy[1], Tommy Cheung[1], Christopher M Rose[1]

1. Microchemistry, Proteomics, and Lipidomics, Genentech, South San Francisco, CA, United States

Mass spectrometry approaches to single cell proteomics (SCP) are gaining momentum as sample preparation methods and mass spectrometry analysis become more sensitive. One approach to SCP labels single cells with isobaric tags and includes a “carrier proteome” at relatively high excess to provide enough ions for peptide identification. Single cells are then quantified utilizing reporter ions 100-200x lower in intensity than the highest multiplexed channel. The carrier proteome channel helps in identification, but makes quantification of the experimental channels more challenging due to the limited dynamic range of the Orbitrap mass analyzer.

Here we introduce an offset mass triggered data acquisition approach that allows a carrier proteome channel to be used without impacting the intra-scan dynamic range of the Orbitrap mass analyzer. This inter-cluster scheme moves the “carrier proteome” peptides into a separate precursor isotopic cluster by utilizing super-heavy TMT (SH-TMT). The inter-cluster approach enables isolation of the experimental peptides without the isolation of the “carrier proteome” and ensures that all resulting reporter ions will be used for quantitation of experimental ratios. Using a recently described multiplexed targeted approach (TOMAHAQ) we explore the impact of the carrier proteome in the context of this inter-cluster approach and typical SCP experimental designs. We then demonstrate how a triggered offset mass approach could be implemented in data dependent acquisition, enabling it to be easily used in SCP experiments.

From these data we make recommendations for the utilization of carrier proteomes and experimental design of SCP experiments, as well as highlight key quantitative metrics (e.g., reporter ion signal-to-noise, injection time, etc.) to ensure meaningful biological results.

40

Super-resolution proteomics method to explore cell heterogeneity at single-cell level

Xinhao Shao[1], Mike Hollas[1], Yu (Tom) Gao[1]

1. University of Illinois at Chicago, Chicago, IL, United States

With current single-cell proteomics approaches, we can confidently detect the proteome of larger single cells, such as frog or fish embryos. Although still very difficult, smaller cells such as mammalian cells can also be detected by various methods. However, certain limitations such as low throughput, tedious sample handling and difficulties in quantification severely limit the use of single-cell proteomics on large-scale studies. Moreover, there are also many important cell species that are far smaller than embryos or mammalian cells. Current instrumentations and methods are simply not sensitive enough to detect the proteome from a single mitochondrion or a single bacterium. Here we present a novel super-resolution proteomics method to differentiate sub-populations of cells at single-cell resolution. With the combination of newly developed chemical derivitization, advanced statistical modeling, and optimized sample preparation process, we have shown that sub-populations of cells could be differentiated at the single-cell level with a typical throughput of >2,000 cells per day and >1,500 proteins quantified per sample.

401

Metastatic melanoma model system selection: implications for glyco-marker discovery

Jodie L Abrahams[1], Matthew P Campbell[1], Andrea Maggioni[1], Daniel Kolarich[1], Mark von Itzstein[1], Nicolle H Packer[2, 1]

1. Institute for Glycomics, Griffith University, Gold Coast, QLD, Australia

2. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

There is a critical need to identify novel predictive and prognostic markers with improved sensitivity and specificity for the management of advanced stage melanoma. Despite a growing understanding of the role of protein glycosylation in melanoma biology, there has been little progress towards using glycosylation changes as melanoma biomarkers.

To date most glycosylation studies have relied on the availability of cell lines. Although these models are an extremely valuable resource for cancer biology research we must question whether the glycosylation patterns of continuously cultured cell lines are truly representative of complex tumour tissue samples. Here we present a comprehensive comparison of commonly used cell lines, patient derived primary cells and lymph node tumour tissue from metastatic melanoma patients to help answer this question.

N-glycans released from melanoma cell membrane proteins were characterised using a porous graphitized carbon (PGC) liquid chromatography mass spectrometry glycomics platform. Structures were fully assigned using MS/MS fragmentation patterns, PGC retention behaviours, and linkages confirmed using exoglycosidase enzymes. The glycosylation profile of cell line panels, patient derived primary cell lines and tumour tissue were compared to identify sample specific glycan features.

Global glycomics identified metastatic melanoma associated glycan features common to all sample sources including branching, sialylation and fucosylation. However, significant differences in the relative abundance of glycan classes were observed between the sample origins, most notably the ratio of mannose to complex type glycans. We show that patient tissue exhibited a more complex profile than cultured cells and that there is a high degree of variation in protein *N*-glycosylation between patient tumours.

This study contributes to our understanding of glycosylation alterations in melanoma and highlights that cell surface glycosylation not only varies between different cell lines, stored cell lines and primary cell lines, but also markedly varies between all cell lines and tissue.

402

The Molecular Stethoscope: RNA of drug metabolising enzymes in circulating vesicles correlates with their specific protein content in the liver

Brahim Achour[1], Zubida M Al-Majdoub[1], Kristi Lea[2], Jeoffrey Schageman[1], Amin Rostami-Hodjegan[1, 3]

1. University of Manchester, Manchester, United Kingdom

2. Life Technologies, Thermo, Austin, Texas, United States of America

3. Certara UK Limited, Simcyp Division, Sheffield, United Kingdom

Background: Precision dosing aims to deliver the right drug dose for a specific patient based on individual characteristics, improving efficacy and reducing toxicity. Multi-omic approaches and 'liquid biopsy' assays are expected to facilitate the use of precision dosing by linking an organ, such as liver (site of drug metabolism) to a minimally-invasive biopsy, e.g. blood.

Methods: Liver tissue (20-250 mg) and plasma (1-3 mL) samples from the same cancer patients (n=9) were analysed using proteomic (targeted and global) and transcriptomic (NGS) methods. Healthy controls (n=5) were used as baseline. In-house QconCAT methodology was used for protein quantification on nanoHPLC-Orbitrap Elite system (Thermo) using DDA mode with an inclusion list. NGS followed Ampliseq workflow performed at Life Technologies (Thermo, Texas) at a depth of 8 million reads/sample. Expression levels of drug-metabolising enzymes were normalized by a novel liver-specific shedding correction factor (LSCF), computed using a combination of 14 liver-specific marker genes measured in plasma.

Results: Liver proteins (~2500) and plasma RNA transcripts (~21000) were measured in cancer samples and normal controls. Data for key drug-metabolising enzymes (the targets) and 14 marker genes were processed. Coverage of targets and markers was 80-100% in the protein data and 64-100% in NGS data. LSCF was higher and more variable in cancer patients than healthy controls; $LSCF_{(cancer)}$=21.62±16.30 rpm, n=9; $LSCF_{(healthy)}$=0.83±0.26 rpm, n=5; *t*-test p<0.01). Tissue protein and LSCF-corrected plasma RNA levels were assessed for correlation; major drug-metabolising enzymes were significantly correlated between plasma (RNA) and liver (protein); CYP3A4 (Pearson R=0.98, p<0.01, R^2=0.95); CYP2C9 (R=0.76, p=0.03, R^2=0.57); CYP1A2 (R=0.93, p=0.02, R^2=0.86); CYP2A6 (R=0.98, p<0.01, R^2=0.96)

Conclusion: A liquid biopsy test for non-invasive assessment of liver content of key enzymes in plasma was established; caveats: (a) enzymes exclusively/predominantly expressed in liver and (b) predominantly shed into the blood. This technology should facilitate efforts towards precision dosing/medicine.

403

System-wide characterization of esophageal squamous cell carcinoma cells by phosphoproteome analysis.

Jun Adachi[1], Akemi Kakudo[1], Ryohei Narumi[1], Yuichi Abe[1], Junko Isoyama[1], Takeshi Tomonaga[1]

1. National Institutes of Biomedical Innovation, Health and Nutrition, Ibaraki, OSAKA, Japan

Phosphoproteome is one of key signatures to understand the differences of patient-derived cancer cell lines at the molecular level, pathway level and system level. We developed sensitive and high-throughput phosphoproteome and tyrosine phosphoproteome analysis platform and performed characterization of 35 esophageal squamous cell carcinoma cell lines.

We obtained phosphoproteome and phosphotyrosine-proteome profiles of thirty-five esophageal squamous cell carcinoma cells established from Japanese patients. Phosphopeptides were enriched by immobilized metal affinity chromatography (IMAC) and

phosphotyrosine peptides were further purified by immuno precipitation using pY 1000 antibody. Peptides were labeled by TMT10 plex reagent and analyzed by Q Exactive plus instrument and MS data were processed by MaxQuant. Search results were filtered to a maximum false discovery rate (FDR) of 0.01 for proteins and peptides. Class I phosphorylation sites (localization probability, $p > 0.75$) were counted as identified sites/peptides.

We identified over 19000 phosphorylation sites including 1402 phosphorylation sites on tyrosine residue across thirty-five cell lines. Moreover 945 phosphorylation sites on protein kinases were quantified. Our preliminary results suggest that phosphorylation status is cell line specific and we observed subgroups in which EGFR, Met, Fyn, Trio or PTK2 are highly phosphorylated. We will present the results of cell viability assays using kinase inhibitors. Phosphoproteomic approach is applicable for characterizing the cellular kinase signaling status and classify subgroups of esophageal squamous cell carcinoma cells.

Phosphoproteome profiling is useful to characterize cellular signaling status and classify subgroups of esophageal squamous cell carcinoma cells.

404

Many multi-transmembrane domain-containing proteins can't conform to current high-stringency MS metrics solely after tryptic digestion

Subash Adhikari[1], Seong Beom Ahn[1], Mark S. Baker[1]

1. Macquarie University, Sydney, NSW, Australia

Background

Since 2010 through neXtProt, the Human Proteome Project (HPP) has progressively assigned protein existence 1 (PE1) status to human proteins according to community-agreed, high-stringency mass spectrometry (MS) metrics. However, the identification of human membrane proteins at this high-stringency remains problematic. Limited arginine (R) and lysine (K) residues in multi-transmembrane domains (TMDs) and restricted tryptic activity to this hydrophobic environment lead to the underrepresentation of TMD tryptic peptides in bottom-up MS experiments. These observations led us to examine the predicted experimental tryptic peptide yield from TMD-containing membrane proteins (TMD-MP), if trypsin were to be precluded from acting on TMD R/K residues. *In silico* analysis of the tryptic peptide repertoire corresponding to the complete sequence and N-/C-terminal strand + ecto- + endo- loop domains (i.e., non-TMD regions) of all TMD-MP highlights that a number of these proteins are unable to generate tryptic peptides that meet high-stringency HPP guidelines.

Method

UniProt defined ecto- and endo- domains including the N- and C- terminal strands (i.e., with only TMDs removed) of TMD-MPs were tryptically digested *in silico*. Proteins that potentially met HPP PE1 MS-based assignment criteria *(i.e., ability to generate 2+ MS-detectable 9+ amino acid (AA), uniquely-mapping, non-nested tryptic peptides)* from these "soluble" hydrophilic domains were determined. Further analyses were performed at lower stringency to visualize how stringency affects the number of proteins that qualify.

Result

In total, 204 of 3,878 (i.e., ~5%) TMD-MPs in the human proteome could not generate peptides as per current HPP MS guidelines, gradually decreasing to 103 (~2.5%) if stringency criteria were relaxed to a single uniquely mapping peptide of 7+ AAs. Olfactory receptors were (by far) the largest protein family that contributed to these statistics. In total, we also observed that 54 (~1.4 %) of all TMD-MPs failed to generate tryptic peptides upon digestion of their complete sequence.

405

Metabolomics Spectral Libraries for data-independent SWATH Liquid Chromatography-Mass Spectrometry Acquisition "SASA, a Novel SWATH Based Metabolomics tool"

Eman Ahmed[2, 1], Hassan Shikshaky[2], Shahd Ezzeldin[2], Aya Osama[2], Sameh Magdeldin[2]

1. Pharmacology Department , Faculty of Veterinary Medicine, Suez Canal University, Ismailia, 41522, Egypt, Ismailia., Egypt

2. Children Cancer Hospital in Egypt 57357-CCHE. , Children Cancer Hospital in Egypt 57357-CCHE, Cairo, CAIRO, Egypt

Metabolomics is considered a potential technique for clinical diagnosis, molecular medicine, therapeutic drug monitoring, and drug development. The conventional metabolomics analysis pipeline depends on IDA technique. Although it is a powerful technique, it is still suffering from stochastic, not reproducible ion selection across samples. Furthermore, even though the presence of different workbenches for metabolomics, metabolites identification remains a tedious task and time-consuming. Consequently, SWATH acquisition has attracted much attention to overcome this limitation. Therefore, a novel SWATH platform for data analysis had been developed with a generation of an accurate mass spectral library for metabolite identification using SWATH mass spectrometry acquisition that relies on the alignment of transition ions. The workflow was validated using standards inclusion/ exclusion compounds list. The false-positive identification was 3.4% from the non-endogenous drugs and the false negative was 10% of the standards with 90% sensitivity and 96.6% specificity. Ions with height ratio samples to blank ≥ 5, ABS retention time shift for each fragment < 0.1417%, ABS peak width at half height for each fragment in relation to precursor ion <17.4965 were kept. The workflow has the availability to subtract the background noise although the complexity of the SWATH sample. Besides, the reality of the identified transition. To demonstrate the feasibility of the workflow strategy, 1282 compounds from HMDB were tested in a variety of biological samples. In the current study, 377 compounds in positive mode and 303 in negative one with 392 unique non-redundant metabolites were recorded. After workflow validation, a free software tool, termed SASA, was developed to analyze SWATH acquired samples.

1. 1. Bonner, R. and G. Hopfgartner, SWATH data independent acquisition mass spectrometry for metabolomics. TrAC Trends in Analytical Chemistry, 2018. 2. Kind, T., et al., Identification of small molecules using accurate mass MS/MS search. Mass Spectrom Rev, 2018. 37(4): p. 513-532. 3. Bruderer, T., et al., Metabolomic spectral libraries for data-

independent SWATH liquid chromatography mass spectrometry acquisition. 2018. 410(7): p. 1873-1884. 4. Anjo, S.I., C. Santa, and B. Manadas, SWATH-MS as a tool for biomarker discovery: From basic research to clinical applications. Proteomics, 2017. 17(3-4).

406

Autoantibody response against tumor-associated antigens in gallbladder carcinoma using immunoproteomics approach

Javed Akhtar[1, 2], Ratna Priya[1, 2], Vaishali Jain[1], Puja Sakhuja[3], Anil Kumar Agarwal[3], Ravindra Varma Polisetty[4], Ravi Sirdeshmukh[5], Sudeshna Kar[2], Poonam Gautam[1]

1. ICMR- National Institute of Pathology, New Delhi, Delhi, India

2. Institute of Molecular Medicine, Jamia Hamdard, New Delhi, Delhi, India

3. Govind Ballabh Pant Institute of Postgraduate Medical Education and Research, New Delhi, Delhi, India

4. Department of Biochemistry, Sri Venkateswara College, University of Delhi, New Delhi, Delhi, India

5. Institute of Bioinformatics, Bangalore, Karnataka, India

Background: Early diagnosis is important for timely treatment of gallbladder carcinoma (GBC) patients leading to the increased survival rate. Here, we have applied serological proteome analysis (SERPA), an immunoproteomics approach, for detection of tumor-associated antigens (TAAs) eliciting humoral response in early stage GBC patients.

Methodology: Immunodepleted tissue proteins from GBC patients (n=7) were resolved by two-dimensional gel electrophoresis (2-DE) followed by immunoblotting using pooled blood plasma from healthy volunteers (n=11) or gallstone (GS) cases (n=11) or early stages of GBC (n=5) or advanced stages of GBC cases (n=9). Image analysis was performed using PDQuest software to identify protein spots with significantly high or specific immunoreactivity in GBC cases. The corresponding protein spots were excised from the 2-D gel followed by in-gel trypsin digestion and mass spectrometric analysis (LC-MS/MS) for identification of proteins. Two of the identified proteins were verified for autoantibody levels in individual plasma samples (30 cases and 20 controls) by dot blot assay.

Findings: 2-D immunoblot analysis led to identification of 25 protein spots showing either significantly high or specific immunoreactivity in early and/or advanced stages of GBC. Mass spectrometric analysis led to identification of proteins from the immunoreactive spots, including annexin A1 (ANXA1), and heat shock protein 60 (HSP60), carbonic anhydrase isoform 1 and 2, aldolase B and cathepsin D. Evaluation of autoantibody levels in individual plasma samples against two of the recombinant proteins, ANXA1 (an immunomodulatory protein implicated in cancer) and HSP60 (chaperonin involved in regulating apoptosis in cancer) using Dot blot assay showed significantly higher levels of autoantibodies against HSP60 (unpaired t-test, p= 0.023) in early stage GBC cases.

Conclusions: The study suggest that the autoantibody levels against HSP60 may be potentially employed for detection of GBC patients at early stages, however, the autoantibody level needs to be validated in larger cohort of clinical samples.

407

Glycan profiling on extracellular vesicles surface using evanescent-field fluorescence-assisted lectin array for biomarker discovery

Kazunari Akiyoshi[1]

1. Kyoto University, Kyoto, KYOTO, Japan

Extracellular vesicles (EVs) are released from various cells and play an important role in cellular communications relating to various diseases [1]. EVs can be used as biomarkers for diagnosis, prognosis, and determining the cell state. Although proteomics or genomics of EVs have been extensively studied, little is known about details of surface glycans on EVs. We report that comprehensive glycan patterns on intact exosomes can be analysed using an evanescent field fluorescence-assisted (EFF) lectin array system [2]. This lectin array system is simple, sensitive, and real-time detection of surface glycan patterns on intact EVs without the destruction of EVs.

EVs were isolated from various kinds of mouse and human cells including cancer cells, undifferentiated and differentiated MSCs, and immune cells by differential ultracentrifugation. Cy3-labeled EVs and their originating cell membranes (CMs) were applied to a glass slide with 45 lectins and fluorescence intensities were detected using an evanescent-field fluorescence scanner without washing. Hierarchical clustering analysis and principal component analysis were performed to evaluate whether surface glycans on EVs have their cell specific patterns. The results indicated that glycan profiling of EVs can be used to classify cell types (normal or cancer) and they can be further divided into each type of cancer, MSC sources, and cell lineages. Sialic acids–coaded EV (sialic acids were highly enriched on the surface of EVs analyzed by the lectin array) specifically interacted with cells expressing sialic acid-binding immunoglobulin (Ig)-like lectins (siglecs), such as HeLa cell *in vitro* and antigen-presenting cells *in vivo*[2].

In conclusion, EFF-lectin array method is a powerful tool for comprehensively glycan analysis of EVs towards functional analysis of glycan on EV surface and also biomarker discovery.

[1] N. Seo, *et al*. *Nature Communications*, 9, 435(2018), [2] A. Shimoda, *et al*. *Biochem. Biophys. Res. Commun.,* 491, 701(2017)

408

Antibodypedia reveals trends in antibody validation

Cristina Al-Khalili Szigyarto[1, 2], Kalle von Feilitzen[2], Mathias Uhlen[2]

1. KTH- Royal Institute of Technology, Stockholm, STOCKHOLM, Sweden

2. Systems biology, SciLifeLab, Solna, Stockholm, Sweden

The importance of antibodies for proteomics studies is widely acknowledged as well as the challenges in using them. Antibodies are versatile reagents used in different applications, under conditions that can influence their performance and their specificity toward the target. Consequently, antibody recognition and binding to the epitope, in a specific context eg. in Western Blot, is not necessarily translatable to other applications like ELISA. Since the performance of these reagents is application dependent, thorough testing is a prerequisite to ensure efficient use of antibodies as reagents. In addition efforts are needed to improve and share information regarding antibody specificity.

Antibodypedia is an online tool that enables scientist to find the right antibody for the right application. Currently the database comprises validation information and citations for more than 4.0 millions of antibodies. Functionality has been integrated to structure experimental data enabling analysis of antibody performance. This resource enables scientists to search among more than 1.9 millions of experimental results and select the most appropriate antibody for their specific application. The evidence regarding antibody performance is also accessible for side-by-side comparison of results and citations.

In accordance to the guideline for systematic validation of antibodies designed by the International Working Group for Antibody Validation (IWGAV) (Uhlén et al 2016), we have developed tools to highlight antibodies validated by: genetic strategies, orthogonal strategies, independent antibody strategies, expression of tagged proteins, and immunocapture followed by mass spectrometry (MS). An increasing number of antibodies are validated according to the IWGAV guideline eg. 4,392 antibodies listed in Antibodypedia are orthogonally validated. The increasing information collected in the database can be used to analyze availability of antibodies against the specific human proteins for specific applications and highlight proteins with few validated antibodies as well as applications lacking suitable reagents.

1. Uhlen et al 2016, Nature Methods 13:823–827.

409

SynGp one-stop shop: synthetic glycopeptides assisted glycoproteomics

Kathirvel Alagesan[1], Chi-Hung Lin[1], Arun Everest-Dass[1], Mark von Itzstein[1], Daniel Kolarich [1]

1. Griffith University, Gold Coast, QUEENSLAND, Australia

Intact glycopeptide characterization is an imperative yet challenging component of glycoprotein analysis. Elucidation of both glycan and peptide requires specific sample preparation workflows that, in combination with multiple tandem mass spectrometry (MS/MS) approaches, enable identification of single glycopeptide species. Here, we systematically optimized and evaluated the strengths and weaknesses of the glycoproteomics workflows using synthetic glycopeptides from sample preparation to its implication on data analysis.

A library of >100 synthetic *N*-glycopeptides representing human serum glycoproteins was synthesized. These synthetic glycopeptides fostered the systematic investigation to explore the advantages and limitations of the glycoproteomics workflow and comment on their suitability for high-throughput glycoproteomic studies. All experiments were performed on a quadrupole-Orbitrap-linear ion trap Tribrid mass-spectrometer.

A simplified approach to purify and produce a panel of glycosylated amino acids carrying *N*-linked glycans with various structures was developed. These building blocks were used to synthesise a library containing >100 glycopeptides and their unglycosylated counterparts. Recently developed "Drop-HILIC" (hydrophilic interaction chromatography) enrichment was further improved to allow for efficient enrichment of both *N*- and *O*-glycopeptides by using DMSO in solubilization buffer. Next, glycopeptide fragmentation characteristics were evaluated using different fragmentation techniques (HCD, SCE-HCD, ETciD, and EthcD) to assess the merit of each method in terms of (a) peptide backbone sequence coverage, (b) glycan composition, (c) proportion of signal in different fragment ion types (e.g., oxonium ions, Y-type ions, and peptide backbone fragment ions, and (d) unambiguous identification of the glycosylation site. We also evaluated and compared various software tools (Byonic, SugarQb) in their ability to identify these glycopeptides reliably.

To our knowledge, this constitutes the first broad and systematic analysis of the LC-MS/MS properties of glycopeptides using synthetic glycopeptides, allowing optimized bottom-up glycoproteomics experiments. The availability of these glycopeptides and spectra will facilitate the development, and improvement of further experimental and computational strategies.

1. Alagesan K, Everest-Dass A, Kolarich D. (2018). 'Isomeric Separation and Characterisation of Glycoconjugates' - Adv Exp Med Biol., 1104: 77-99. Alagesan K, Kolarich D. (2019). 'Improved strategy for large scale isolation of sialylglycopeptide (SGP) from egg yolk powder' – MethodsX ., 9;6:773-778. Alagesan K, Khilji SK, Kolarich D. (2017). 'It is all about the solvent: on the importance of mobile phase for ZIC-HILIC glycopeptide enrichment' - Anal Bioanal Chem., 409(2):529-538.

410

Analysis of proteome and phosphoproteome in response to higher temperatures in rice cultivars

Dalal Alonazy[1], Ramesh Katam[1], Peter Scott[2], Sara Bergin[3, 4], Raja K Reddy[5]

1. Department of Biological Sciences, Florida A&M University, Tallahassee, Florida

2. University of Florida , Gainesville, FL, US

3. Biological Sciences , Florida A&M University , Tallahassee, Florida, US

4. Des Moines Area Community College, Ankeny, IA, USA

5. Mississippi State University, MS, Mississippi State, USA

Rice is the third most important agricultural commodity worldwide. Annual increase in the global temperature has an adverse effect on the crop productivity and nutritional quality. During the temperature stress, plants undergo various physiological, biochemical changes occur in the plant to adapt to high temperatures levels. The present study was aimed to investigate the

relative changes in proteome and phosphoproteome in response to higher temperatures in rice cultivars. The specific objectives are to: (i) characterize of heat stress responsive proteins in tolerant and susceptible rice cultivars (ii) determine phosphorylated protein modifications in response to heat stress (iii) evaluate the proteins associated with heat tolerance mechanism and their biosynthetic pathways. Heat tolerant (Cv-5) and susceptible (Cv-13) rice cultivars were used to investigate the changes in proteome changes to heat stress in rice panicle. Plants were treated with three temperature levels; i) Optimum control 28/20°C, day/night, ii) Moderate32/24°C, and iii) high 36/28°C. Plant growth and yield-related parameters were determined after 137 days sowing till harvest. Proteins were extracted from frozen panicle tissue and iTRAQ labeled, following hybrid quadrupole-TOF QSTAR Elite MS/MS system and TripleTOF™ 5600. A total of 779 proteins were identified in control and treated samples. Twenty-seven proteins found to be phosphoproteins. When compared to the proteins responded to temperature, Cv-5 showed the abundance of highest number of proteins in response to higher temperature treatments. Gene ontology enrichment analysis with agriGO indicated the majority of the identified functions include response to abiotic stimulus, metabolites. Cv-5 displayed highest amount of proteins that are associated with defense responses and environmental stresses. The proteome data was validated with the expression levels of transcripts. This study provides new insight on proteins and phosphoproteins contribute to the temperature tolerance.

411

A different perspective of circulatory biomarker discovery in neurodegenerative diseases: combined use of alternative proteomics analyses to create a comprehensive "proteomics signature of the blood"

Sandra I. Anjo[1, 2], Rafael Silva[1], Miguel Rosado[1, 3], Inês Baldeiras[1, 2, 4], Alexandre Mendonça[5], Cristina Januário[4], Isabel Santana[1, 2, 4], Ana Verdelho[5, 6], Diana Santos[7], Andreia Gomes[7], Miguel Castelo-Branco[7], Bruno Manadas[1]

1. Center for Neuroscience and Cell Biology – University of Coimbra (CNC-UC), Coimbra, PORTUGAL, Portugal

2. Faculty of Medicine, University of Coimbra, Coimbra, Portugal

3. Faculty of Pharmacy, University of Coimbra, Coimbra, Portugal

4. Hospital and University Center of Coimbra, Coimbra, Portugal

5. Faculty of Medicine, University of Lisbon, Lisbon, Portugal

6. Hospital de Santa Maria, Lisbon, Portugal

7. Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT), Coimbra, Portugal

Background: Blood biomarker discovery has been dominated by targeted analysis of disease-associated proteins or conventional untargeted proteomics strategies. However, these attempts have failed to identify high confidence biomarkers of Neurodegenerative Diseases (NDs), which we believe is in part due to: i) the use of conventional proteomics analyses in a very challenging sample, such as blood, and ii) the difficulties in classifying the patients' groups.

Moreover, most of these studies are focused on a single fraction of the blood, such as plasma or serum, which only reveals an incomplete set of circulating proteins. Thus, the present work aimed at obtaining a comprehensive characterization of the circulatory proteins by combining different proteomics analyses of blood samples from the same individuals, including: i) undepleted plasma; ii) peripheral blood mononuclear cells (PBMCs);and iii) high and low molecular weight fractions of the plasma/serum, usually misrepresented in a conventional analysis.

Methodologies: To achieve this goal, state-of-the-art quantitative proteomics, SWATH-MS, was used in the three analyses referred above. These analyses were performed for a representative group of the two most common NDs and a control group. The diagnostic model achieved was further tested in some potential cases of Alzheimer's and Parkinson's disease.

Findings: From this combinatory study it was possible to quantify thousands of proteins, with some of these being capable of distinguishing the NDs from the control samples, and AD from PD patients.

Conclusions: The results achieved in this preliminary study strengthened the importance of combining alternative proteomics strategies to obtain a deeper characterization of patient's samples, creating a "blood signature" of each sample, which has a higher potential as a source of biomarkers. Future work will comprise the introduction of more disease groups (samples are already being collected), and the addition of other analysis such as PTMs-focused and metabolomics analysis of these samples.

412

Automated quality assessment for genome-wide protein structure prediction

Brendan R. E. Ansell[1], Bernard J Pope[2, 3], Peter Georgeson[2, 3], Samantha J Emery-Corbin[1], Melanie Bahlo[1], Aaron R Jex[1]

1. Walter & Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

2. Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, VIC, Austrlia

3. Melbourne Bioinformatics, Parkville, VIC, Australia

Computational protein structure prediction is increasing in feasibility and accuracy, as demonstrated by recent advances from alphaFold and recurrent geometric networks (AlQuraishi et al 2019). Protein structure prediction will soon likely form a standard part of genome reannotation pipelines for both model and divergent organisms. Prediction of human protein structures can greatly assist in understanding the roles of different isoforms, and nonsynonymous genetic mutations across tissues, and in disease.

However in order to harness the potential of large-scale protein structure prediction, methods are needed to independently and efficiently assess the quality of thousands of predicted structures, without human assistance.

To test a novel quality assessment method we used I-TASSER software to predict the structure of 5000 proteins encoded in the human parasite *Giardia*, a simple eukaryote that lacks introns and thus encodes thousands of full-length proteins. We used discrete protein sequence annotations (Pfam codes) assigned to peptides encoding predicted structures, and their closest

empirically-determined homologues in the PDB, to bin the predicted structures into a high-confidence (matching IPR code) category, or lower-confidence category (non-matching).

Continuous metrics output by I-TASSER were used to construct a random forest machine learning model that predicted the high-confidence category, yielding structural insight into ~1000 proteins including enzymes important for drug resistance and redox maintenance. The classifier also produced a second tier of predicted structures that have features of the high-confidence structures, but lack matching PFAM domains with their closest crystal structure homologue.

High-confidence models exhibited greater transcriptional abundance, and the classifier generalized to selected human proteins, indicating the broad utility of this approach for automatically stratifying predicted structures.

This work provides a method for assigning confidence in predicted protein structures *en masse* in a software-agnostic manner, and can be used to vastly increase knowledge of the structural proteome in humans and other medically relevant species.

413

Integrated phospho-glycomics identified specific target networks of cancer stem cells

Norie Araki[1], Akiko N Nambu[1], Atit Silsirivanit[1, 2], Yuki N Ikeda[1], Daiki Kobayashi[1]

1. Tumor Genetics & Biology, Kumamoto University, Kumamoto, Japan

2. Biochemistry, Khon Kaen University, Khon Kaen, Thiland

Background

Cancer stem cells (CSC) have been considered responsible for the therapeutic-resistance/recurrence of cancers, and thus proposed as the therapeutic target. To clarify molecular mechanisms and develop clinical targets against CSC, we established CSC-clones from glioblastoma patient's tissues as glioma stem cells(GSCs), having a potential of differentiation/promotion of glioblastoma, and subjected to a unique integrated Phospho-glycomics. Using original GSC-iPEACH database, we tried to identify specific phospho-glyco networks associated to the maintenance/differentiation of GSC.

Methods

We established 12 GSC clones from patient's gliomas and glioma cell lines having the potential to differentiate into glioblastomas, extracted their Proteins and mRNAs, and subjected to iTRAQ/TMT, DNA array, Rectin/glyco-gene array, and quantitative phospho-proteome (HAMMOC) analyses. All of the data was integrated by iPEACH, and used for GO and knowledge-based network analyses. Biological validations were performed with immunocyto/histochemistry, western-blotting, mouse xenograft analyses.

Results

The data integration and extraction of specific phospho- and glyco-proteins, and their related genes revealed that, during the GSC differentiation, cell surface glyco-proteins and RAS-MAPK/PI3K signalings were significantly up-regulated, meanwhile, SOX2, CD133, and specific proteoglycans/synthetic-enzymes/metabolic pathways were obviously down-regulated. Interestingly, GSC differentiation was significantly associated with the decrease of chondroitin Sulfate(CS)-modified proteins/related enzymes. By the CS-degradation enzyme treatment, GSC differentiation was dramatically induced and this was associated with the up-regulation of specific cellular phosphokinase dependent signalings. Importantly, these differentiation processes were also associated with the interaction of CS-proteins and adhesion molecules such as integrins, and suppressed by integrin-inhibitors/CS administrations significantly. Combination treatments of a cancer-drug Temozolomide and these GSC-differentiation inhibitors suppressed glioma progression, increased the chemosensitivities, and led the longer survival of mouse xenograft-models.

Conclusions

Functional integrated phospho-glycomics for the first time demonstrates that the GSC induces the specific glyco-phospho network/signalings to regulate GSC stemness/differentiation, and these information will provide the new therapeutic strategies/clinical targets against malignant gliomas.

414

Novel sources of peptide antigens in the melanoma immunopeptidome

Ritchlynn Aranha[1], Anthony Purcell[1], Pouya Faridi[1], Nathan Croft[1], Divya Duscharla[1], Andreas Behren[2]

1. Monash University, Clayton, VIC, Australia

2. Cancer Immunobiology, Olivia Newton-John Cancer Research Institute, Austin Hospital, Heidelberg, Victoria, Australia

Background: Antigen-presenting cells play a crucial role in directing T-cells to identify and eliminate cancer cells by presenting peptides on the cell surface in complex with Human Leukocyte Antigens (HLA). The peptides presented by HLA class I molecules originate from proteins that are degraded by the proteasome into linear and spliced (non-contiguous sequences derived from the same or different proteins) peptide antigens. Mapping the peptide ligands presented to the immune system – not only on cancerous tissue but also on cells within the tumour microenvironment – will greatly enhance the range and diversity of targets available for immuno-therapy and enable personalised cancer treatment.

Methods: This study combines peptide sequencing performed by mass spectrometry with PEAKS Studio software and novel bioinformatics algorithms to identify linear and spliced peptides presented by six patient-derived melanoma cells under normal and interferon-treated conditions.

Results: We created a database consisting of more than 30,000 unique peptides per cell line, including linear, cis- and trans-spliced peptides. Spliced peptides made up around 15-20% of each peptide dataset. Interferon treatment of cells dramatically remodelled the repertoire of peptides presented by all cell lines, with only ~40% overlap between peptides presented under constitutive and stimulated conditions. 3265 (~8% of the total dataset) presented by the HLA-A*02:01 allele were common to all

cell lines. This database also includes nearly 1300 peptides derived from 79 known melanoma-associated antigens (MAA) and across all cell lines, about 50% of peptides derived from MAA were spliced in nature.

Conclusion: This study created the largest known database of melanoma patient-derived immunopeptidomes. We observed that interferon treatment leads to a change in the peptide repertoire presented by HLA-I and that peptide splicing further diversifies the immunopeptidome. A subset of HLA-A*02:01 binders were common to all cell lines, offering potential as viable vaccine targets across populations.

415

Proteomics on the way to reverse Avogadro number as basis of precision medicine

Alexander Archakov[1]

1. Institute of Biomedical Science, 10, 8, Russia

The main advances major challenges of the C-HPP – insufficient analytical sensitivity of proteomic technologies and the complexity of the proteome. The increased analytical sensitivity of proteomic technologies and the combined results from transcriptomic and proteomic analyses of a single chromosome of individual sample has facilitated the study of proteoforms. For example, of the 275 protein-coding genes on human chromosome 18, 85% of the transcripts were discovered using next-generation sequencing (NGS) and polymerase chain reaction (PCR), and 45% of the proteins were detected using shotgun and selected reaction monitoring technology, in both liver tissue and the HepG2 cell line. To increase proteome coverage, a combination of shotgun technology and selected reaction monitoring with two-dimensional alkaline fractionation has been recently developed by using UPS 1 and 2. Sigma Aldrige sets as the “gold standard”. To detect proteoforms that cannot be identified by such technologies, nanotechnologies such as combined atomic force microscopy with molecular fishing and/or nanowire detection were used. Both technologies provide a powerful tool for single molecule analysis, by analogy with nanopore sequencing during genome analysis. We believe that such approach could be used for detection of the «missing» proteins of the single chromosome during C-HPP. According our point of view the technology sensitivity drives the depth and width of proteome.

416

Plasma proteomic profiling of patients with haemorrhagic fever with renal syndrome from puumala and dobrava hantavirus infections

Stuart D Armstrong[1], Miša Korva[2], Julian A Hiscox[1], Tatjana Avšič Županc [2]

1. University of Liverpool, Liverpool, United Kingdom

2. Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

Triaging patients with acute febrile illness, particularly with haemorrhagic fever, can provide better clinical focus and assessment of medical countermeasures. Identifying unique molecular signatures associated with a particular infection and each stage of subsequent disease – such as acute phase through to decline or recovery, would be of predictive value. For most infectious diseases a blood sample is easiest to obtain and usually the most informative. In most cases translating data to an easily applicable ELISA is most clinically relevant. Therefore, proteomics provides an ideal avenue to characterise real world samples. In order to evaluate whether this type of approach – from samples to predictive capability, patients with hantavirus were characterised using a proteomic approach.

Haemorrhagic fever with renal syndrome (HFRS) in Eurasia and hantavirus pulmonary syndrome in the Americas (HPS) are caused by hantaviruses. Clinical symptoms of HRFS can range from headaches and fever to thrombocytopenia and acute kidney injury. Two hantaviruses, Dobrava (DOBV) and Puumala (PUUV) cause HFRS disease in humans. PUUV infection usually has a mild clinical path with a case fatality rate < 1%. DOBV has a case fatality rate of approximately 10%. The pathogenesis behind the variance in severity of infection between individuals infected with different hantavirus genotypes is poorly understood.

Longitudinal plasma samples from 12 individuals with either mild or severe DOBV or PUUM were compared with healthy controls (n=5) using a high-resolution label free proteomics approach. Changes in the plasma proteome pointed to underlying biological consequences. From these, several proteins, including SPINK1, APOF, HPR and CFHR1 were identified as potential prognostic biomarkers for severe disease. Comparing the hantavirus profiles found here to the plasma proteomes of individuals exposed to other pathogens (e.g. EBOV and spirochetes) will stratify biomarkers for hantavirus that in turn will enhance the diagnosis and prognosis of this emerging pathogen.

417

Performance evaluation of a the new Orbitrap Exploris 480 mass spectrometer

Tabiwang N. Arrey[1], Rosa Rakownikow Jersie-Christensen[1], Julia Kraegenbring[1], Kerstin Strupat[1], Markus Kellmann[1], Catharina Crone[1], Thomas Moehring[1], Alexander Harder[1]

1. Thermo Fisher Scientific, Bremen, Germany

Since its introduction, Orbitrap-based-MS have and are still playing a pivotal role in many different research areas such as proteomics, metabolomics, biopharma, etc. Each of these applications comes with different challenges to mass spectrometers. To address some of these challenges, new technological developments, as well as improvements on existing mass spectrometers is a necessity. Here we evaluated the Orbitrap Exploris 480 mass spectrometer for proteomics applications, with focus on data dependent acquisition and data independent acquisition. Additionally, we assess the use of the PHiSDM processing algorithm on TMT11plex labeled samples.

The initial results from the comparison of DDA methods between the Q Exactive HF-X and the Orbitrap Exploris 480 MS gives a yield of approximately 10% more peptide and protein identifications when using the same sample and experimental conditions. For the DIA experiments, we were able to reduce the sample amount by a factor of 2 and still achieve the same results as on the

Q Exactive HF-X. Furthermore, to demonstrate the qualitative capability of the quadrupole Orbitrap, a two-proteome mixture was analyzed. We determined the precision and accuracy for LFQ using yeast spiked into a constant HeLa background on the Orbitrap Exploris 480 MS and found the highest deviation in accuracy to be only around 10% with high precision. On the Orbitrap Exploris 480 MS a dedicated algorithm, phase-constrained spectrum deconvolution method(SDM), has been implemented to reach higher resolution in shorter times. This is especially beneficial when using reporter ion quantitation such as TMT11plex, as it enables the use of shorter transients to achieve the same mass resolution relative to conventional FT based approaches without sacrificing data quality. With the PHiSDM algorithm activated we were able to boost the total number of quantified proteins in a HeLa digest labeled with TMT 11plex, by 25 %.

418

Unrevealed mystery of cell dust:extracellular vesicles and tumor-derived exosomes

Deanna Ayupova[1], Renee Goreham[2], Paul Teesdale-Spittle[1]

1. Victoria University of Wellington, Wellington, WELLINGTON, New Zealand

2. Univeristy of Newcastle, Newcastle, Australia

Lung cancer is the leading cause of cancer death. Early diagnosis of cancer is highly desirable, as treatment is more successful with early stage cancers. However, early detection of lung cancer remains a challenge, with most people diagnosed only after the cancer has progressed. Since their discovery as a subset of extracellular vesicles (EVs), exosomes have attracted interest, as they provide potential biomarkers for various diseases including cancer.

The potential of EVs, and especially exosomes, as sources of biomarkers is very high as has been identified through several comparative proteomic analyses of exosomes from different biological fluids (plasma, serum, urine, saliva). Despite the fact that exosomes can serve as supporting evidence in acancer diagnosis, there is no defined and standardized method for isolating and purifying exosomes that gives substantial yield, purity, and optimised protein abundance for proteomic analysis. In this study, we used ultracentrifugation and size-exclusion column-based isolation for exosome isolation from conditioned cell culture medium and exhaled breath condensate. Exosomes characterization was confirmed by transmission electron microscopy (TEM), scanning electron cryomicroscopy (cryoSEM), dynamic light scattering (DLS), immunoblotting, and protein quantification assay. A preliminary proteomic investigation is underway. This approach is the most effective method to isolate exosomes as it keeps vesicles intact. Application of Immunoblotting with gold nanoparticles (AuNP) proved the identification of general (non-specific) markers for exosomes (tetraspanines CD63 and 81) on the membrane for all isolated exosomes. Finally, we propose that exosomes can be isolated in a non-invasive way from breath, and that these exosomes will contain protein biomarkers that can be used in early detection of lung cancer.

419

SWATH-MS proteomic analysis can discriminate between actinic keratosis, Bowen's disease and cutaneous squamous cell carcinoma

Ali Azimi[2, 1], Pengyi Yang[3], Marina Ali[2, 1], Vicki Howard[4], Graham Mann[5, 6], Kimberley Kaufman[7, 8], Pablo Fernandez-Penas[2, 1]

1. Centre for Translational Skin Research, The University of Sydney, Westmead, NSW, Australia

2. The Department of Dermatology, Westmead Hospital, The University of Sydney, Westmead, NSW 2145, Australia

3. School Mathematics and Statistics, The University of Sydney, Camperdown, NSW, Australia

4. Dermatopathology, Skin and Cancer Foundation Australia, Darlinghurst, NSW, Australia

5. Centre for Cancer Research, Westmead Institute for Medical Research, The University of Sydney, Westmead, NSW, Australia

6. Melanoma Inistitute Australia, The University of Sydney, Wollstonecraft, NSW, Australia

7. School of Molecular Bioscience, Faculty of Science, The University of Sydney, Camperdown , NSW 2050, Australia

8. Brain and Mind Centre, The University of Sydney, Camperdown, NSW 2050, Australia

Actinic keratosis (AK), Bowen's disease (BD) and cutaneous squamous cell carcinoma (cSCC) are heterogeneous keratinocytic skin lesions (KSL). Biomarkers that can accurately stratify these lesion types are needed to support a new paradigm of personalised, precise management of skin neoplasia. In this paper, we used the data independent acquisition (DIA) proteomics workflow, SWATH-MS, to analyse formalin-fixed paraffin embedded (FFPE) samples of normal skin and KSL including well differentiated (WD), moderately differentiated (MD) and poorly differentiated (PD) cSCC. We quantified 3574 proteins across 93 samples studied. Differential abundance analysis identified 19, five and six protein markers exclusive to AK, BD and cSCC lesions, respectively. Among cSCC lesions of various levels of tumour differentiation, 118, 230 and 17 proteins showed potential as biomarkers of WD-, MD- and PD-cSCC lesions, respectively. Bioinformatics analysis revealed that AK and cSCC lesions were associated with decreased apoptosis, and BD lesions with over-representation of DNA damage repair pathway. Differential expression of FGFR2 alternative splicing, Rho GTPase signalling, and RNA metabolism proteins were associated with the level of cSCC tumour differentiation. Proteome profiles also separated KSL subtypes on principal components analysis. Overall, protein markers have excellent potential to discriminate KSL subtypes and facilitate new diagnostic and therapeutic strategies.

420

Physiological and proteomic changes to heat and water stress in soybean leaf

Ramesh Katam[1], Mahya Bahmani[1], Kambham Raja Reddy[2]

1. Department of Biological Sciences, Florida A&M University, Tallahassee, FL, USA

2. Department of Plant and Soil Science, Mississippi State University, Mississippi State, MS, USA

Production of Soybean (*Glycine max*), an important oil seed crop is severely influenced by several abiotic factors, such as flooding, and drought. Water stress is one of the ruminating factors of soybean crop. Plants adapt to water and temperature stresses via modulating signal molecules associating the protein changes thus altering several metabolic pathways. To obtain insight into the effects of WS and HS on molecular and cellular functions of soybean, changes in leaf protein composition were studied by 2-DE Gel electrophoresis complemented with Mass spectrometry. Two soybean cultivars, drought tolerant (A) and susceptible (B) were exposed to different heat and water conditions. PD Quest analysis revealed at least 200 proteins in both cultivars, of which, 61 proteins were differentially expressed in response to WS and HS. Differentially expressed leaf proteins were excised from 2-DE gel, and trypsin digested. The peptide sequence tags generated from the spots were queried through MASCOT search. Gene ontology analysis for each protein reveals functional categories including photosynthesis, metabolism, transport, stress and defense, and glycolysis.

The majorities of heat responsive-proteins were up regulated during heat stress and combined stress in cultivar B; these proteins were down regulated to water stress. However in cultivar A, the heat shock proteins were generally down regulated to all levels stress. Proteins involved in folding and biosynthesis were either over expressed or did not change due to heat stress. Our studies showed that differentially expressed proteins involved in antioxidant defense were mostly up-regulated, whereas proteins associated with photosynthesis, secondary metabolism, and amino acid and protein biosynthesis were down-regulated in response to heat stress. In both cultivars, proteins involved in folding and biosynthesis were either highly abundant or remains unchanged to heat stress. Furthermore, a pattern of cross-tolerance phenomenon was detected in both cultivars for two different stresses when subjected independently. Proteins involved in antioxidant defense were most relatively abundant, whereas proteins associated with photosynthesis, secondary metabolism, and amino acid and protein biosynthesis were detected in low abundance to heat stress.

421

SWATH library construction including recombinant proteins allows identification and quantification of lower abundance human plasma cancer biomarkers

Seong Beom Ahn[1], Karthik Kamath[2], Abidali Mohamedali[3], Zainab Noor[3], Dana Pascovici[2], Jemma Wu[2], Subash Adhikari[1], Matthew J. McKay[2], Mark S. Baker[1]

1. Department of Biomedical Sciences, Macquarie University, North Ryde, NSW, Australia

2. Australian Proteome Analysis Facility, Macquarie University, North Ryde, NSW, Australia

3. Department of Molecular Sciences , Macquarie University, North Ryde, NSW, Australia

Human plasma is the most informative, accessible biofluid for assessing the status of human health. However, detection and quantification of low abundance cancer-related proteins (e.g., CEA, cytokines, shed proteins) remains one of the principal challenges in proteomic biomarker discovery - due to high abundance plasma proteins obscuring biomarkers. Antibody technologies suffer from batch variation and non-specific detection, therefore, quantitative MS techniques such as SWATH-MS and similar DIA methods are an attractive alternative to assess plasma cancer biomarkers. These approaches rely on prior library construction using IDA/DDA data and the challenge of detecting low abundance biomarkers can be hampered by the same dynamic range issues which limit the proteome discovery space for plasma proteomics.

Here, we report on the use of a SWATH mini-library comprised of 32 previously-reported cancer biomarkers for the quantitative assessment of non-depleted pooled human plasma cohorts from clinically-staged (20 healthy or 20 stages I-IV) CRC patients by SWATH-MS. To ensure validity, we employed two independent SWATH analysis software (Skyline and PeakView) to identify quantifiable peptides. Of the 32 cancer biomarkers used to construct the SWATH mini-library, we reliably identified and quantified 25 proteins in human plasma of CRC patients. In all cases, a significantly higher peptide count for each protein allowed better quantification (e.g., 12 for CEA, 8 for IL-6) compared to prior results. In CRC, CEA expression was significantly upregulated, recapitulating many prior studies. Similarly, we recapitulated our observation that plasma ADAMDEC1 is upregulated in early-stage CRC. We propose the use of expanded recombinant protein SWATH libraries for the discovery of diagnostic, prognostic and theranostic protein biosignatures for cancers and other diseases.

422

BloodKB: an open community-scale knowledge base for plasma proteome diversity

Julie Wertz[1], Benjamin Pullman[1], Jeremy Carver[1], Seungjin Na[2], Nuno Bandeira[1]

1. University of California, San Diego, La Jolla, CA, United States

2. Computer Science, Hanyang University, Seoul, South Korea

Blood, plasma, and serum are key for biomarker discovery as they're the predominant types of samples used for diagnostic analyses in clinical practice and are available in biobanks from thousands of clinical studies. Nevertheless, MS/MS analysis of these types of samples typically identifies less than one third of all acquired spectra and little is known about the diversity and quantitative variation of post-translational modifications across samples from different individuals or over time series from the same individuals. BloodKB aggregates identifications from deep reanalysis of >35 million spectra in >1TB of public mass spectrometry data deposited in MassIVE, containing >1,300 samples from dozens of individuals and covering a broad range of variations in age and gender, health and disease, and time series in response to health interventions.

Altogether, our reanalysis doubled the identification rate for spectra in these datasets and Maestro open search for unexpected post-translational modifications revealed an unprecedented level of diversity, with hundreds of known modification types

significantly detected at minimum false localization rate of 1% and supported by very significant correlations between the spectra of modified and unmodified peptides (i.e., p-values <1e-10). A further 120+ putative novel modifications were also detected under the same stringent conditions even after consideration of possible combinations of all known modifications.

Reanalysis further reveals >800 hypermodified peptides with 10+ distinct combinations of modifications, with up to over 200 unique modification variants for a single peptide sequence. Peptide identifications were also uniquely mapped to ~20,000 exons (out of a total of ~29,000 exons), with ~7,000 distinct peptides mapped to exon splice junctions and hundreds more mapped to functional regions associated with disease or regulatory interactions.

As an open community resource, BloodKB will continue to grow in volume and knowledge as more data becomes publicly available and additional analyses or metadata are contributed over time.

423

Analytical validation of run-to-run and site-to-site performance of a human immune profiling assay and automated data analysis solution for CyTOF mass cytometry technology

Stephen Li[1], Daniel Majonis[2], Vladimer Baranov[1], Olga Ornatsky[1], C. Bruce Bagwell[3], Benjamin C. Hunsberger[3], Dmitry Bandura[1]

1. Fluidigm Canada Inc., Markham, Ontario, Canada

2. Fluidigm Corporation, South San Francisco, CA, United States

3. Verity Software House, Topsham, Maine, USA

Immune profiling is an essential method for quantifying changes in immune population numbers and states over time in health and disease. A cornerstone in translational and clinical research, it is frequently used to investigate chronic inflammation, infectious disease, autoimmune diseases, and cancer. The diversity of immune populations demands a high-parameter approach to fully and efficiently quantify changes. Mass cytometry, which utilizes CyTOF® technology, is a single-cell analysis platform that has used as many as 50 metal-tagged antibodies[1] to resolve discrete cell populations, all in a single tube of sample. It is an ideal solution for routine enumeration of immune cell populations.

We developed a sample-to-answer solution for human immune profiling using mass cytometry: the Maxpar® Direct™ Immune Profiling System. It includes an optimized 30-marker immune profiling panel provided in a single-tube format, validated SOPs for human whole blood and PBMC staining, an instrument data acquisition template, instructions for data acquisition on a Helios™ system, and automated Maxpar Pathsetter™ software for data analysis.

Here we present analytical validation data on repeatability, reproducibility, software precision / accuracy, and site-to-site reproducibility. The repeatability of eight identical donor samples acquired on a single Helios instrument resulted in CVs <12% for whole blood and <9% for PBMC for all major populations (>5% in frequency). Reproducibility of three identical samples acquired on three different Helios instruments resulted in CVs <10% for whole blood and <20% for PBMC for all major populations. We also performed Deming regression to test differences in results obtained by manual gating and Maxpar Pathsetter. For all major populations analyzed, 95% confidence interval of the correlation coefficient contained the value of 1, suggesting no proportional difference between the two methods of analysis. Finally, a multi-site study of whole blood from a single donor stained and analyzed at five institutions independently resulted in CVs of <10% for populations ≥5% in frequency, demonstrating that the Maxpar Direct Immune Profiling System shows a high degree of inter-site reproducibility.

We conclude that this assay provides a robust solution for broad immune profiling using mass cytometry, reducing sources of variability and subjectivity in sample preparation and data analysis.

1. 1 Simoni, Y., Becht. E., Fehlings, M. et al. “Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates.” Nature 557 (2019): 575–579

424

DeepY: A deep learning model for biosimilarity evaluation of antibody drug using intact glycoproteins analysis by LC-MS

Geul Bang[1], Ji won Lee[1], Myung Jin Oh[2], Hyun Joo An[2], Heeyoun Hwang[1]

1. Korea Basic Science Institute, Cheongju, CHUNGBUK, South Korea

2. Chungnam national university, Daejeon, South Korea

Heterogeneity and complexity of the glycosylation on biotherapeutics greatly depend on expression system, process conditions, and environment of cell culture of products. In order to evaluate biosimilarity of antibody drug such as Herceptin®, we have developed a deep learning model, DeepY, using intact glycoprotein analysis by LC-MS. Briefly, each antibody drug was independently analyzed to identify its intact glycoprotein composition. As a result, the list of identified intact glycoprotein compositions from each MS data was merged in the intact glycoprotein database, where a total of 34 intact glycoprotein compositions was identified from all three antibody drugs. Independently, the deconvoluted masses and their abundances of antibody drugs generated by MaxEnt were used as data sets such as training, validation, and test set, for development of a deep learning model using convolutional and fully connected neural network. The accuracy was approximately 90% at training, validation and test set. The DeepY could predict the biosimilarity and distinguish the low quality of mass spectra from all antibody drugs. We will further test the antibody drugs in batch to batch, expand to other original antibody drugs and their biosimilars.

425

Modelling life and death in mammalian cells to generate rational engineering strategies

Craig Barry[1], Esteban Marcellin[1]

1. Australian Institute for Bioengineering and Nanotechnology, Fairfield, QLD, Australia

Dysregulation of checkpoint mechanisms which govern mammalian cell cycle and programmed death results in important cell proliferation phenotypes, spanning research interests from cancer studies to industrial cell line development. For example, the circadian oscillator has evolved into an autonomous timekeeping mechanism which coordinates metabolic requirements in synchrony with the sun. Each cell is equipped with its own circadian clock to confer rhythm across an entire organism. This metabolic control requires a coordinated protein interactome.

A protein interactome principally represents the mammalian cell cycle, and is subject to intrinsic and extrinsic noises which affect the trajectory of cell fate. *In silico* mathematical models have demonstrated an ability of accurately representing the dynamics of cell cycle regulatory networks and experimentally observed phenotypes, yet computational models of this kind require parameter training using experimental data.

Utilising a high resolution Q-Exactive HF-X Orbitrap mass spectrometer, we can capture and quantify crucial regulators of cell proliferation during the cell cycle. The Q Exactive HF-X enables MS/MS acquisition above 40 Hz at 7500 resolution resulting in impressive proteome coverage, achieving 3000 proteins in a single 1 h injection in capillary mode. Having established reliable methods for protein detection and quantification we have constructed mammalian cells which express four cell cycle phase-dependant fluorescent markers (FUCCI4) for discrete fluorescence activated cell sorting (FACS) for subsequent intracellular protein detection and quantification. This will allow us to build proteomic distributions of important nodes within the cell cycle interactome for parameter training of a detailed mathematical model. Using the protein expression profiles of discrete cell cycle phases, we aim to identify the boundary of intrinsic noise which determines normal cell proliferation and dysregulated phenotypes, leading to cell death. Identifying these protein markers will facilitate rational engineering strategies in areas of targeted medicine and mammalian cell line development of biologics-producing strains.

426

Dehydration-responsive nuclear proteome and phosphoproteome profiling of a grain legume chickpea (*Cicer arietinum* L.)

Niranjan Chakraborty[1], Pragya Barua[1]

1. National Institute of Plant Genome Research, New Delhi, DELHI, India

Background

Human health is deeply rooted with the natural resources, which provide nutrition, and many current research emphasize in underpinning sustainable planetary health. With the socio-economic impacts of climatic drifts on the rise, the yield of nitrogen fixing legumes is being adversely affected by dehydration stress. Identifying the organelle-specific regulators for adapting to such environmental constraint would not only aid in understanding the molecular basis of stress-response, but developing fortified varieties (1). Nucleus (PM), designated as the cell's control centre, hosts genetic information and regulates gene expression, we therefore aimed at understanding the dehydration-induced alterations in the expression patterns of the proteins and phosphoproteins hosted by the nucleus.

Methodologies

Four-week-old seedlings of a grain legume, chickpea, were subjected to gradual dehydration (2) and nuclear proteins (NPs) were extracted from unstressed control as well as from 72 and 144 h stressed tissues. Phosphopeptides were enriched by titanium-dioxide treatment followed by detection and relative quantification.

Findings

We identified 4832 NPs and 478 phosphosites, corresponding to 299 unique nuclear phosphoproteins (NPPs) involved in multivariate cellular processes including protein modification and regulation of gene expression, among others (3). The identified proteins included several novel kinases, phosphatases and transcription factors, besides 660 uncharacterised proteins. Spliceosome complex and splicing related proteins were dominant among differentially regulated NPPs, indicating their dehydration-modulated regulation. Phospho-motif analysis revealed stress-induced enrichment of proline-directed serine phosphorylation. Association mapping of NPPs revealed predominance of differential phosphorylation of spliceosome and splicing associated proteins. Also, regulatory proteins of key processes viz., protein degradation, regulation of flowering time and circadian clock were observed to undergo dehydration-induced dephosphorylation.

Concluding statement

This inventory comprising several novel regulatory proteins and their precise sites of phosphorylation, would provide new insights into stress adaptation and enable directed genetic manipulations for developing climate-resilient crops.

1. 1. Gayen, D., Barua, P., Lande, N. V., Varshney, S., Sengupta, S., Chakraborty S, Chakraborty, N. (2019). Dehydration-responsive alterations in the chloroplast proteome and cell metabolomic profile of rice reveals key stress adaptation responses. Environ Exp Bot 160: 12-24.
2. 2. Gayen, D., Gayali, S., Barua, P., Lande, N. V., Varshney, S., Sengupta, S., ... & Chakraborty, N. (2019). Dehydration-induced proteomic landscape of mitochondria in chickpea reveals large-scale coordination of key biological processes. J Proteomics, 192: 267-279.
3. 3. Barua P, Lande NV, Subba P, Gayen D, Pinto S, Keshav TS, Chakraborty S, Chakraborty N. (2019) Dehydration-responsive nuclear proteome landscape of chickpea (Cicer arietinum L.) reveals phosphorylation-mediated regulation of stress response. Plant Cell Environ 42: 230-244.

427

Enhancing extracellular vesicle isolation of human plasma verified by high resolution lipidomics

Amani Batarseh[1], Alex Chen[2], Kim Ekroos[3], Susannah Hallal[4], Kim Kaufman[5], Michael Mariani[6]

1. BCAL Diagnostics, Eveleigh, NSW, Australia

2. LC-MS Application Specialist ANZ, Thermo Fisher Scientific, Scoresby, NSW, Australia

3. Founder & CEO at Lipidomics Consulting Ltd. , Esbo, Finland

4. Discipline of Pathology, Brain and Mind Centre, Sydney Medical School, University of Sydney, Camperdown, NSW, Australia

5. Head of Brain Cancer Research, Chris O'Brien Lifehouse, 1-Department of Neurosurgery, Chris O'Brien Lifehouse, 2-Discipline of Pathology, Brain and Mind Centre, Sydney Medical School, University of Sydney, University of Sydney, Camperdown, NSW, Australia

6. Sales & Applications Manager ANZ - LSMS , Thermo Fisher Scientific, North Ryde, NSW, Australia

Extracellular vesicles (EVs) are secreted from many cell types and play important roles in intercellular communication. EVs carry a range of biomolecules that reflect the identity and molecular state of their parental cell and are found in biological fluids. Omics studies have extensively focused on characterisation of the protein and nucleic acid cargo of EVs while lipids are less studied. EVs are increasingly being utilised in disease diagnosis as they are considered to carry valuable information about the disease state. Thus, novel disease biomarkers might be identified in EV lipidomes.

EVs were enriched from 1ml human plasma samples using ultracentrifugation, considered the gold standard approach for EV enrichment, and size exclusion chromatography(SEC) (Izon). Lipids extracted according to Matyash et al. (2008) were loaded on a C30 Acclaim column and analysed using targeted and untargeted lipidomics approaches using a Vanquish liquid chromatography (LC) system and Fusion orbitrap mass spectrometer (MS). LipidSearch software was used to annotate lipid species.

More than 250 lipid species were identified and quantified in the plasma EVs following both enrichment methods. The two methods generated highly similar lipid profiles, indicating that SEC may be a viable alternative to the cumbersome UC method. Interestingly, the SEC approach yielded less lysophosphatidylcholine lipids, which may be related to a more homogenous vesicle population captured by SEC. Various literature reviews refer to glycerolipids, likely originating from co-isolating vesicles such as low-density lipoproteins, as contaminants in the EV fractions. We detected these lipids and propose that if they are differentially expressed in states of disease, they can be used as biomarkers independent of their origin.

This study presents a workflow for comprehensive lipidomics of EVs using two isolation methods that are compatible with downstream state-of-the art LCMS, improving our ability to study the lipid components of EVs and identifying new disease biomarkers.

428

Powerful, standardized and versatile low-flow LCMS platform for deep dive proteomics and high-throughput analysis with maximum MS utilization

Oleksandr Boychenko[1], Christopher Pynn[1], Runsheng Zheng[1], Peter Jehle[1], Mike Baynham[1]

1. Thermo Fisher Scientific, Germering, Germany

Standardization of technology platforms and methods is a key for wide adoption of LCMS in different applied areas. However, in reality, the number of different instruments, setups, configurations, and methods for typical LCMS proteome profiling is extremely high that complicates the transfer of the developed method.

The typical proteomics workflow includes experiments for deep dive proteome analysis followed by an analysis of large sample cohorts with high-throughput, sensitive and robust LCMS. Whilst nanoLCMS is capable of both specific and significant proteome depth, its widespread adoption has been limited by the low throughput and insufficient robustness of the methods employed.

Here we demonstrate a new set of capillary-flow LC-MS (capLC-MS) methods capable of large sample cohort analysis using a Thermo Scientific™ UltiMate™ 3000 RSLCnano system coupled to a new Thermo Scientific™ Exploris™ 480 Hybrid Quadrupole-Orbitrap™ mass spectrometer.

The five low-flow LC-MS methods are capable of throughputs of 180, 100, 60, 30 and 24 samples per day affording MS utilization from 75 to over 90% respectively. These were validated using HeLa protein and crude plasma digests and showed excellent long term reproducibility and good protein coverage with more than 150 protein groups identified using the 8 min LC-MS method increasing to over 250 protein groups for the 60 min LC-MS plasma runs. The results were compared with standard nanoLCMS with 45 and 90 min peptide elution window.

The low-flow LCMS methods in combination with the high sensitivity and fast acquisition speed of the Exploris 480 mass spectrometer provide robust, fast and quantitative and deep profiling for complex sample matrices including crude plasma protein digests and cell lysates. Furthermore, the consistent results generated over hundreds of replicate injections, prove that the methods are suited to the analysis of large sample cohorts and biomarker studies.

429

Understanding complex post-translational modifications in CHO cells

Aiden Beauglehole[1], Lucia Zacchi[1, 2], Axayacatl Gonzalez[1], Esteban Marcellin[1], Ben Schulz[1, 2]

1. Australian Institute for Bioengineering and Nanotechnology, AIBN, Brisbane, Queensland, Australia

2. School of Chemistry and Molecular Biosciences, SCMB, Brisbane City, Queensland

Biologics production in CHO cells is a rapidly expanding field that makes up over 1/3 of the pharmaceutical industry. The expansion of biologics in yield and quality has been focused around physical processes such as bioreactors, media quality, and cell density rather than focusing on cellular machinery. The lack of focus on cellular machinery has left complex proteins with large amounts of post-translational modifications (PTMs) to have comparatively small yields and qualitative issues to other biologics. Compounding issues, biologics production is centered on the stable integration of DNA into the genome. Stable integration into the genome is random meaning that the gene can be inserted anywhere with multiple gene copy numbers, collectively called 'position effects'. In order to understand biologic production variation, position effects need to be minimised. Factor IX is a biologic that has numerous PTMs with 10 fold lower yields in contrast to other biologics. Utilising CRISPR/Cas9, we have developed CHO cell lines that are able to study protein production independent of position effects. By removing position effects, comparative SWATH based proteomics and metabolomics can occur to understand why Factor IX has poor yields compared to high producing biologics. Additionally, SWATH based proteomics allowed us to observe recombinant protein expressions effect on Factor IX's post-translational modifications and other biologics as a whole. Our findings can be used to engineer CHO cells to cope with the metabolic and biological demands of making Factor IX and other biologics with large amounts of PTMs.

430

Establishing the proteomic profile of human neuroblastoma cells during neuronal differentiation

Fiona Bicknell[1], Jacqueline Melvold[1], Bruce Milthorpe[1], Matt Padula[1]

1. University of Technology, Sydney, Broadway, NSW, Australia

The immortalized human neuroblastoma cell line SH-SY5Y has long been utilized for *in vitro* studies, its usefulness being attributed to its dopaminergic characteristics and ability to differentiate into functional neuron-like cells. Notably, retinoic acid (RA) treatment of SH-SY5Y cells has yielded multiple insights into neuronal differentiation in both the developing and adult brains. In addition, RA-differentiated SH-SY5Y cells provide an *in vitro* model to study diseases related to abnormalities in dopamine signalling such as Parkinson's and schizophrenia. Although the proteome of undifferentiated SH-SY5Y cells has been extensively studied, a comprehensive proteomic profile of these cells during and after differentiation is required. Therefore, this study aims to expand on existing proteomic data of RA-differentiated SH-SY5Y cells via quantitative LC-MS/MS. The data generated will enhance understanding of the differentiation process, as well as uncover potential targets for therapeutic agents, with the long-term outcome being improved treatments for relevant neuronal diseases.

431

Improved middle-down characterization of antibodies using multiple ion activation techniques and Proton Transfer Charge Reduction on an Orbitrap Eclipse mass spectrometer

Romain Huguet[1], Steve Binos[2], Kristina Srzentic[1], John E.P. Syka[1], Christopher Mullen[1], Joshua Silveira[1], Jennifer Sutton[1], Luca Fornelli[3]

1. Thermo Fisher Scientific, San Jose, CA, United States

2. Thermo Fisher Scientific, Parkville, VIC, Australia

3. University of Oklahoma, Normal, OK, United States

Relative to peptide-based mass spectrometry approaches, middle-down (MD) strategies offer the advantage of higher molecular sequence integrity in the characterization of biotherapeutics. However, reaching the protein sequence coverage required to confidently verify the complete mAb primary sequence or map post-translational modifications in MD experiments can be challenging. We apply advanced ion activation techniques, including ultraviolet photodissociation (UVPD) and electron transfer dissociation (ETD) in combination with Proton Transfer Charge Reduction (PTCR) reactions to improve IgG1 mAb characterization via liquid chromatography (LC)-MS. We have demonstrated that rapid (8 min long) gradient middle-down LC-MS/MS analyses of digested NIST mAb collectively yielded sequence coverage in excess of 80%. 90% sequence coverage was observed for Fc/2 and LC and around 80% for Fd mAb subunits. ETD and UVPD spectra of polypeptide precursors over 15 kDa are extremely complex and therefore are challenging to process with current *m/z* to mass spectral conversion software. The high spectral density of these product ion peaks is such that isotopic peak clusters often overlap and are frequently not sufficiently resolved to be differentiated from noise, nor to accurately assign a charge state and monisotopic peak, and thus not properly converted to neutral masses. To achieve more complete sequence coverage, product ion peak clusters associated with fragmentation in the middle region of the mAb subunits must be observed and properly mass deconvoluted. We extend our LC-MS/MS analyses of mAb subunits by utilizing PTCR subsequent to ETD and UVPD to enhance sequence coverage. In a single LC run, we obtained for Fc/2 and Fd over 50% and around 60% for LC mAb subunits. All the Complementarity Determining Regions (CDRs) were partially to fully sequenced. In case of Fc/2, this strategy unraveled an extensive series of large *z*-ions, unambiguously confirming the glycosylation site of the most abundant glycan variant.

432

Antibody Panel Based N-glycan imaging of patient serum for cancer biomarker discovery

Alyson P Black[1], Peggi M Angel[1], Richard R Drake[1], Anand S Mehta[1]

1. Medical University of South Carolina, Charleston, SC, United States

The vast majority of biomarkers used in the detection of cancer are glycoproteins, and studies indicate that the N-glycan component of the glycoprotein can act as a better marker of cancer than the protein component. However, accurate glycoprotein

biomarker assays are lacking, and there is a need for higher throughput biomarker discovery. We have developed a new platform for multiplexed N-glycoprotein biomarker analysis from patient serum. This platform combines matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) workflows with antibody slide arrays. Antibody Panel Based (APB) N-glycan imaging allows for specific capture of N-glycoproteins by antibodies and N-glycan analysis in a protein-specific and multiplexed manner. Development of this technique has focused on human serum glycoproteins alpha-1-antitrypsin, haptoglobin, hemopexin, immunoglobulin G, low molecular weight kininogen, and transferrin. Using purified standard solutions and human serum samples, glycoproteins were specifically captured and followed by enzymatic release of N-glycans. N-glycans are detected with a MALDI mass spectrometer in a localized manner along the array. Glycoproteins were captured in a concentration-dependent manner while maintaining specificity of capture. As a proof of concept, cirrhotic patient serum samples were compared to healthy serum, and a previously shown increase in an IgG N-glycan was observed corresponding to disease status. This novel approach to protein-specific N-glycan analysis from an antibody panel can be further expanded to include any glycoprotein for which a validated antibody exists, allowing analysis of potentially 100s of individual glycoprotein targets from a patient sample in just one imaging run. Additionally, this platform can be adapted for analysis of any biofluid or biological sample that can be analyzed by antibody arrays. This technique has exciting potential to be applied in the clinic as both a biomarker discovery tool as well as a screening tool in readily available clinical biofluid samples with minimal consumption.

433

Proteomic analysis of cellular secretory responses in a human lung-challenge model of tuberculosis disease

Jonathan Blackburn[1], Clemens Hermann[1], Malika Davids[1], Anil Pooran[1], Bridget Calder[1], Keertan Dheda[1]

1. University of Cape Town, Cape Town, WESTERN CAPE, South Africa

Background: Despite increased global efforts, Tuberculosis (TB) remains the biggest killer amongst infectious diseases today. Due to the rise of multi-drug resistant strains and the absence of effective vaccines, the search for correlates of protection has become a major objective in TB research, to enable the development of novel and effective vaccines. Surprisingly, the nature of the lung mucosal proteome during TB infection is barely known, as most studies have only characterised a small subset of cyto- and chemokines, thus leaving the possible involvement of the majority of the lung mucosal proteome in the dark.

Methodology: In this first-in-mankind clinical study, we used a controlled human lung-challenge model to focus on neglected basic immunobiological aspects of *Mycobacetrium tuberculosis* (Mtb) infection in the lung. Bronchoalveolar lavages (BAL) were performed on patients with different susceptibility profiles, ranging from patients with presumed sterilising immunity, to those with latent TB infection, to patients with one or more episodes of active TB. BAL samples were taken both at baseline and after a three-day challenge with either live Bacillus Calmette–Guérin (BCG) or purified protein derivative (PPD). The mucosal cellular immune response was characterised using in-solution tryptic digestion and LC-MS/MS measurement on a Q Exactive mass spectrometer.

Findings: Overall, 3650 protein groups were identified in BAL fluid. Both challenges with BCG or PPD challenge induced dysregulation of the mucosal proteome compared to the baseline sample. However, PPD challenge resulted in twice as many dysregulated proteins than BCG challenge. Furthermore, patients from the four patient groups responded differently towards the both challenges according to their previous exposure to *Mycobacterium tuberculosis*.

Conclusion: Intra-lung challenge with PPD induced stronger mucosal immune responses compared to BCG challenge. We propose this proteomic approach to monitor intra-lung challenges with different vaccines, ultimately to enable improved design of effective vaccines and therapeutic interventions.

434

Predicting immunotherapy treatment outcomes in melanoma patients

Jonathan Blackburn[1], Jessica Da Gama Duarte[2], Jonathan Cebon[2], Andreas Behren[2]

1. University of Cape Town, Cape Town, WESTERN CAPE, South Africa

2. Olivia Newton John Cancer Research Institute, Melbourne, Victoria, Australia

Background: Immunotherapies with checkpoint inhibitors have revolutionized the treatment of advanced melanomas. Nonetheless, not all patients benefit clinically from these immunotherapeutics and some develop immune-related adverse events (irAEs). Therefore, the ability to predict treatment outcomes in a non-invasive manner is of utmost value. The presence of tumour antigen-specific antibodies in circulation is indicative of immune engagement with the tumour, while an over-reactive autoantibody repertoire may predispose patients to develop irAEs. Thus, we aimed to determine the unique antibody repertoire of melanoma patients prior to undergoing immunotherapy with checkpoint inhibitors, to investigate if it may predict the likelihood of treatment success and the onset of irAEs.

Methods: Blood samples from 15 advanced melanoma patients were collected at baseline before treatment with anti-CTLA-4 or anti-PD-1 checkpoint inhibitors under HREC approved study protocol, with informed consent. Patients clinically developed high-grade, low-grade or no irAEs following treatment. Humoral immune responses were investigated using the Sengenics Immunome protein array, a high-content array with 1627 full-length, folded, immobilized tumour and self-antigens that enabled the interrogation of the depth and breadth of the immune response.

Results: In patients treated with CTLA-4 or PD1 checkpoint blockade, the detection of antibodies against 10 to 60% of all antigen specificities on the array content at baseline was predictive of clinical response to treatment, including both partial and complete responses. In instances where clinical response was accompanied with the onset of irAEs, antibodies were detected against more than 80% of the array content. However, when patients progressed or developed irAEs without clinical benefit, antibodies were detected against less than 10% of the array content.

Conclusions: Investigating humoral immunity may predict clinical responses and the onset of irAEs in patients treated with immunotherapies. Personalized therapeutic decisions will prevent patients from undergoing unnecessary treatments, while leading to improved clinical responses and outcomes.

435

Sub-lethal Rifampicin effects on the Mycobacterium smegmatis cell wall and on survival in macrophage infection models

Jonathan Blackburn[1], Alexander Giddey[1], Nelson Soares[1]

1. University of Cape Town, Cape Town, WESTERN CAPE, South Africa

Background: Tuberculosis disease, caused by *Mycobacterium tuberculosis*, is one of the leading causes of mortality globally. Mycobacterial drug resistance is a growing concern with frontline drugs such as rifampicin in danger of becoming ineffective. We report the effects of sub-lethal rifampicin exposure on the cell wall proteome and lipidome of both susceptible (WT) and drug-resistant (SL) isogenic engineered strains of *Mycobacterium smegmatis* and validate a data derived hypothesis regards bacterial survival.

Methods: Both rifampicin susceptible and resistant *Mycobacterium smegmatis* strains were treated with sub-lethal doses of rifampicin. The lysate was enriched for the cell wall, cytosolic, and cellular debris fractions; peptides prepared via FASP and analysed on our Q-Exactive by LC-MS/MS. Proteins were identified and quantified via MaxQuant and statistically analysed in R for dysregulation and pathway enrichment via STRINGdb.

Results: We identified a total of 2632 proteins with 646 and 258 found to be dysregulated in susceptible and resistant strains respectively. GO term enrichment showed enrichment for 75 and 46 KEGG pathways respectively with ABC Transporters enriched in the top 10 terms, with majority of proteins decreased abundance, in both strains. Porphyrin and chlorophyll metabolism was among the top 25 terms in both with dysregulation of specific enzymes conserved across drug-sensitivity. Both strains showed evidence for downregulation of mammalian cell entry proteins known to be important in infection of some cell types and survival of mycobacteria in macrophages. Pre-treatment with sub-lethal concentrations of rifampicin showed reduced uptake and survival (24 hours) of rifampicin susceptible mycobacteria in a macrophage infection model.

Conclusions: Mycobacteria showed conserved responses to sub-lethal rifampicin treatment across drug-sensitivity suggesting a specific, non-stress-induced {adaptive} response to this important front-line drug. Mycobacterium smegmatis appears to respond to rifampicin treatment by reducing cell wall permeability and trans-cell wall transport with the consequence of impaired infectivity and survival in macrophages.

436

Enhanced protein identification using harsh cell disruption combined with bead-based protein preparation (SP3-protocol) in a *Legionella pneumophila* cell culture infection model

Sascha Blankenburg[1], Christian Hentschker[1], Christina Herkt[2], Kristin Surmann[1], Petra Hildebrandt[1], Stephan Michalik[1], Anna Nagel[1], Denise Dittmar[1], Manuele Gesell Salazar[1], Bernd Schmeck[2], Uwe Völker[1]

1. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Mecklenburg-Vorpommern, Germany

2. Philipps-University Marburg, Institute for Lung Research, Marburg, Hessen, Germany

Background

Community acquired pneumonia (CAP) is of high importance being the third most common cause of death worldwide. One third of all CAP-patients, which need admission to intensive care, derive from Legionnaires' disease. However, the molecular interactions between host and pathogen are only partially resolved. Therefore, we used a human alveolar macrophages *Legionella pneumophila* infection model to profile the interplay of both partners by proteomics. Since such studies are often hampered by low sample amounts, we optimised a bead-based sample preparation protocol to enable comprehensive proteomics analysis of low *L. pneumophila* cell numbers and its corresponding host upon internalization.

Methodologies

L. pneumophila was cultivated in BCYE medium and $2x10^6$ or $5x10^6$ bacteria were filtered. Bacteria on filters were comparatively processed using three different cell disruption and digestion methods. A combination of SDS based cell disruption und tryptic digestion with an adapted single pot solid phase sample preparation (SP3) protocol was finally used in an infection assay. Samples of THP-1 cells and *L. pneumophila* Corby were generated by cell sorting using a FACSAria™ 8 h, 12 h, and 16 h post infection. Finally, peptides of host and pathogen prepared with the optimized workflow were analyzed by nanoLC-MS/MS with a Q Exactive™ Plus mass spectrometer.

Results

Coverage of the *L. pneumophila* proteome was increased up to 300% at the protein level and up to 620% at the peptide level with accompanied improvements in reproducibility, protein quantification, and data completeness. Overall, 1650 *L. pneumophila* proteins and 2967 THP-1 proteins were identified and quantified from the infection setting. Time-resolved, interdependent changes in the proteomes of pathogen and host will be discussed in the presentation.

Concluding statement

The adapted SP3-protocol in combination with harsh cell disruption enables strongly increased protein and peptide identification and therefore, allows new insights into host-pathogen interactions during infections.

437

Proteins co-ordinating cell permeability and nutrient transport in the rumen epithelium of sheep.

Jude J Bond[1], Karthik S Kamath[2], Matthew McKay[2], David Wheeler[3], Victor H Oddy[1]

1. NSW Department of primary industries, Armidale, NSW, Australia

2. Department of Molecular Science, Australian proteome analysis facility, Macquarie university, Ryde, NSW, Australia

3. Orange Agricultural Institute, NSW Department of primary industries, Orange , NSW, Australia

We examined the processes represented by cytosol and membrane rumen epithelial proteins controlling nutrient transport using proteomics.

Ventral rumen tissue from 20 sheep with variation in CH4, CO2 emissions and mean retention time of digesta were used to prepare cytosol and membrane fractions using the method described by Bond et al. (2018). Protein extracts were identified and quantified using SWATH MS using the Ensembl Ovine database. Bioinfomatic tools allocated proteins to their sub-cellular compartments.

We identified 3162 proteins of which 2599 proteins were quantified. Those that control paracellular permeability include 2 claudins (CLDN 1 and 4). We identified 43 plasma membrane solute carrier family (SLC) transporters involved in the transport of inorganic ions, amino acids, lipids, sugars and metabolic intermediates. These are regulated by electrochemical gradients. In addition, 8 ATP-binding cassette (ABC) transporters require ATP for active transport of ions and phospholipids were identified. Nine subunits of the V-ATPase proton pump which acidify intracellular compartments such as the lysosome and phagosome where identified. Transporters identified include a maxi-Cl- anion channel sub-unit and several less well characterised Cl- channels not previously reported in the rumen epithelium. We did not detect any significant difference in the abundance of transporter proteins, or those allocated to cell compartments bound by membranes included endosome, lysosome, exosome, phagosome and peroxisome associated with differences in the sheep sampled.

The in-depth coverage of cytosol and membrane proteins in isolated epithelium has revealed some of the specific proteins which regulate the permeability and transport of nutrients in the rumen epithelium of sheep. It has helped us to assemble a model of the processes regulating nutrient transport that occur in ruminants. Our comprehensive results inform linkage between phenotypes of nutrient use efficiency with cellular protein markers in the rumen epithelium.

Bond et al. 2019 Journal Animal Science and Biotechnology 10

438

Protein-based Cardiogenic Shock Patient Classifier

Eva Borràs[1, 2], Ferran Rueda[3], Cosme García-García[3], Oriol Iborra-Egea[3], Antoni Bayés-Genís[3], Eduard Sabidó[1, 2]

1. Proteomics Unit, Universitat Pompeu Fabra (UPF), Barcelona, Spain

2. Center for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

3. Heart Institute, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain

Cardiogenic shock (CS) is associated with high short-term mortality, substantial morbidity and resource utilization. Despite widespread use of early coronary reperfusion, CS prevalence remains unaltered and it is the leading cause of in-hospital death. Although contemporary risk scores are available (i.e. CardShock and IABP-SHOCK II) they mostly rely on classical clinical acumen and conventional laboratory variables and more accurate risk stratification strategies are needed to guide interventions to improve patients outcome. Here we developed a circulating protein-based classifier able to predict short-term mortality risk among patients with CS.

In an initial proteomics screening, a cohort of 48 CS patients (Barcelona cohort) was used to select candidate proteins to predict short-term mortality risk. Then, 51 selected proteins were quantified by targeted proteomics (PRM) in an independent European multicenter cohort of 97 patients (CardShock cohort). Concretely, chromatographic profiles were obtained for all measured peptides and compared to their internal references for accurate relative protein quantification.

The classification power of these proteins was evaluated resulting in CS4P (Cardiogenic Shock 4 Proteins), the combination of four plasma proteins able to predict 90-day risk of mortality among CS patients (AUC=0.83). CS4P comprises the abundances of liver fatty acid-binding protein (L-FABP), beta-2-microglobulin (B2MG), fructose-bisphosphate aldolase B (ALDOB), and SerpinG1 (IC1). Moreover, the CS4P model in combination with contemporary CardShock risk score improved mortality prediction and patient reclassification compared with the CardShock risk score alone (AUC 0.84 *vs*. AUC 0.78). The CS4P patient classification power was confirmed by enzyme-linked immunosorbent assay (ELISA). This validation will favor its translation into routine clinical practice and ultimately guide clinicians in selecting patients for advanced therapies.

1. Rueda et al, European Heart Journal (In press)

439

Understanding the storage protein biosynthesis and protein compensation in hordein double-null barley lines using SWATH mass spectrometry

Utpal Bose[1], James A. Broadbent[1], Keren Byrne[1], Malcolm Blundell[2], Crispin A. Howitt[2], Michelle L. Colgrave[1]

1. Agriculture and Food, CSIRO, St Lucia, QLD, Australia

2. Agriculture and Food, CSIRO, Canberra, ACT, Australia

Background

The hordeins are the major barley seed storage proteins and are elicitors of coeliac disease - a condition that affects ~1% of the world population. There are four multi-genic hordein families in barley: B-hordeins representing 70-80% of total gluten; C-hordeins (10-20%); D-hordein (<5%); and, γ-hordeins (1-2%). Using conventional breeding a series of null lines were developed in which different classes or combinations of classes of hordeins are absent.

Methodology

Data-independent acquisition (DIA) mass spectrometry (MS) was used to measure proteome-wide abundance differences between wild-type and selectively-bred hordein double-null barley lines. Data were acquired in information-dependent acquisition (IDA) and sequential window acquisition of all theoretical fragment-ion spectra-mass spectrometry (SWATH-MS) on a TripleTOF 6600 MS (SCIEX, USA). Statistical and functional analyses were performed on the proteins perturbed between wild-type and hordein double-null barley lines.

Findings

A total of 6,138 peptides mapping to 1,907 proteins were quantified at a 1% false discovery rate. Pairwise comparisons revealed proteome-wide alterations for BC-, BD- and CD-null lines in the order of ~16%, ~10% and ~14%, respectively. As an example, the comparison between wild-type and the BC-null line identified 151 up-regulated proteins (7.9%), while 145 (7.6%) proteins were down-regulated in the BC-null line. Contextualization within a protein-protein interaction network reveals the up-regulation of proteins associated with primary metabolism, transcription and enzymatic biosynthesis processes while down-regulation of heat shock proteins can be found in the BC-null line. Gene Ontology (GO) analysis (molecular function) revealed that enzymatic activities were up-regulated, whilst nutrient reservoir activities were down-regulated in the BC-null line.

Concluding remarks

Overall, the GO-based analysis provides an overview of the functional classes that are perturbed during the breeding process. Grain proteome profiling delivers an informative molecular portrait of the hordein double-null lines and the underlying storage protein biosynthesis mechanisms thereby shedding light on mechanism of proteome compensation.

44

Mass spectrometry– teaching us new lessons in immunity

Anthony Purcell[1, 2]

1. Monash University , Clayton, VIC, Australia

2. Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Immunopeptidomics is the study of the repertoire of peptide ligands selected for antigen presentation by MHC class 1 molecules. In recent years, this has been revolutionised by the use of in depth qualitative and quantitative mass spectrometry techniques. In particular, fundamental and novel insights have been made that shed light on the dynamic range of antigen presentation, the role of HLA polymorphism, antigen presenting cell type and mode of antigen processing on the resultant immunopeptidome, as well as diversity introduced by degenerate proteolysis and extensive post-translational modifications. I will highlight recent advances in analytical workflows and how these have provided new insights into antigen processing and presentation by selecting recent published and unpublished examples from our laboratory (***1-7***).

References:

1. Illing, P. T. et al. HLA-B57 micropolymorphism defines the sequence and conformational breadth of the immunopeptidome. **Nature communications** 9, 4693 (2018).

1. Pymm, P. et al. MHC-I peptides get out of the groove and enable a novel mechanism of HIV-1 escape. **Nat Struc Mol Biol** 24, 387-94 (2017).

1. Ooi, J. D. et al. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. **Nature** 545, 243-7 (2017).
2. Faridi, P. et al. A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. **Science Immunol** 3, pii: eaar3947 (2018).

1. Croft, N. P. et al. Most viral peptides displayed by class I MHC on infected cells are immunogenic. **Proc Natl Acad Sci U S A**, 116:3112-17 (2019).

1. Koutsakos, M *et al.* Human CD8+ T cell cross-reactivity across influenza A, B and C viruses, **Nat Immunol**, 20:613-25 (2019)
2. Woods, K. *et al.* The diversity of the immunogenic components of the melanoma immunopeptidome, **bioRxiv**, https://doi.org/10.1101/623223 (2019)

440

Differential Glycosylation and Novel *N*-Glycan Site Analysis of ECM Proteins in Gynaecological Cancers

Matthew T Briggs[1], Mark R Condina[1], Yin Ying Ho[2], Georgia Arentz[2], Parul Mittal[2], Manuela Klingler-Hoffmann[1], Gurjeet Kaur[3], Martin K Oehler[4], Peter Hoffmann[1]

1. Future Industries Institute, The University of South Australia, Mawson Lakes, SA, Australia

2. The University of Adelaide, Adelaide, SA, Australia

3. Institute for Research in Molecular Medicine, Universiti Sains , Pulau Pinang, Malaysia

4. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

Gynaecological cancers are a group of cancers that originate from the vulva, vagina, cervix, uterus (endometrium), placenta, fallopian tubes and ovaries in adult women. These cancers are known to undergo an epithelial-mesenchymal transition (EMT) which results in the metastasis of epithelial cells and an increase in their resistance to apoptosis by altering the extracellular matrix (ECM). ECM proteins play an imperative role in cell health as they provide the scaffold upon which cells and tissues are built, hence ECM proteins directly and indirectly influence almost all cellular processes, including cell differentiation, proliferation, and motility. The majority of ECM proteins are known to be glycosylated, which is the most prevalent type of post-translational modification (PTM). In this study, we extracted proteins from cervical, vulvar, endometrial and ovarian cancer tissues with their corresponding normal tissues. These tissues were then treated sequentially with PNGase F, allowing for the analysis of cleaved *N*-linked glycans, followed by trypsin, allowing for the analysis of tryptic peptides. Furthermore, consecutive tissue sections were treated with trypsin followed by zwitterion chromatography-hydrophilic interaction liquid chromatography (ZIC-HILIC) enrichment, allowing for the analysis of glycopeptides. The *N*-glycan, tryptic peptide and glycopeptide samples were analysed by liquid chromatography tandem mass spectrometry (LC-MS/MS), in order to identify and quantify the oligosaccharide chains and the *N*-glycosite of peptides. *N*-glycosite peptide analysis revealed differential glycosylation of ECM proteins from cancerous and normal tissues. However, further investigation with a larger patient cohort is required to assess and validate these findings.

441

PromarkerD as an immunoaffinity mass spectrometry assay for diabetic kidney disease

Scott Bringans[1], Kirsten Peters[1], Tammy Casey[1], Richard Lipscombe[1], Sarah Thomas[1], Stephen R Pennington[2], Orla Coleman[2], Holger A Ebhardt[2]

1. Proteomics International, Broadway, Nedlands, WA, Australia

2. Atturos, Dublin, Ireland

The PromarkerD assay for diabetic kidney disease was originally developed as a multi-protein targeted mass spectrometry assay directly from depleted (top 14 proteins) digested human plasma. This assay has been adapted to an ELISA format and more recently an immunoaffinity mass spectrometry assay. The immunoaffinity method utilises bead-based antibody binding for the specific PromarkerD protein biomarkers in single multiplex capture step. The captured protein biomarkers are reduced, alkylated and digested in situ on the beads with injection onto a microflow LCMS system for targeted mass spectrometry. The results obtained with the immunoaffinity method applied to a 100-person cohort were compared to the original direct plasma digestion method with correlation between the two methods confirmed by Bland and Altman plot analysis. To test the robustness of the process between laboratories the assay was also performed in laboratories in Australia and Ireland using the same samples. The advantages of the immunoaffinity technology developed for this assay are three-fold. Firstly, increased throughput of analysis with a 96 well based robotic handling capable system which also minimises human intervention and handling. Secondly, samples that are injected onto the LCMS system are in a much cleaner form than a crude plasma digest which enhances sensitivity and reduces machine down time due to less frequent source cleaning and minimal LC blockages (in micro flow format). Thirdly, the immunoaffinity method has been designed to be available as a simple technology transfer process to partner laboratories that have LCMS capability. These advantages may make such an immunoaffinity-MS technology approach a superior choice for multi-protein biomarker diagnostic assays.

442

Synaptic dysfunction investigated in clinical cohorts by immunoprecipitation/mass spectrometry.

Ann Brinkmalm[1], Annika Öhrfelt[1], Henrik Zetterberg[1], Kaj Blennow[1]

1. University of Gothenburg, Mölndal, NA, Sweden

Synaptic dysfunction is a key pathogenic event in neurodegenerative and psychiatric diseases. Cerebrospinal fluid (CSF) biomarkers reflecting synaptic integrity would be highly valuable tools to monitor synaptic dysfunction directly in patients. We have recently developed mass spectrometry based CSF-assays for the pre-synaptic proteins SNAP-25 and synaptotagmin-1 and employed them in the study of neurodegenerative and psychiatric diseases.

The assays have undergone several improvements in sensitivity and through-put which will be discussed. Briefly, immunoprecipitation of SNAP-25, SYT-1 or SNAP-25/SYT-1 was performed on a KingFisher™ Duo or KingFisher™ Flex Purification System. Isotopically labeled standards were added and the enriched proteins were digested with trypsin or trypsin/Lys-C. MS based quantification was performed either with LC-PRM-MS on a Q Exactive/Ultimate 3000 system (Thermo Fisher Scientific) or LC-MRM-MS on a 6495 Triple Quadrupole LC/MS system (Agilent Technologies).

CSF SNAP-25 and SYT-1 concentrations in AD compared to controls were found to be significantly higher also at the very early stages of the disease. Interestingly, a longer soluble form of SNAP-25 provides improved differentiation between the different groups of patients and controls than all soluble SNAP-25. A study on patients with small cell lung cancer indicate that SNAP-25 is expressed both in presynaptic and postsynaptic parts of the neuron and that CSF levels of the different forms mirror different populations of the SNAP-25. The assay has also been used to monitor secretion of SNAP-25/SYT-1 during cortical neuronal differentiation of human induced pluripotent stem cells. Recently the combined SNAP-25/SYT-1 assay has been employed in a study of more than 2000 patients from several neurodegenerative diseases.

We conclude that IP-MS can be employed in clinical studies of neurodegenerative and psychiatric diseases with sufficient number of patients to reveal biologically relevant results.

443

Identification of covalently cross-linked Aβ dimers in Alzheimer's disease brain by combining top-down and bottom-up mass spectrometry

Gunnar Brinkmalm[1], Wei Hong[2], Tiernan T O'Malley[2], Wen Liu[2], X Sun[2], Matthew P Frosch[3], Erik Portelius[1], Henrik Zetterberg[1], Kaj Blennow[1], Dominic M Walsh[2]

1. Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden

2. Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital, Boston, MA, USA

3. Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Objective

The amyloid β-protein (Aβ) is believed to play an initiating role in Alzheimer's disease (AD) and aqueous extracts of certain AD brains can cause Aβ-dependent toxicity. When examined using denaturing SDS-PAGE such extracts contain two broad bands at ~4 and ~7kDa. While ~4kDa-Aβ has been shown to contain a variety of Aβ primary structures, the identity of ~7kDa-Aβ, has been controversial. Here, we report methods that allowed the identification of components of ~7kDa Aβ.

Methods

~4kDa-Aβ and ~7kDa-Aβ were isolated from the aqueous phase of human brain and solubilized amyloid plaques. Long-term potentiation (LTP) and live-cell imaging of iPSC-derived neurons was used to assess bioactivity, and portions of the same materials were used for mass spectrometry. Analysis of intact and proteolytically digested ~4kDa-Aβ and ~7kDa-Aβ was performed using LC-MS/MS under both acidic and alkaline conditions. Several types of database search softwares, in-house written scripts, and manual analysis, were utilized to identify ~7kDa-Aβ.

Results

~7kDa-Aβ, but not ~4kDa-Aβ, blocked hippocampal LTP and disrupted the neuritic architecture of iPSC-derived neurons. Analysis of ~4kDa-Aβ from six separate AD brains identified more than 35 Aβ primary structures - including the previously elusive p3 fragment, Aβ17-42. When alkaline conditions were used to analysis ~7kDa-Aβ solubilized from plaques, 11 covalently cross-linked Aβ heterodimers were detected. These included: 1-38×1-40, 1-40×1-40, 2-40×1-40, and 1-42×1-42. Intriguingly, LC-MS/MS analysis of trypsin-treated ~7kDa-Aβ identified Aβ dimers linked between Asp1 and Glu22.

Conclusions

Our demonstration of covalently cross-linked Aβ heterodimers in human brain recommends further investigation of dimers as therapeutic targets and potential AD biomarkers. However, large amounts of material were required for our analysis and only the most abundant dimers were positively identified. In future studies it will be essential to increase analytical sensitivity and develop software to allow the identification of low abundant dimers.

444

Proteome survey of wound fluid from non-healing wounds reveals key biological processes associated with poor healing outcomes

Daniel A Broszczak[1, 2], James A Broadbent[1, 2], Tony J Parker[1, 2]

1. School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

2. Tissue Repair and Translational Physiology, Institute of Health and Biomedical Innovation, Kelvin Grove, Queensland, Australia

Non-healing wounds are a significant problem for patients and healthcare systems worldwide. The underlying biochemistry, which drives non-healing outcomes in self-perpetuating leg wounds, is poorly understood. To address this knowledge deficit, a study of the proteins that compose the fluid, which exudates from these wounds, may provide important insight regarding treatment response and healing outcome for patients. In this respect, we have conducted a clinical study that included the collection of biological samples and clinical / psychosocial data over a 24 week period, during which time patients received best-practice care. Biological samples were analysed using data-independent acquisition SWATH-MS to identify and quantify the protein complement of the wound fluid. The resulting data were integrated with clinical measurements and contextualized by gene ontology annotations to enable deeper insight into the dynamic biological processes taking place within non-healing wounds. This identified key biological processes that may indicate specific underlying issues for a sub-set of wounds and their recalcitrant nature towards clinical care. A number of biological markers that are indicative of the wound healing outcome were also derived from these analyses. Unravelling the complex biology of non-healing wounds through proteome and clinical data integration provides some insight into the mechanisms associated with a patient's adverse or positive responses to clinical care. Such information can be developed further to inform clinical practices and enable the meaningful personalisation of wound management.

445

A polyomic approach to understand the molecular basis of virulence in Leishmania

Richard Burchmore[1], Abduabset Kabli[1], Kathryn Crouch[1], Michael Barrett[1]

1. University of Glasgow, Glasgow, SCOTLAND, United Kingdom

Leishmania have the ability to subvert the host immune system and adopt sophisticated strategies to develop and survive within the mammalian host. Leishmania parasites survive within macrophages, inside a parasitophorous vacuole. The molecular communication between host and parasite decides the outcome of infection, but is incompletely understood. We have compared genotype and phenotype of an attenuated *Leishmania mexicana*line with a virulent, isogenic wild type precursor. We aim to identify key virulence factors and to explore the potential of the attenuated line as a vaccine candidate.

This study was conducted on promastigotes of *Leishmania mexicana*and involvedcomparative polyomics approaches to identify the molecules that contribute to Leishmania virulence. Log phase promastigotes of wild-type and gentamicin-attenuated (H-line) were grown in parallel in media containing 10% FBS. For comparative proteomic analysis, protein extracts were labelled using 6-plex TMT and analyzed with LC-MS/MS. For metabolomics, metabolites were extracted with Chloroform/ Methanol/ Water (1:3:1) and analyzed with LC-MS. For transcriptomics, RNA was isolated and converted into a library of cDNA molecules for cluster generation and DNA sequencing. We found 18 proteins differentially expressed in attenuated Leishmania (FC ≥1.5 and FDR ≤0.05) and 26 identified metabolites (FC ≥1.5 and FDR≤0.05), whereas transcriptomics data found 481 transcripts were differentially expressed (FC ≥1.5 and FDR ≤0.05). Most of the differentially expressed proteins and transcripts were metabolic enzymes, and the majority of the differentially expressed metabolites were substrates that involved in nucleotide, carbohydrate and amino acid metabolisms. Correlation of polyomics datasetsreveals that nucleotide metabolism (pyrimidine and purine) is significantly altered in gentamicin-attenuated Leishmania. Furthermore, nucleobase and nucleoside transporters were significantly down regulated in proteomics analysis. Modulation of gene expression may relate to gentamicin selection. ΔNT3 cells became more sensitive to allopurinol (purine analogue) compared to wild-type cells (EC_{50}= 141.1), suggesting that this may contribute to Leishmania virulence.

446

Applying LC-MS/MS to provide insights into eukaryotic nitrogenase engineering

Keren Byrne[1], Michelle Colgrave[1], James A Broadbent[1], Rob Allen[2], Shoko Okada[2], Ema Johnstone[2], Christina Gregg[2], Craig Wood[2]

1. CSIRO, Agriculture and Food, Brisbane, QLD, Australia

2. CSIRO, Agriculture and Food, Black Mountain, ACT, Australia

Introduction

Improving the ability of plants and plant-associated organisms to fix and assimilate atmospheric nitrogen has driven plant biotechnologists for decades. The potential benefits are manifold: reduction in the negative environmental effects from chemical fertilizer use; availability of reactive nitrogen; potential economic benefits; and, opportunities to provide affordable crops to farmers in developing countries to increase yield. The combination of recent advances in bioengineering and increased knowledge about the biochemistry and biosynthesis of the nitrogenase enzyme has made this vision more possible. Of critical importance is the mitochondrial/chloroplast targeting and processing of the nitrogenase proteins in eukaryotes if the goal of nitrogen fixation in plants is to be a reality.

Methods

Liquid chromatography-mass spectrometry (LC-MS) offers many advantages over traditional methods for protein detection and quantitation, especially when dealing with transgenic membrane proteins that are often difficult to express or generate antibodies against. In this study, proteomics aimed to identify, characterise and quantify 16 nitrogenase (Nif) proteins that had been engineered into tobacco (*Nicotiana benthamiana*) and *Escherichia coli*. Samples were analysed with data-dependent micro-flow LC-MS/MS acquisition (SCIEX TripleTOF 6600) and multiple-reaction-monitoring LC-MRM-MS (SCIEX 6500 QTRAP).

Results

A panel of 168 peptides were tested for their suitability as Nif-specific MRM peptide markers. MRM methods have been developed for 14 nitrogenase proteins. A targeted LC-MS/MS method was employed to study the N-terminal region of the expressed NifK protein after engineering into tobacco to assess the mitochondrial processing and putative cleavage site. We have identified a mitochondrial processing peptidase (MPP) putative cleavage site and developed an assay capable of monitoring the efficiency of cleavage.

Conclusion

MRM technology is a valuable strategy to provide insights into the generation of modified plants where the N_2 fixation machinery is introduced and highlights the usefulness of interdisciplinary collaboration to progress the engineering of plants to express nitrogenase.

447

Pulse Data Independent Acquisition (PulseDIA) on Orbitraps for deeper proteomic profiling with relatively small sample consumption

Xue Cai[2, 1], Weigang Ge[2, 1], Rui Sun[2, 1], Xiao Yi[2, 1], Tiansheng Zhu[2, 1], Yi Zhu[2, 1], Tiannan Guo[2, 1]

1. Institute of Basic Medical Sciences, westlake university, Hangzhou, Zhejiang, China

2. 1. Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, westlake university, Hangzhou, ZHEJIANG, China

Background

Data independent acquisition (DIA) is rapidly progressing due to its high degree of accuracy, reproducibility, high-throughput data acquisition and small sample consumption. However, the sensitivity of DIA-MS method on Orbitraps is limited by the relatively

slow scan rate and window width. To reach a deeper proteome, multiple dimensional fractionation coupled with data dependent acquisition (DDA) usually consumes 100 µg peptides per sample.

Methods

A PulseDIA-MS method was developed to fractionate peptides at MS2 level to maximize the proteome depth with 0.5ug peptides per injection. In PulseDIA, each DIA window was divided into multiple small parts by multiple injections of the same sample with complementary scanning windows. Four parameters were tested and optimized to maximize the performance of PulseDIA: i) number of injections; ii) length of LC gradient; iii) fixed or variable window; iv) the width of overlaps between adjacent sub-windows. HeLa cell digest, breast cancer cell digests were used for the method benchmarking and optimization. The PulseDIA method was then applied to analyze 24 clinical tissue samples from 12 cholangiocellular carcinoma (CCC) patients. All data were analyzed using Spectronaut.

Results

PulseDIA analysis of breast cancer cell sample led to identification of 45,592 peptides and 5,346 proteins using a library containing 60,687 peptide precursors and 6,239 proteins, which is 57.9% and 28.2% higher than conventional DIA method. Five injections of HeLa cell digest increased the peptide and protein identifications by 100.1% and 26.5% respectively compared to conventional DIA. We further applied the PulseDIA method to analyze 24 CCC samples, and quantified 59,887 peptides (58.3% increase) from 5,426 proteins (8.3% increase) using two injections of 0.5 µg peptides. The pulseDIA analysis identified novel regulated proteins in this CCC cohort.

Conclusions

We present a novel DIA method called pulseDIA which allows deeper proteomic analysis using relatively small amount of peptide samples.

448

Identification of novel N-glycosylation sites from the bacterium *Campylobacter jejuni*

Joel Cain[2, 1], Stuart Cordwell[2, 1, 3, 4]

1. School of Molecular Bioscience, University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. Mass Spectrometry Core Facility , University of Sydney, Sydney, NSW, Australia

4. Discipline of Pathology, University of Sydney, Sydney, NSW, Australia

Introduction:

One of the prominent causative agents of acute gastroenteritis in the developed world is *Campylobacter jejuni.* An important factor for virulency, as well as its ability to establish a commensal relationship within avian guts (the primary vector for infection), is its capacity to perform protein N-linked glycosylation. Recently we established that loss of N-glycosylation was associated with an array of phenotypes including aberrant chemotactic responses, altered respiratory preferences, changes in cell morphology and increased susceptibility towards temperature and osmotic stress. Given the lack of concrete connections between known targets of this PTM and the affected phenotypic traits, we looked to broaden the list of known N-glycoproteins within *C. jejuni* with the aim of identifying putative causative agents.

Methodologies:

Membrane proteins from *C. jejuni* 11168 O were enriched using Na_2CO_3 precipitation, following protein aliquots were digested using an array of commercially available proteolytic enzymes. Intact N-glycopeptides were then enriched using ZIC-HILIC. N-glycopeptides were analysed on an Orbitrap Fusion™ Tribrid™ mass spectrometer using HCD and product-ion triggered CID fragmentation.

We also looked to identify N-glycoproteins from 11168 O ΔpglB::pglB with the rationale that the reported elevated abundance of the oligosaccharyltransferase, PglB, within this strain would lead to elevated levels of low abundance N-glycopeptides.

Key findings:

Of the known 131 modification sites, we were able to unequivocally demonstrate occupation of ~80% of known N-glycopeptides. We were able to identify 54 novel modification sites, including 15 novel N-glycoproteins.

449

Proteomics reveals multiple phenotypes associated with *N*-linked glycosylation in *Campylobacter jejuni*

Joel Cain[2, 1], Ashleigh Dale[2, 1], Paula Niewold [2, 3], William Klare [2, 1], Lok Man [2, 1], Melanie White [2, 3], Nichollas Scott [1, 4], Stuart Cordwell [2, 3, 1, 5]

1. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. Discipline of Pathology, University of Sydney, Sydney, NSW, Australia

4. Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria, Australia

5. Sydney Mass Spectrometry, University of Sydney, Sydney, NSW, Australia

Introduction:

Campylobacter jejuni is a major gastrointestinal pathogen generally acquired *via* consumption of poorly prepared poultry. *N*-linked protein glycosylation encoded by the *pgl* gene cluster targets >80 membrane proteins and is required for both non-symptomatic chicken colonization and full human intestinal epithelial cell virulence. Despite this, the biological functions of *N*-glycosylation remain unknown.

Methodologies:

Here we examined the effects of *pgl* gene deletion on the *C. jejuni* proteome using label-based liquid chromatography / tandem mass spectrometry (LC-MS/MS) and validation using data independent analysis (DIA-SWATH-MS). Targeted metabolomics was used to assess the impact of loss of *N*-glycosylation on intracellular levels of key respiratory metabolites. Other conventional microbial assays, tube-based chemotactic responses, biofilm formation, cell morphology and susceptibility to temperature and osmotic stress, were performed to establish how protein-level changes translated into macro-level phenotypes critical to cell function and pathogenicity.

Key findings:

We quantified 1359 proteins corresponding to ~84% of the predicted *C. jejuni* NCTC 11168 genome. Deletion of the *pglB* oligosaccharyltransferase (*ΔpglB*) resulted in a significant change in abundance of 185 proteins, 137 of which were restored to their wild-type levels by reintroduction of *pglB (ΔpglB*::*ΔpglB*). Deletion of *pglB* was associated with significantly reduced abundances of known *pgl* targets and increased levels of stress-related proteins. Commensurately, *pglB* mutants demonstrated reduced survival following temperature (4°C and 46°C) and osmotic (150 mM NaCl) shock, and showed altered biofilm phenotypes compared to wild-type *C. jejuni*. Targeted metabolomics established that glycosylation negative *C. jejuni* were depleted of all respiration-associated proteins that allow the use of alternative electron acceptors under low oxygen conditions, which correlated with protein-level changes to known small-molecule transporters. These data indicate a multi-factorial role for *N*-glycosylation in *C. jejuni* physiology.

45

Proteomic analysis of a small animal model for ricin exposure reveals underlying pathways and responses to toxicity

Graeme C Clark[1], Stuart Armstrong[2], Abbie Harrison[2], Angela Essex-Lopresti[1], Jonathan David[1], Christopher Green[1], Julian A Hiscox[2]

1. Defence Science Technology Laboratory, Porton Down, UK

2. University of Liverpool, Liverpool, MERSEYSIDE, United Kingdom

Publish consent withheld

450

A comprehensive O-glycoprotein repository to facilitate O-glycosylation study

Weiqian Cao[1], Jiangming Huang[1], Mengxi Wu[1], Pengyuan Yang[1]

1. Fudan University, Shanghai, China

There's growing demands to comprehensive and precise analysis of protein O-glycosylation in many fields like cancer studies, biopharmaceutical researches and clinical trials. However, the lack of updated and convenient database made it a big obstacle to refer to and store emerging O-glycoprotein data. To solve this problem, an O-glycoprotein repository named OGP was constructed based on a collection of different sourced O-glycoprotein information including existed databases and public datasets (O-Glycbase 6.00 and Simple-cell dataset), and manually extracted data from literatures published since 1998. A total of 9354 O-glycosylation sites and 11633 site-specific O-glycans mapping to 2133 O-glycoproteins were currently recorded in the repository. OGP is the largest O-glycoprotein repository by far. Based on recorded site data, an O-glycosylation site prediction tool was developed. Moreover, the OGP-backed website was already online (*http://www.oglyp.org/*). There are four functional modules comprised in the website: Statistic Analysis, Database Search, O-Glycosylation Site Prediction and Data Submit. Each modules are specially designed and user-friendly. The first version of OGP and OGP-backed website allow users to obtain various information of O-glycoproteins, such as protein access, sequence, function, site-specific glycan structures, experimental methods, potential glycosylation sites, *etc*. O-glycosylation information mining can be done in one site efficiently, which will greatly facilitate the researchers on study of O-glycosylation.

451

Uncoupling the mechanisms of protein degradation

Tom TC Carew[1], Mary MD Doherty[1], Phil PW Whitfiled[1]

1. UHI, Inverness, INVERNESS, United Kingdom

Uncoupling the mechanisms of protein degradation

Protein degradation is a fundamental process in the cell which is essential to maintain normal function. Intracellular protein degradation is governed by two major pathways. Proteasomal degradation targets specific proteins and allows sensitive and rapid control of cellular protein composition by the ubiquitin:proteasome (UPS) system. Autophagy involves the mass degradation of proteins in the lysosome. It is believed that these systems are compensatory, however, little is known about how each process contributes to the overall stability of individual proteins. There is a clear need to understand which process is dominant during times of normal cell function and when the cell is subjected to stress. Importantly, disturbances of the protein degradative pathways have been associated with the development a number of prevalent diseases including neurodegenerative disorders. The ability to examine these key pathways at the protein level is therefore critical to the understanding of the molecular basis of disease states. In this study the stability of multiple intracellular proteins have been quantitatively assessed under defined conditions in HEK-293 cells. SILAC-based mass spectrometry methods in conjunction with proteome simplification have been employed to explore the degradative pathways. Gene ontology analysis has revealed novel changes in protein expression when

UPS and autophagy are perturbed. These experimental strategies have permitted dissection of the relative control of both pathways and permit further understanding of how protein degradative mechanisms interact and function in cellular systems.

Keywords

Protein degradation, UPS, Autophagy, SILAC

Authors

1Thomas Carew, 2Iain S. Young, Phillip D. Whitfield, Mary K. Doherty

1Division of Biomedical Sciences, University of the Highlands and Islands, Inverness

2School of Medicine, University of Liverpool, Liverpool, UK

452

DCLK1: a novel promoter of gastric cancer progression

Annalisa Carli[1, 2], Shoukat Sterle[1], Alex Azimpour[1], Janson Tse[1, 2], David Greening[2, 3], Pierre Faou[2], Isabelle Lucet[4], Matthias Ernst[1, 2], Michael Buchert[1, 2]

1. Olivia Newton-John Cancer Research Institute, Heidelberg, VIC, Australia

2. La Trobe University, Heidelberg, VIC, Australia

3. The Baker Institute, Melbourne, VIC, Australia

4. The Walter and Eliza Hall Institute of Medical Research, Parkville , VIC, Australia

Doublecortin-like kinase1 (DCLK1), a microtubule associated protein (MAP), has recently gained interest in the cancer research field. Whole-genome sequencing suggests that DCLK1 is a novel tumour driver and DCLK1 overexpression correlates with epithelial-to-mesenchymal transition (EMT) in pancreas, intestine and colon cancer. A recent meta-analysis in gastric cancer (GC) showed that DCLK1 overexpression correlates with advanced and poorly differentiated GC, lymph node metastasis and reduced overall patient survival.

Our analysis of the stomach adenomas (STAD) dataset from the Cancer Genome Atlas (TCGA), showed that DCLK1-high expressing tumours significantly clustered within the genomic stable molecular subtype and the histologically diffuse type. We are currently evaluating DCLK1 expression off 300 stomach cancer patients by immunohistochemistry on tissue microarrays.

We established a DCLK1-overexpressing MKN1 gastric cancer cell-line. The overexpression resulted in increased migration and invasion in vitro and in vivo. These findings support our TCGA-STAD data analysis where high DCLK1 levels correlated with EMT, chemokines, and stromal- and immune cell markers. Strikingly, we observed an overall increase in chemokine secretion when DCLK1 is overexpressed, ex vivo. CXCL12 is the one of the main upregulated chemokines; this is further supported by findings in the TCGA-STAD data set, which shows that DCLK1 and CXCL12 expression levels significantly correlate with each other. Furthermore, a DCLK1-inhibitor reversed migration, invasion and chemokine secretion in the DCLK1-overexpressing MKN1 cells to parental MKN1 cell levels, in vitro and in vivo. This suggests that DCLK1 could be a good target for poor prognosis GCs with high DCLK1 levels.

Thus far, the signalling cascade in which DCLK1 can induce EMT or increased chemokine secretion is poorly understood. Our aim is to answer these questions using SILAC mass spectrometry studies by comparing proteomics, phospho-proteomics and secretomics analysis on parental MKN1 and DCLK1-overexpressing MKN1 cells, with and without DCLK1-inhibitor.

453

Monitoring Physical Healing and Psychological Wellbeing in Paediatric Burns through Proteomic Analysis of Saliva.

Morgan Carlton[1, 2, 3], Tony Parker[1, 2], Joanne Voisey[1], Chamindie Punyadeera[1], Leila Cuttle[1, 2, 3]

1. School of Biomedical Science, Faculty of Health, Queensland University of Technology, Kelvin Grove, QLD, Australia

2. Tissue Repair and Translational Physiology Program, Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD, Australia

3. Centre for Children's Burns and Trauma Research, Center for Children's Health Research, South Brisbane, Queensland, Australia

Burn injuries are extremely traumatic, impacting both physical and psychological wellbeing. Currently, procedures for monitoring the healing process are subjective, and they rely on the expertise level of clinicians. Particularly for children, it is important to have objective, simple methods for monitoring healing. The identification and measurement of biomarkers present in biological fluids have the potential to allow clinicians to diagnose and monitor the healing progression of patients in a quantitative manner. Saliva is a non-invasive bio-fluid that has been underutilised in biomarker research. Therefore, the focus of this research is to identify salivary markers in paediatric burns patients that could be used to monitor both physical and psychological healing progression.

Advanced proteomic methods, such as Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH-MS), will be used to detect and quantify protein abundances in saliva collected from children with burns 6-11 days post-burn. These protein abundances will be correlated with psychological measurements and wound healing data from the patient, and subsequently compared to protein profiles of healthy, non-traumatised children recruited from a community setting who have not been exposed to burns.

This research will enable the identification of a panel of proteins that could be indicative of physical healing status and/or psychological wellbeing state of children with burns. Most importantly, the information identified through this research could lead to better understanding of the biological mechanisms related to healing and the trauma response which may allow the diagnostic utilisation of saliva in burns care and treatment.

454

Role of NCL1 in Cysteine induced toxicity

Rahul Chakraborty[1], Ajay Bhat[1], Khushboo Adlakha[1], Monu Chadar[1], kausik Chakraborty[1], Shantanu Sengupta[1]

1. CSIR- Institute of Genomics and Integrative Biology, New Delhi, NEW DELHI, India

Cysteine, a thiol containing amino acid synthesized via the transsulfuration pathway is a rate limiting precursor for glutathione synthesis and is also utilized for protein synthesis. Among the biological thiols elevated levels of homocysteine is known to be associated with various diseases and is considered to be an independent risk factor for cardiovascular disease. Recent evidences actually support this fact and it has been shown that cysteine induced growth defect in yeast was more severe than homocysteine. Reports also suggest that elevated levels of cysteine may be associated with cardiovascular disease. Using Yeast as a model system in this study we have tried to figure out the mechanism of cysteine induced toxicity. To characterize the cellular response in presence of high level of cysteine we have performed a quantitative proteomics experiment and found several differentially expressed proteins in presence of high level of cysteine which includes aminoacid metabolic proteins, glycolytic – TCA cycle proteins, ribosomal proteins. Even after genomewide mutant screening we came to know that Ncl1 (SAM dependent t-RNA methyl transferase) plays a crucial role in Cysteine toxicity. Δncl1 is much more sensitive towards cysteine and it plays a crucial role in protein biosynthesis and energy metabolism.Intracellular aminoacid measurement by using o-phtalaldehyde reveals that cysteine causes aminoacid imbalance in the cells and by using s35 labeled methionine we also found that cysteine induces translational arrest. Further we have found that supplementation of high levels of leucine and pyruvate can rescue cysteine induced toxicity.

455

An Orbitrap Eclipse Tribrid mass spectrometer with real time search enhances multiplexed proteome coverage and quantitation accuracy.

Aaron M Robitaille[1], Romain Huguet[1], Ryan Bomgarden[2], Derek Bailey[1], Graeme McAlister[1], Arne Kreutzmann[3], Daniel Mourad[3], Woong Kim[1], John C Rogers[2], Daniel Lopez-Ferrer[1], Andreas Huhmer[1], Vlad Zabrouskov[1], Zoon Chan[4]

1. Thermo Fisher Scientific, San Jose, CALIFORNIA, United States

2. Thermo Fisher Scientific, Rockford, IL, USA

3. Thermo Fisher Scientific, Bremen, Germany

4. Thermo Fisher Scientific ANZ, SCORESBY, Victoria, Australia

Quantitative proteomics strategies using Tandem Mass Tags™ (TMT™) enable precise measurement of peptides or proteins from samples multiplexed into a single high-resolution LC/MS experiment. Interference can suppress ratio quantitation and thereby mask true differences in abundance. Here we evaluate if an Orbitrap Eclipse Tribrid mass spectrometer including real time search (RTS), advanced spectral processing algorithms, and modified hardware can enhance TMT quantification accuracy and proteome coverage. Synchronous precursor selection (SPS) based methods provided higher accuracy compared to MS2 methods for TMT quantitation. However, depending on which fragments are selected for MS3 quantitation, accuracy can still be distorted. To improve upon this, we implemented RTS between MS2 and MS3 scans. Using this approach, MS3 scans are only triggered if a peptide is identified from the preceding MS2. This increased the number of peptides identified with RTS by 30%. Secondly, RTS selects fragment ions for MS3 quantitation that are generated from the identified peptide on the fly. Thus, quantitation can be improved to be 95% interference free. The Orbitrap Eclipse Tribrid mass spectrometer has an optimized quadrupole that improves ion transmission, enabling narrower isolation widths to improve TMT quantitation accuracy. Additionally, we evaluated how next generation isobaric mass tags could increase multiplexing capacity on the new instrumentation. Overall, the Orbitrap Eclipse™ Tribrid mass spectrometer includes features such as TurboTMT and Precursor Fit which facilitate intelligent acquisition methods that improve TMT quantitation accuracy, precision, and proteome coverage.

456

AWVP: An automated workflow for variant peptide production

Ian Yi-Feng Chang[1], Chi Yang[1], Shiao-Wen Li[1], Wei-Chao Liao[1], Jau-Song Yu[1], Yu-Sun Chang[1]

1. Molecular Medicine Research Center, Chang Gung University, Kwei-Shan Dist., Tao-Yuan City, Taiwan

To maximize the number of identified novel peptides from proteome data, variations derived from genomic and transcriptomic data can be used to generate variant peptides that are absent from the reference sequences. We developed AWVP, automated workflow for variant peptide production, to generate variant peptides from FASTQs of whole exome sequencing (WES) and RNA sequencing (RNA-Seq) data.

AWVP utilizes GATK4 to produce analysis-ready BAMs from WES and RNA-Seq FASTQs. HaplotypeCaller will be used to identify germline variations from BAM files of normal samples. Mutect2, VarScan2, MuSE, and SomaticSniper are utilized to call somatic variations from normal-tumor matched WES and RNA-Seq BAMs. Customprodbj is used to create SNV variant peptides. To identify additional variant peptides from RNA-Seq data, AWVP uses StringTie to perform transcript assembly from RNA-Seq BAMs with annotation-guided mode and de novo assembly mode. Open reading frames (ORF) will be translated from the assembled transcripts by TransDecoder.LongOrfs. AWVP also uses MiXCR to predict and translate T-cell receptors into CDR3 peptides from RNA-Seq. Besides, AWVP also integrates KNIFE to retrieve back-splicing junctions of circular RNAs (circRNAs) from RNA-Seq data. A six-reading-frame translation tool, sixpack, is used to identify peptides that across back-splicing junctions of circRNAs. Finally, proteins or peptides that are identical to the standard reference proteins will be discarded.

We used AWVP to generate variant peptides for 50 WES and 39 exome capture RNA-Seq data of normal-tumor matched OSCC samples. The preliminary study identified 69,359 germline and 9,389 somatic mutations. We found ~410,000 proteins derived from assembled transcripts. About 2,253 and 2,684 distinct T-cell receptors were identified in normal and tumor samples,

respectively. Finally, we found 74,349 and 78,673 peptides across circRNA back-splicing junctions in normal and tumor samples, respectively.

With AWVP, we can automatically produce comprehensive variant peptides for proteogenomic studies in standardized, repeatable, and reproducible manners.

457

Circadian rhythm impacts the synaptic glycosylation machinery but not the *N*-glycosylation signatures in mice brains

The Huong Chau[1], Edward Moh[1, 2], Lasse Dissing-Olesen[3], Nima Sayyadi[1, 2], Beth Stevens[3], Nicolle Packer[1, 2], Morten Thaysen-Andersen[1]

1. Biomolecular Discovery and Design Research Centre, Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. ARC Centre of Nanoscale BioPhotonics, Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

3. Department of Neurology, F.M. Kirby Neurobiology Center, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

Circadian rhythm, the "body clock", plays central roles in diverse facets of physiology. Sleep is crucial for memory consolidation, a notion supported by the fact that structural and functional changes of the neuronal synapses occur during sleep. Thus, we hypothesised that circadian rhythm regulates the synaptic protein *N*-glycosylation and that alterations of the synaptic *N*-glycome, in turn, may impact the synapse modulation. As a first step to test this hypothesis, the *N*-glycome and the underpinning *N*-glycosylation enzymes of neuronal synaptosomes of mice brains isolated during the light (n = 5, "sleep") and dark phase (n = 4, "awake") were investigated using quantitative LC-MS/MS-based glycomics and proteomics. The synaptosomes were density-separated, their purity verified using synaptic protein markers and their membrane protein extracts obtained. In accordance with previous literature, glycomics identified 56 biosynthetically-related, mostly asialylated structures displaying prominent core fucosylation and GlcNAc- or mannose-capped and Lewis-type epitopes. Surprisingly few sleep-wake cycle *N*-glycome differences were observed within the detailed synaptic *N*-glycome map. Label-free proteomics confidently identified and quantified 193 glyco-enzymes in the synaptosomes of which 42 glycosyltransferases and glycosidases were found to be involved in *N*-glycoprotein biosynthesis. Importantly, circadian regulation of known clock proteins (e.g. BMAL1, PER1) was observed, which verified the experimental design and the proteomics data. Interestingly, approximately half of the *N*-glyco-enzymes including both the catabolic hydrolases and anabolic glycosyltransferases were significantly regulated during sleep, indicating that these are not rate-limiting in the *N*-glycoprotein biosynthesis. In summary, our data indicate that the neuronal protein *N*-glycosylation machinery but not the resulting synaptosomal *N*-glycome fluctuates with circadian rhythm. This study also provides a high-resolution quantitative map of the murine synaptic *N*-glycome useful for future glycobiological explorations.

458

PERCEPTRON - An Open-source GPU Accelerated Proteoform Identification Pipeline for Top-Down Proteomics

Safee Chaudhary[1]

1. Lahore University of Management Sciences (LUMS), Lahore, PUNJAB, Pakistan

Perceptron is a freely available web-based proteoform search and identification platform for top-down proteomics (TDP). Perceptron search pipeline brings together algorithms for intact mass tuning, *de novo* sequencing, identification of chemical and post-translational modifications and candidate protein scoring. High-throughput performance is achieved through the execution of multiple search threads in parallel, on graphical processing units (GPUs) using NVidia Compute Unified Device Architecture (CUDA) framework. The accuracy and performance of the tool have been validated on several TDP datasets and against available TDP software. Specifically, results obtained from searching three published TDP datasets demonstrate that Perceptron outperforms all other tools by at least 71% in terms of reported proteins and 80% in terms of runtime. Together, the proposed software significantly enhances the state-of-the-art in TDP search software and is publically available at http://perceptron.lums.edu.pk. Users can also create in-house deployments of the tool by building code available at the GitHub repository (http://github.com/BIRL/Perceptron).

459

Protein Biomarker Discovery for the identification of Methicillin-sensitive, heterogeneous Vancomycin-Intermediate and Vancomycin-Intermediate *Staphylococcus aureus* Strains by Data-Independent Acquisition Quantitative Proteomics

Yu-Ching Liu[1], Jang-Jin Lu[2], Fuu-Jen Tsai[1], Chao-Jung Chen[1]

1. China Medical University, Taichung, Taiwan

2. Chang Gung University, Taoyuan , Taiwan

Differentiation of Methicillin-sensitive *Staphylococcus aureus* (MSSA), heterogeneous Vancomycin-Intermediate *S. aureus* (hVISA) and Vancomycin-Intermediate *S. aureus* (VISA) is important for proper therapy and timely intervention of outbreaks. In this study, a total of 90 *S. aureus* isolates were enrolled, including 21 MSSA, 25 VSSA, 29 hVISA and 15 VISA. The protein biomarkers for distinguish among MSSA, hVISA and VISA strains were discovered by label-free DIA quantitative proteomics. In the label-free DIA results, 8786 non-redundant peptides, corresponding to 418 different annotated non-redundant proteins were identified. 5 proteins were identified as potential biomarkers with high sensitivity and specificity for the separate identification of MSSA, hVISA and VISA strains.

46

Proteomics of laser-captured microdissected glomeruli and tubulointerstitium reveals extracellular matrix remodelling of kidney allografts with antibody-mediated rejection

Sergi Clotet-Freixas[1, 2], Caitriona McEvoy[1, 2], Ihor Batruch[3], Max Kotlyar[4], Chiara Pastrello[4], Julie Van[1, 2], Yun Niu[4], Andrea Bozovic[5], Vathany Kulasingam[5], Peixuen Chen[6], Sofia Farkona[1, 2], Igor Jurisica[4, 7], Andrzej Chruscinski[2], Rohan John[1, 5], Ana Konvalinka[1, 2, 6]

1. Toronto General Research Institute, University Health Network (UHN), Toronto General Hospital, Toronto, Ontario, Canada

2. Department of Medicine, Division of Nephrology, University Health Network, Toronto, Ontario, Canada

3. Department of Laboratory Medicine and Pathobiology, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

4. Krembil Research Institute, University Health Network, Toronto, Ontario, Canada

5. Laboratory Medicine and Pathobiology, University Health Network, University of Toronto, Toronto, Ontario, Canada

6. Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada

7. Departments of Medical Biophysics and Computer Science, University of Toronto, Toronto, Ontario, Canada

BACKGROUND

Kidney transplantation is the optimal treatment for end-stage kidney disease, but most grafts fail prematurely. Antibody-mediated rejection (AMR) accounts for >50% of graft loss. AMR is caused by antibodies against HLA and non-HLA antigens in two main renal compartments: glomeruli and tubulointerstitium. We hypothesized that compartment-specific proteome alterations may uncover the mechanisms of early antibody-mediated injury.

METHODOLOGIES

We performed laser-capture microdissection to isolate glomeruli and tubulointerstitium from FFPE kidney biopsies, and subjected unfractionated samples to label-free proteome analysis by LC-MS/MS on Q-Exactive-Plus mass spectrometer. Analyses were performed using MaxQuant, Perseus, IID and pathDIP. We compared 7 biopsies with AMR to 23 non-AMR biopsies with cellular rejection or acute tubular necrosis. Primary human glomerular microvascular endothelial cells (HGMEC) were studied *in vitro*.

FINDINGS

We identified 2,026 proteins in glomeruli and 2,399 in tubulointerstitium (FDR=0.01). 120 proteins were differentially expressed in AMR vs. non-AMR glomeruli and 180 in the tubulointerstitium ($p<0.05$). Proteins involved in HLA-mediated antigen presentation were increased in AMR. Basement membrane and cytoskeletal proteins were significantly decreased in AMR. Reduced glomerular protein levels of LAMC1, NPHS1, and PTPRO in AMR were verified by immunostaining. Levels of basement membrane proteins correlated directly and significantly ($R>0.7$; $p<0.05$), suggesting co-regulation in AMR. Protein expression of CCT8 (cytoskeleton dynamics) and CALU (protein folding) correlated with histological features of AMR, namely glomerulitis and peritubular capillaritis ($p=0.017$). Protein-protein interaction and comprehensive pathway analysis of our glomerular protein signature revealed enrichment of inflammatory pathways, such as IL-8 signaling. Stimulation of HGMECs with anti-HLA class I antibody increased the secretion of IL-8 and MCP-1 cytokines($p<0.05$).

CONCLUDING STATEMENT

Basement membranes are often remodeled in late chronic AMR and are the targets of non-HLA antibodies, suggesting that our findings may represent early, important alterations in AMR. Targeting early basement membrane remodeling in AMR may represent a new therapeutic opportunity.

460

WinProphet: a user-friendly pipeline management system for proteomics data analysis based on Trans-Proteomic Pipeline

Ching-Tai Chen[1], Chu-Ling Ko[2], Wai-Kok Choong[1], Jen-Hung Wang[1], Wen-Lian Hsu[1], Ting-Yi Sung[1]

1. Institute of Information Science, Academia Sinica, Taipei, Taiwan

2. Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan

Protein and peptide identifications are essential tasks in proteomics research and involve a series of steps in analyzing tandem mass spectrometry (MS/MS) data. Several software packages have been developed for such purpose, among which Trans-Proteomic Pipeline (TPP) is frequently used and supports various tools for different steps, e.g., database sequence search and statistical validation. TPP provides a web-based user interface named Pentunia to launch most of its functions. Despite its powerful functionality, TPP requires manual intervention to launch each step, including selecting tools and setting parameters. When a large number of MS/MS files need to be processed or the workflow is complicated, it is desirable to be free of manual effort. Thus, a software tool for pipeline creation and execution is in pressing need.

We have developed WinProphet, a tool to create and automatically execute a pipeline for proteomic analyses. It seamlessly integrates with TPP and other external command-line programs, supporting various functionalities including database search for protein and peptide identification, spectral library construction and search, DIA (Data-Independent Acquisition) data analysis, isobaric labeling and label-free quantitation. WinProphet is a standalone and installation-free tool with graphical interfaces for users to configure, manage, and automatically execute pipelines. The constructed pipelines can be exported as XML files with all of the parameter settings for reusability and portability. The executable files, user manual, and sample data sets of WinProphet are freely available at http://ms.iis.sinica.edu.tw/COmics/Software_WinProphet.html.

461

Scalable and Automated Plasma Workflow Based on the Thermo Scientific Q Exactive HF-X MS Platform

Jing Wang[1], Ryan Bomgarden[2], Sergei Snovida[2], Sarah Trusiak[1], Emily Chen[1]

1. Thermo Fisher, Cambridge, MA, United States

2. ThermoFisher Scientific, Rockford, IL, USA

Purpose – Variability and dynamic range of protein abundance substantially influence the human plasma proteome analysis. To develop novel markers indicative of diseases using proteomics-based approaches, the plasma workflow has to be high-throughput and robust for hundreds of runs to make a reliable conclusion out of a clinical study. In this study, we developed a standardized high throughput (HT) plasma proteomics analysis workflow focusing on balancing the depth identification and scalability for sampling large population cohorts. The workflow consists of an automated sample preparation method and an Evosep LC system coupled to an Q Exactive HF-X platform. This poster presents a high throughput serum and plasma proteomics analysis workflow for large population cohorts.

Methods – To reduce analytical variability of plasma sample preparation for LC-MS/MS analysis, we automated Thermo Scientific™ EasyPep™ Mini MS Sample Prep Kit using the Hamilton Microlab STARlet liquid handling system with [MPE]2 positive pressure and evaporation modules. The Evosep LC system was used to run high throughput and automated LC methods. Thermo Scientific™ Q Exactive™ HF-X MS and data-dependent acquisition (DDA) were used to generate quantitative LFQ plasma proteome data. Skyline was used for retention time analysis and Thermo Scientific™ Proteome Discoverer™ 2.3 software was used for database search and post-data analysis.

Results – The automated sample preparation can process 96 samples within 4 hours with ~80% recovery. The throughput of LC analysis for the standardized workflow is > 50 samples per day with 10% overhead to minimize sample carrying over. Around 150 and 200 core proteins (high confident) could be reproducibly identified and quantified for the undepleted serum and the depleted plasma samples, respectively. An example of this workflow applied to small scale depleted plasma lung cancer samples is presented in this poster as well.

462

High-efficiency and low-cost production of pure home-brew Tn5 transposase protein and its application in the development of novel single cell resolution Next Generation Sequencing (NGS) methods.

Gary G Chen, Robert Chen, Gustavo Turecki

One of the major issues in the production of home-brew Tn5 transposes is the availability of lab equipment such as HPLC facilities. Although some methods for producing home-brew Tn5 without the need of such facility have been published, it was reported to be quite difficult to replicate. We have also encountered difficulties while attempting to follow these exact methods. As such, changes to this method have been made to facilitate the process. We have integrated a large-scale production of pure Tn5 into our routine, which is then used for the assembly of Tn5 transposomes. This presentation provides a description of a novel Tn5 based NGS library construction method using a unique Tn5 transposome assembling strategy. We believe it will improve single cell based NGS applications such as single cell Assay for Transposase Accessible Chromatin Sequencing (ATAC-seq) and Whole Genome Bisulfite Sequencing (WGBS).

KEYWORDS: Tn5 transposase; Single cell barcode; ATAC-seq; Next Generation Sequencing (NGS)

463

The Chinese immunopeptidome project and a platform for validation of naturally presented HLA-binding peptides

Lei Chen[1], Xuan Dong[1], Le Zhang[2], Xinxin Lin[2], Zhilong Lin[3], Yan Ren[3], Ying Gu[1], Yong Hou[1]

1. BGI Research, Shenzhen, China

2. Genoimmune, Wuhan, China

3. BGI Genomics, Shenzhen, China

Immunopeptidome is the repertoire of peptides bound to human leukocyte antigen (HLA), the allelic types of which vary from races and populations. With the improvement in mass spectroscopic techniques and bioinformatic analysis, increasing numbers of immunopeptidome database have been established worldwide. Few studies have yet been published focusing on the alleles with high frequency among Chinese, which consists of 20% human population. Last year, our group initiated the Chinese Immunopeptidome Project in the virtue of high throughput sample preparation and mass spectrometry for immunopeptidome profiling. In this project, small amount of peripheral blood was collected from 250 volunteers to generate immortalized B cell lines for HLA-binding peptide identification, as well as genomic and transcriptomic analyses. The database established in the project was also utilized to facilitate machine learning based models for neoantigen prediction and discovery. Meanwhile, we have developed a protocol to validate naturally processed and presented epitopes from clinical specimens with targeted mass spectrometric (parallel reaction monitoring) approach, which serves as the basis to following investigation on affinity and immunogenicity of HLA-binding peptide *in vitro.* Altogether, the database and the protocol will contribute to the current anti-cancer therapeutic interventions and shed light on the implementation of multi-omics study in cancer immunology and system biology.

464

Proteomic and phosphoproteomic networks of Arabidopsis stomatal immune responses

Sixue Chen[1], Qiuying Pang[1], Tong Zhang[1], Wenwen Kong[1]

1. University of Florida, Gainesville, FL, United States

Stomatal openings are a major route of pathogen entry into the plant, and plants have evolved mechanisms to regulate stomatal aperture as innate immune response against bacterial invasion. However, the mechanisms underlying stomatal immunity are not fully understood.

Taking advantage of high-throughput liquid chromatography tandem mass spectrometry (LC-MS/MS) and a TiO_2/ZrO_2-based phosphopeptide enrichment protocol, we performed a label-free proteomic and phosphoproteomic analyses of enriched guard cells in response to a model bacterial pathogen *Pseudomonas syringae* pv. tomato DC3000 .

In total, 495 proteins and 1229 phosphoproteins were identified as differentially regulated. These proteins are involved in a variety of signaling pathways, e.g., hormone and reactive oxygen species signaling. The pathways form interactive molecular networks that regulate stomatal immunity. Based on the motifs of the phosphoproteins, we tentatively identified different protein kinases responsible for the phosphorylation activities.

This study is the first comprehensive investigation of the guard cell phosphoproteome, and has provided new insights into the multifaceted mechanisms of stomatal immunity. The differential phospohoproteins are potential targets for engineering or breeding for enhancing plant defense

465

Phosphoproteomics for Identification of Activated Signalling Pathways in Breast Cancers

Yanfang Chen[1, 2], Trisha Almazi[1, 2], Heather Murray[1, 2], Abdul Mannan[1, 2], Nikita Panicker[1, 2], Matthew D Dun[1, 2], Nicole Verrills[1, 2]

1. School of Biomedical Sciences & Pharmacy, University of Newcastle, Callaghan, NSW, Australia

2. Cancer Research Program, Hunter Medical Research Institute, New Lambton, NSW, Australia

Breast cancer is a major national and global health problem. Outcomes for patients with triple negative or Luminal-B subtype tumours remains poor. Activation of growth and survival signalling pathways, in large part mediated by phosphorylation cascades, is a key feature of all human tumours, including breast cancer. Low gene and proteins level of PPP2R2A, encoding for the B55α regulatory subunit of the ser/thr protein phosphatase, PP2A, is associated with aggressive breast tumours, and predicts for worse outcome in breast cancer patients. Our laboratory has found that molecular inhibition of PPP2R2A in normal mammary epithelial cells results in a tumourigenic phenotype, supporting a role for PPP2R2A as a tumour suppressor in breast cancer. However, the pathways regulated by PPP2R2A are unknown. Comparative phosphoproteomics was performed on BT474 cells (human Luminal-B-like breast cancer), transduced with either control shRNA or shRNA sequences directed toward two different regions of the *PPP2R2A*. Titanium dioxide enriched phosphopeptides were quantified by LC-MS/MS and differential proteins and pathways analysed using Ingenuity Pathway Analysis software. A significant increase in phosphorylation of 56 sites from 48 proteins, and decrease in phosphorylation of 24 sites from 22 proteins were identified in both of the *PPP2R2A* knockdown cell lines. These included YAP1, a transcriptional regulator involved in suppressing apoptotic genes, which was confirmed using phospho-PRM analysis. Follow up studies revealed increased nuclear localization of YAP1 in both knockdown lines, and this was reversed upon re-activation of PP2A. Nuclear localization of YAP1 in human breast tumours has recently been associated with poor outcome, thus functional loss of PPP2R2A may contribute to poor outcome in breast cancer at least in part by dysregulation of YAP1. this study sheds light on the pathways regulated by this important tumour suppressor, and paves the way for novel therapies for breast cancer patients.

466

Study of plant systemic migrating proteins in response to nitrate, phosphate, and potassium deficiency using quantitative proteomics approach

Kai-Ting Fan[1], Byung-Kook Ham[2], Szu-Yu Liu[1], William J. Lucas[3], Yet-Ran Chen[1]

1. Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

2. Global Institute for Food Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

3. Department of Plant Biology, University of California, Davis, California, USA

Primary macronutrients nitrate (N), inorganic phosphate (Pi), and potassium (K) are major elements to build-up fundamental biomolecules, function in energy transfer and the regulation of enzyme activity in plants. In response to the limited macronutrient availability in soil, plants display a high degree of adaptive responses, depending on both local and systemic regulation networks to coordinate nutrient homeostasis within the whole plant. There are several types of molecules such as metabolites, peptides, proteins, and RNAs are known to be involved in the transduction of systemic signaling. However, the roles of the proteins in systemic regulating different types of nutrient deficient stress are still elusive. In this study, the mobile proteins were collected from the phloem sap of heterografted plants under different nutrient stresses and quantitated by the isobaric labeling approach. Among 3200-3700 quantified cucumber and watermelon proteins in the phloem sap of heterografted cucumber/watermelon, 675,

785, and 796 cucumber proteins were found to be systemically translocated under N, Pi, and K deficiency, respectively. Those long-distance translocated cucumber proteins with more than 1.3- or less than the 0.76-fold change in abundance could be involved in carbohydrate metabolic process, molecular biosynthetic process, ubiquitin/protein degradation, and cytokinesis. Interestingly, some potential RNA-binding proteins could be long-distance trafficking in response to N, Pi, or K-deficiency. For instance, one of cucumber protein which is the homolog of one Arabidopsis heterogeneous nuclear ribonucleoprotein showed reduced abundance upon Pi-deficiency in the root-to-scion direction, supporting its potential role as a negative regulator in salt and osmosis stress responses. These mobile proteins with the change in relative quantities in response to N, Pi, or K deficiency should provide a better understanding of the mobile signaling mechanisms and homeostasis for nutrient starvation.

467

Proteogenomics landscape of early-stage lung adenocarcinoma patients identifies stratification hallmarks of early onset and progression

Yi-Ju Chen[1], Theodoros I Roumeliotis[2], Ya-Hsuan Chang[3], Ching-Tai Chen[4], Chia-Li Han[5], Miao-Hsia Lin[1], Ze-Shiang Lin[6], Wei-Hung Chang[7], Ke-Chieh Huang[1], Hao-Chin Yang[1], Pei-Shan Wu[1], Pei-Yi Lin[1], Kuen-Tyng Lin[1], Jen-Hung Wang[4], Gee-Chen Chang[8, 9], Yet-Ran Chen[7], Ana I Robles[10], Henry Rodriguez[10], Jin-Shing Chen[11], Sung-Liang Yu[6], Ting-Yi Sung[4], Jyoti S Choudhary[2], Hsuan-Yu Chen[3], Pan-Chyr Yang[12], Yu-Ju Chen[1]

1. Institute of Chemistry, Academia Sinica, Taipei, Taiwan

2. The Institute of Cancer Research, London, UK

3. Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

4. Institute of Information Science, Academia Sinica, Taipei, Taiwan

5. Master Program in Clinical Pharmacogenomics and Pharmacoproteomics, College of Pharmacy, Taipei Medical University, Taipei, Taiwan

6. School of Medical Technology, National Taiwan University, Taipei, Taiwan

7. Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

8. Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

9. Comprehensive Cancer Center, Taichung Veterans General Hospital, Taichung, Taiwan

10. Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

11. Division of Thoracic Surgery, Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan

12. Center of Genomic Medicine, National Taiwan University, Taipei, Taiwan

Lung cancer in East Asian population is characterized by the large proportion of never smoker and female patients, the predominance of lung adenocarcinoma (LUAD) histological subtype and the high frequency of EGFR mutations . However, the etiology and the molecular mechanisms underlying the disease remain poorly understood making treatment effective only in a small number of cases.

In this study, we report proteogenomic profiles of paired tumor and adjacent normal tissues from ~100 LUAD patients of which 83% were never smoker and 87% were diagnosed at early stage. The genomic landscape confirmed the distinct mutational profile of our cohort compared to previously reported studies. We observed gender-specific differences in driver and passenger mutations, which likely contribute to the disease heterogeneity and the different clinical outcomes. For example, we note the APOBEC signature was predominantly associated with females, especially with early onset or without EGFR activating mutation. The integrated mRNA, protein abundance and site-specific phosphorylation revealed molecular signatures associated with APOBEC mutagenesis which included DNA repair as well as alterations in wide range of functional modules. Proteome subtypes highlight molecular differences that extend the classification beyond the level of clinical staging and genomic driver mutation, these signatures may provide clues on patient outcome and progression. Proteomics landscape also revealed the stage-specific progression signatures characterized by dramatic molecular reorganization at early stage to regulate cancer cell survival, migration and proliferation. Further validation by immunohistochemistry staining in an independent retrospective cohort showed that our candidate extracellular biomarkers were associated with poor overall survival.

In conclusion, this study reveals the distinct genetic profile of the Asian cohort, highlights key molecular signatures associated with early onset in females, and provides a transformative view on the progression of early stage NSCLC.

468

Quantitative N-terminomic and phosphoproteomic analyses of primary neurons and rodent models of neurotoxicity reveal distinct signalling networks governing neuronal death in excitotoxicity

Heung-Chin Cheng[1], Syeda Sadia Ameen[1], Ashfaqul Hoque[1], Joe Ciccotosto[1], Iqbal M Hossain[1], Carli Roulston[2], Laita Bokhari[1], Nicholas Williamson[1], Ching-Seng Ang[1]

1. University of Melbourne, Parkville, Victoria, Australia

2. Florey Neuroscience Institute, Parkville, Victoria, Australia

Background: Excitotoxicity, initiated by over-stimulation of ionotropic glutamate receptors (iGluRs), is a major pathological process directing neuronal death in stroke and neurodegenerative diseases. The over-stimulated iGluRs allow massive influx of calcium ions into the affected neurons, leading to over-activation of neurotoxic enzymes such as the neurotoxic proteases calpains, protein kinases and phosphatases to perturb the structures, expression and phosphorylation of specific neuronal proteins. These perturbed proteins form signalling networks that direct neuronal death in excitotoxicity.

Methods: We used the quantitative “Terminal Amine Isotopic Labelling of Substrates” (TAILS) and phosphoproteomics methods to define the neurotoxic signalling networks governed by the perturbed proteins in cultured primary neurons. Specifically, we aim

to identify the calpain substrates and the neuronal proteins of which the phosphorylation levels are perturbed in primary neurons undergoing excitotoxic cell death.

Results: We identified the cleavage sites in ~300 neuronal proteins proteolytically processed by proteases activated in excitotoxicity. Among them, the tyrosine kinase Src was cleaved by calpains at the unique domain to generate a neurotoxic truncated fragment (ΔNSrc), which acts as a mediator of excitotoxic neuronal death. We also demonstrated that blockade of cleavage of Src to form ΔNSrc could protect against excitotoxic neuronal death *in vivo* in a rat model of neurotoxicity. Additionally, we identified ~6,500 phosphosites in over 4,000 neuronal proteins exhibited dynamic changes in phosphorylation levels in excitotoxicity. Bioinformatic analysis predict over 20 protein kinases as the upstream regulatory kinases directly phosphorylating some of these phosphosites, suggesting that aberrant regulation of some of these kinases are key events directing neuronal death in excitotoxicity.

Conclusion: Results of our proteomic analyses form the conceptual framework for future investigation to define the molecular mechanism governing neuronal death in diseases.

469

Site-Specific Glycoprofiling of Recombinant Human Corticosteroid-Binding Globulin

Anastasia Chernykh[1], Krishnatej Nishtala[1], Sayantani Chatterjee[1], Harry C. Tjondro[1], Hannes Hinneburg[1], Zeynep Sumer-Bayraktar[2], Emily J. Meyer[3, 4], David J. Torpy[3, 4], Morten Thaysen-Andersen[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Charles Perkins Centre, School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

3. Endocrine and Metabolic Unit, Royal Adelaide Hospital, Adelaide, SA, Australia

4. Discipline of Medicine, University of Adelaide, Adelaide, SA, Australia

The liver-derived corticosteroid-binding globulin (CBG) is a heavily *N*-glycosylated transport protein of anti-inflammatory cortisol in human plasma. We have previously demonstrated that the Asn347-glycosylation of the exposed reactive centre loop (RCL) of human CBG impacts the delivery mechanism of cortisol by modulating the neutrophil elastase cleavage rate. However, the structure-function relationship of this complex *N*-glycoprotein remains largely unexplored. Recombinant human CBG variants may be valuable in aiding our understanding of CBG glycobiology. To this end, the glycosylation of newly available HEK293-expressed recombinant human CBG products including polyhistidine- (His-CBG) and biotin-tagged (biotin-CBG) forms were here site-specifically profiled using mass spectrometry-based glycomics and glycopeptide analysis. Glycomics profiling indicated that both the His- and biotin-CBG carry similar, but not identical, *N*-glycosylation despite their common expression system. Complex-type core-fucosylated bi-, tri-, and even higher antennary sialoglycans were identified, glyco-features that are recapitulated in native human CBG. Detailed glycopeptide analyses involving multiple protease digestion strategies, glycopeptide enrichment, and orthogonal dissociation methods used with advanced mass spectrometry detection were employed to site-specifically determine several glyco-features unique to His- and biotin-CBG including the presence of GalNAc(β1,4)GlcNAc (LacdiNAc) and NeuAc(α2,3/6)GalNAc(β1,4)GlcNAc *N*-glycan motifs, and oligomannosidic *N*-glycans with and without mannose-6-phosphate. Interestingly, the glycoprofiling also demonstrated the presence of *O*-glycosylation of both His- and biotin-CBG, a feature not previously reported for native human CBG. Importantly, the two identified *O*-glycans ($Gal_1GalNAc_1NeuAc_2$ and $Gal_2GalNAc_2NeuAc_2$) could be located to the RCL region in proximity to the neutrophil elastase cleavage site and hence may impact cortisol delivery. In conclusion, this site-specific glycoprofiling study has accurately documented the glycosylation of two recombinant human CBG variants, an important prerequisite for future glycobiological and structure-function studies using recombinant human CBG.

47

Identification of key pro-survival proteins in isolated colonic goblet cells of *Winnie*, a murine model of spontaneous colitis

Richard Wilson[1], Rohit Gundamaraju[2], Nuri Gueven[3], Rajaraman Eri[2]

1. Central Science Laboratory, University of Tasmania, Hobart, Tas, Australia

2. School of Health Sciences, University of Tasmania, Launceston, Tasmania, Australia

3. Pharmacy, School of Medicine, Faculty of Health, University of Tasmania, Hobart, TAS, Australia

Background: Accumulating evidence suggests that the goblet cell-derived mucin-2 (Muc2) is a major component of the immune system and that perturbations in Muc2 lead to an ulcerative colitis-like phenotype. The animal model *Winnie* carries a missense mutation in Muc2 that causes Muc2 misfolding, accumulation in goblet cells and ER stress. Excessive ER stress is a hallmark of many diseases, including ulcerative colitis, cancer, diabetes and Parkinson's disease. However, rather than committing to cell death, the typical outcome of unresolved ER stress, *Winnie* goblet cells are characterised by hyper proliferation, suggesting additional regulation of this cellular stress response.

Methods: To elucidate the molecular mechanisms underlying ulcerative colitis in the *Winnie* model we isolated goblet cells from *Winnie* and wild type mice and used label-free quantitative proteomics and bioinformatics to understand the functional consequences of Muc2 misfolding and accumulation.

Results: A large number of changes were identified that highlight a dramatic reprogramming of energy production, including enhanced utilization of butyrate, a key energy source of colonic cells. A major finding was the marked up-regulation of the coiled-coil-helix-coiled-coil-helix domain proteins Chchd2, Chchd3 and Chchd6. In particular, we identified and confirmed the upregulation and nuclear translocation of Chchd2, a protein known to inhibit oxidative stress induced apoptosis.

Conclusions: This study is the first to apply proteome-level analysis to the pre-clinical *Winnie* model of ulcerative colitis. Identification of proteins and pathways affected in isolated *Winnie* goblet cells provides evidence for novel adaptive mechanisms underlying cell survival under conditions of chronic ER stress.

470

Efficient sample preparation of human tears proteomic workflow using S-Trap™

Jimmy Ka-Wai CHEUNG[1], Yin-Hon SZE[1], King-Kit LI[1], Thomas Chuen LAM[1]

1. Laboratory of Experimental Optometry, Centre for Myopia Research, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Tear fluid proteomics offers a non-invasive approach in sample collection for ocular research. The low complexity of tear content has made it an ideal source for detecting and monitoring differential expressed proteins in various ocular conditions. However, the small volume of tears has hampered comprehensive protein analysis. A new S-Trap method for sample preparation offers an efficient protocol for down-stream analysis by mass spectrometry.

Tears collected from healthy HK-Chinese subjects and were pool together and lysed with lysis buffer (10% SDS, 0.1M TEAB). It was then reduced, alkylated and digested on the S-Trap micro column. The digested peptides were analyzed by the Eksigent ekspert NanoLC 400 system coupled with a TripleTOF 6600 system (SCIEX). SWATH window size was adjusted in Analyst 1.7 (SCIEX). Data were exported using ProteinPilot 5.0 and PeakView 2.2 (SCIEX). Functional classification was analyzed using PANTHER classification system.

A much shorter sample preparation time (within a day) compared to typical in-house in-solution MS preparation while having a high peptide recovery yield (~81%) from S-Trap method. A total of 271 non-redundant proteins (4056 distinct peptides) were identified from a 3 ug injection at 1% Global FDR (increase of 20% in unique proteins compared to 2 µg injection). A cut-off of 0.4 (LOG2 fold change) was determined for detection of differential protein expression from SWATH quantitation experiment.

We have established an efficient proteomic workflow for human tears- having a relatively short sample preparation time and high recovery yield for MS acquisition. Higher injection amount had a higher impact on the tear protein ID identified. The fold change of differential expression cut off filter was determined basing on the results of SWATH quantitation, which will be used for our future quantitative setup relating to human tears for a more rapid quantitation of tears proteomic in ocular research.

471

Profiling the membrane and secretory proteomes to discover the signaling drives the malignancy of oral cavity cancers

Lang-Ming Chi[1]

1. Chang Gung Memorial Hospital, Guishan Dist.,, TAOYUAN CITY 333, Taiwan

Background: The main reasons of poor prognosis of oral cavity cancer are most lesions diagnosed in advanced stage, secondary cancer occurrence, local recurrence, and/or metastasis. Currently, the key molecules for oral squamous cell carcinoma to associate with cancer progression and as specific markers to support treatment in clinic remain unavailable.

Methodologies: Isobaric-tagging and SILAC-assisted quantitative proteomics were implemented to determine the dys-regulated protein and phosphoprotein candidates in conditional media and membrane parts of oral cancer cell lines with high and low metastatic ability.

Findings: Eighty five dys-regulated secretory protein candidates were discovered from the 3616 proteins identified in conditional media of the cultured oral cancer cells with various metastatic abilities. Using the enriched membrane compartments, 52 dys-regulated protein candidates were found from the 4111 quantified proteins. Three secretory protein candidates were further confirmed to increase in the metastatic cancer cell lines by Western blot. One of membrane protein candidate, claudin-1, was confirmed to be overexpressed in oral cancer cell lines with nodal metastatic ability in mouse model.

Concluding: Using the established oral cancer cell lines with high and low metastatic ability to explore the protein candidates and signaling drive the malignancy of oral cancer cells. Eighty five dys-regulated secretory proteins and 52 dys-regulated protein candidates in membrane parts were discovered. The overexpression of claudin-1 in oral squamous cell carcinoma has been reported by several research teams, which support the confidence of our finding. Currently, the abundance of the discovered secretory and membrane candidates in tissues and bio-fluid samples from oral cancer patients are under evaluated by immunohistochemical staining, ELISA and MRM MS assays. The roles of those candidates in tumor development will be examined in animal models to find out the molecules play a key role in metastasis of oral cavity cancers.

472

MUC1 glycopeptides by chemoenzymatic synthesis revealed distinct specificities of anti-MUC1 antibodies

Yasunori Chiba[1], Yayoi Yoshimura[1, 2], Kaori Denda-Nagai[3], Yoshie Takahashi[1], Izuru Nagashima[1], Hiroki Shimizu[1], Toshimitsu Kishimoto[2], Miki Noji[3], Tatsuro Irimura[3]

1. National Institute of Advanced Industrial Science and Technology, Tsukuba, IBARAKI, Japan

2. Japan Bioindustry Association, Tokyo, Japan

3. Graduate School of Medicine, Juntendo University, Tokyo, Japan

Mucin 1 (MUC1) has been known as a carcinoma-associated mucin-like glycoprotein and anti-MUC1 antibodies have been used as serum markers for tumor burden to monitor therapeutic outcomes of breast cancer and as a serum marker for lung fibrosis. Though some of the anti-MUC1 antibodies were already investigated, the specificities of many anti-MUC1 monoclonal antibodies, particularly their specificities toward glycoforms at different positions remain unclear. In this study, we produced a total of 20 glycopeptides representing a unit of the tandem repeat domain of MUC1 with GalNAc (Tn-antigen), Galβ1-3GalNAc (T-antigen), NeuAcα2-3Galβ1-3GalNAc (sialyl T-antigen), or NeuAcα2-6GalNAc (sialyl Tn-antigen) at each threonine or serine residue by chemoenzymatic synthesis to analyze their capacity to bind 13 monoclonal antibodies specific for MUC1. First, glycopeptides with a GalNAc residue at each one of the five possible *O*-glycosylation sites were chemically synthesized and then extended with Gal and/or NeuAc residues using glycosyltransferases with donor sugar nucleotides. In the case of synthesis of NeuAcα2-6GalNAc attached to the serine at position 9 or 19 of the peptide, Galβ1-3(NeuAcα2-6)GalNAc was prepared and then β1-3,4 galactosidase was used to remove galactose. Next, the binding capacity of 13 antibodies for the glycopeptides was analyzed by Enzyme-Linked ImmunoSorbent Assay. The results indicated that anti-MUC1 monoclonal antibodies can be classified into several subgroups based on their specificity toward different glycopeptides having specific glycan structures and glycosylation sites. Considering that some of these anti-MUC1 antibodies were developed to aid cancer diagnosis and therapy, the present results are potentially useful as the basis for further clinical use of various anti-MUC1 antibodies.

473

Quantitative iTRAQ-based proteomic analysis of exogenous cytokinin induced callus and protocorm-like bodies (PLBs) formation from the root tips of *Vanilla planifolia*

Chiew Foan Chin[1], Li Chin Chai[1], Qingsong Lin[2], Nadiatul Akmar Abd Aziz[1]

1. University of Nottingham Malaysia Campus, Semenyih, SELANGOR, Malaysia

2. Biological Sciences, National University of Singapore, Singapore

Natural vanilla is a popular flavouring essence which is in high demand globally. In order to meet the demand, superior vanilla plant varieties are needed for high quality vanilla pods production. Clonal propagation is a viable means of producing high yielding vanilla plantlets. However, the production of vanilla plantlets through the tissue culture route is still very low. We have developed an efficient regeneration protocol by using cytokinin as the sole plant growth regulator to clonal propagate the plantlets from the root tips of *Vanilla planifolia* (*V. planifolia*) Andrews. The key steps involved in the regeneration protocol involve induction of callus from the excised root tips followed by the conversion to PLBs before eventual formation of shoot primordia in the MS media supplemented with 1 mg/L BAP. We use isobaric tags for relative and absolute quantitation (iTRAQ) proteomic approach to investigate the underlying molecular and cellular mechanisms involved in the induction of callus (VR-VC) and PLBs conversion (VC-VP). A total of 595 and 1237 unique proteins with abundance changes of > 1.5 or < 0.67 - fold were identified in VR-VC and VC-VP respectively. The results indicated that for the dedifferentiation of root tip into callus, the GO terms enriched in upregulated proteins were involved in carbohydrate metabolic process, oxidation-reduction and chloroplast while the down regulated proteins were associated with protein transport, protein binding and gene expression. For the differentiation of callus to protocorm, the up-regulated proteins involve thylakoid membrane, plastids and photosystem I while the down regulated proteins involve RNA binding and metabolic process. The findings provide insights into the steps involved in the tissue culture of vanilla orchids which will be useful in devising effective strategies for the micropropagation and breeding programme of *V. planifolia*.

474

ImCal: Internal calibration curves for accurate quantitation in clinical proteomics.

Cristina Chiva[1, 2], Olga Pastor[1, 2], Eduard Sabido[1, 2]

1. Proteomics Unit, Centre de Regulació Genòmica (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, España

2. Universitat Pompeu Fabra, Barcelona, España

Targeted proteomics has become the method of choice for biomarker validation in human biopsies due to its high sensitivity, reproducibility, accuracy, and precision. However, for targeted proteomics to be transferred to clinical routine there is the need to reduce its complexity, make its procedures simpler, increase its throughput, and improve its analytical performance. Here we present the Isotopologue Multipoint Calibration (ImCal) quantification strategy, which uses a mix of isotopologue peptides to generate internal multipoint calibration curves to accurately quantify biomarker peptides in clinical applications without the need of expert supervision.

The performance of ImCal was demonstrated quantifying 2 peptides from HER2 in standard and in clinical samples. Five different isotopically-labeled internal standard were synthesized per each peptide of interest. The peptides were isotopologues, with the same sequence but different nominal mass. All samples were analysed in a triple quadrupole in SRM mode.

HER2 protein was quantified from FFPE tissue samples from breast cancer tumours. The isotopologue standard peptides were spiked in the samples at different amounts to generate an internal calibration curve per each endogenous peptide covering 2 orders of magnitude. To determine the accuracy and precision of the internal calibration method in peptide quantification, the method was compared with the standard methods using known amounts of the unlabelled peptide. Results show that ImCal renders high accuracy and reproducibility in peptide quantification, and it is suitable to generate ready-to-use quantification kits for selected biomarkers of clinical relevance.

The use of ImCal overcomes several of the current limitations for peptide biomarker quantification in clinical proteomics applications. It is compatible with both high- and low-resolution mass spectrometers as well as with different levels of endogenous peptide. It eliminates the need for blank matrixes, and it allows the evaluation of matrix effects and definition of a valid quantification range in each individual sample.

475

MetaboKit: a comprehensive data processing workflow for DDA and DIA-MS data in untargeted metabolomics

Hyungwon Choi[1], Guoshou Teo[1], Pradeep Narayanaswamy[2], Stephen Tate[3]

1. National University of Singapore, Singapore, NA

2. SCIEX, Singapore

3. SCIEX, Toronto, Ontario, Canada

Compound identification and quantification in untargeted metabolomics has long depended on mass fingerprinting of precursor ions and verification by retention time (RT) matching of standards, forgoing systematic fragmentation of precursor ions and MS/MS scans. This practice was influenced by several major factors: (i) RT is a crucial feature for identification of a compound, but it varies by column types; (ii) MS/MS spectra are variable across instrumentation parameters such as collision energy and the spectra contain only a few fragment ions in many small molecules; and (iii) there often exist isomers with similar RT, adding to the ambiguity that cannot be resolved by MS/MS. All these factors contribute to the difficulty in establishing a universally applicable compound identification workflow.

To address some of these challenges, we developed a software package MetaboKit that embodies comprehensive data extraction workflows for DDA and DIA-MS data processing in metabolomics. MetaboKit has special emphasis on active utilization of MS/MS for both identification and quantification tasks. The first part of MetaboKit performs MS/MS-based compound identification and MS1-based quantification for DDA data, which automatically constructs a library of MS/MS spectra validated by external spectral libraries, each element annotated with in-house retention time for a given chromatographic column. Using the library of spectra, the second part of the tool performs semi- and fully-targeted quantification of precursors and fragments from DIA data. Throughout this process, we account for user-specified adducts and in-source fragments in both identification and quantification steps.

To evaluate the performance, we generated a mixture of 90 metabolite standards and analysed the sample with LC-MS/MS in IDA and DIA mode of scans. We show that MetaboKit has outstanding specificity in the identification and great quantitative accuracy, a feature not demonstrated by other bioinformatics workflows that do not systematically incorporate MS/MS scans in the compound identification.

476

Social, economic and demographic correlates of drug and food consumption identified by wastewater-based epidemiology

Phil M Choi[1], Ben Tscharke[1], Saer Samanipour[2], Wayne D Hall[1], Coral E Gartner[1], Jochen F Mueller[1], Kevin V Thomas[1], Jake W O'Brien[1]

1. University of Queensland, Brisbane, QUEENSLAND, Australia

2. Norwegian Institute for Water Research, Oslo, Norway

Publish consent withheld

477

Proteogenomic approach to UTR peptides identification

Seunghyuk Choi[1], Shinyeong Ju[2, 3], Jinwon Lee[1], Seungjin Na[1], Cheolju Lee[2, 4, 5], Eunok Paek[1]

1. Department of Computer Science, Hanyang University, Seoul, Republic of Korea

2. Center for Theragnosis, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea

3. Department of Bio-Medical Science and Research Institute for Natural Sciences, Hanyang university, Seoul, Republic of Korea

4. Division of Bio-Medical Science and Technology, KIST School, Korea University of Science and Technology, Seoul, Republic of Korea

5. Department of Converging Science and Technology, KHU-KIST, Kyung Hee University, Seoul, Republic of Korea

There have been many reports showing translation of 5' untranslated regions (5'-UTRs) and 3' untranslated regions (3'-UTRs), mostly identified by ribosome profiling and/or tandem mass spectrometry (MS/MS) based proteomics. We propose a proteogenomic approach to identify UTR peptides from a MS/MS assay. Firstly, we construct a translated UTR peptide database with an assumption that UTR may be translated due to single nucleotide errors in recognizing START or STOP codon. After that, we apply a multi-stage search strategy (Madar et al., 2018), which is a method of rigorously identifying novel peptides. As a result, we identified 52 5'-UTR peptides and 9 3'-UTR peptides from a H1299 cell line dataset. There was a total of 45 and 9 genes corresponding the 5'-UTR and 3'-UTR peptides, respectively. Almost a half of 45 genes were commonly observed in a previous study. We further decided alternative start codon of 5'-UTR peptides based on codon frequencies, and then estimated the strength of its kozak context. We classified contexts into strong/weak/non-kozak classes (Lee et al., 2012). The kozak class composition of novel translation initiation sites (TISs) was comparable to that of the annotated translation initiation sites. As for read-through (RT) events at 3' end of coding region, we identified a translation of 3'-UTR of MDH1 gene, which is consistent with a previous report (Stiebler et al., 2014). Furthermore, we could identify that the stop codon was substituted to tryptophan, which was not detected by Ribo-Seq. Finally, we also validated expression of 29 UTR peptides by MS/MS analysis of synthetic peptides. Among

them, 28 UTR peptides were verified. Peptide identification using tUTR DB together with multi-stage strategy could rigorously identify UTR peptides translated due to single nucleotide errors.

1. Madar, I.H.; Lee, W.; Wang, X.; Ko, S.-I.; Kim, H.; Mun, D.-G.; Zhang, B.; Paek, E.; Lee, S.-W. Comprehensive and sensitive proteogenomics data analysis strategy based on complementary multi-stage database search. International Journal of Mass Spectrometry 2018, 427, 11–19.
2. Lee, S.; Liu, B.; Lee, S.; Huang, S.-X.; Shen, B.; Qian, S.-B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc. Natl. Acad. Sci. U. S. A. 2012, 109, E2424–32.
3. Stiebler, A.C.; Freitag, J.; Schink, K.O.; Stehlik, T.; Tillmann, B.A.M.; Ast, J.; Bölker, M. Ribosomal Readthrough at a Short UGA Stop Codon Context Triggers Dual Localization of Metabolic Enzymes in Fungi and Animals. PLoS Genetics 2014, 10, e1004685.

478

Subcellular mapping of protein re-localisation in response to ionising radiation: A new approach to understanding DDR

Josie Christopher[1], Mercedes Vazquez-Chantada[2], Derek Barratt[2], Hilary Lewis[3], Kathryn Lilley[1]

1. Cambridge Centre for Proteomics (CCP), University of Cambridge, Cambridge, UK

2. Discovery Science , AstraZeneca, Cambridge , UK

3. Oncology Biosciences , AstraZeneca, Cambridge , UK

Genomic instability is an enabling hallmark of cancer, often triggered by abnormal expression and behaviour of DNA damage repair (DDR) proteins [1] High levels of mutations and aberrant posttranslational modifications (PTMs) can lead to changes in interacting partners and/or subcellular location, affecting the efficiency of DDR and causing subsequent mutagenesis [2,3]. These deficiencies in DDR can sensitise cancer cells to DNA damaging agents, such as chemotherapeutics and radiotherapy. Despite this, cancer cells use alternative DDR pathways to bypass cell death and become resistant to such therapeutics. Therefore, contributing genes and proteins to these reliant pathways have become prime candidates for oncology drug design and diagnostics. The prime examples are PARP1 and BRCA1/2, , targeted by the therapeutic inhibitor olaparib and diagnostic markers for breast and ovarian cancers, respectively [4].

Ionising radiation (IR) is detrimental for DNA and causes prolific cellular changes, such as cell morphology and activation of multiple signalling cascades. Whether response to IR is pro or antiapoptotic, trafficking between organelles is required for a proportion of proteins to function, as it is well documented that some proteins have localisationdependent roles and 'moonlight' between subcellular compartments [3,5]. Previous work studying relocalisation have been performed using lowthroughput techniques, which rely on effective fluorescent antibodies or GFPtagging [6]. To investigate this trafficking, we applied a holistic and novel spatial proteomics workflow, LOPITDC. This technique uses the combination of differential ultracentrifugation for organelle separation, multiplexed quantitative mass spectrometry and machine learning to produce subcellular proteome maps [7]. Currently, we have successfully produced extremely reproducible, high-resolution spatial proteome maps of dynamic protein re-localisation in a lung carcinoma cell line (A549) in response to ionising radiation. We hope this will give us a deeper understanding of these proteins' roles and behaviour to this stimulus and to aid DDR-targeting drug design.

1. Hanahan D, Weinberg RA (2011) "Hallmarks of cancer: the next generation". Cell 144: 646–674.
2. Alshareeda AT, Negm OH, Green AR, Nolan CC, Tighe P, Albarakati N, Sultana R, Madhusudan S, Ellis IO, Rakha EA (2015) "KPNA2 is a nuclear export protein that contributes to aberrant localisation of key proteins and poor prognosis of breast cancer". Br J Cancer 112: 1929–1937.
3. Li R, Hannon GJ, Beach D, Stillman B (1996) "Subcellular distribution of p21 and PCNA in normal and repair-deficient cells following DNA damage". Curr Biol 6: 189–199.
4. O'Connor MJ (2015) "Targeting the DNA Damage Response in Cancer". Mol Cell 60: 547–560.
5. Boisvert F-O-M, Lamond AI (2010) "p53-Dependent subcellular proteome localization following DNA damage". Proteomics 10: 4087–4097.
6. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, et al. (2017) "A subcellular map of the human proteome". Science (80-) 356: eaal3321.
7. Geladaki A, Kočevar Britovšek N, Breckels LM, Smith TS, Vennard OL, Mulvey CM, Crook OM, Gatto L, Lilley KS (2019) "Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics". Nat Commun 10: 331.

479

Proteomics investigation of lesion-dependent salivary proteins for OSCC monitoring

Lichieh Julie Chu[1, 2], Yu-Tsun Lin[1], Wei-Fan Chiang[3, 4], Jau-Song Yu[1, 2, 5]

1. Molecular Medicine Research Center, Chang Gung University, Tao-Yuan, NA, Taiwan

2. Liver Research Center, Chang Gung Memorial Hospital at Linkou, Tao-Yuan, NA, Taiwan

3. Department of Oral & Maxillofacial Surgery, Chi-Mei Medical Center, Liouying, Tainan, NA, Taiwan

4. Department of Dentistry, Chi-Mei Medical Center, Liouying, Tainan, NA, Taiwan

5. Department of Cell and Molecular Biology, Chang Gung University, Tao-Yuan, NA, Taiwan

Oral squamous cell carcinoma (OSCC) is one of the most malignant diseases with high prevalence in Taiwan, frequently results in poor prognosis and mortality because of delay diagnosis, tumor recurrence or metastasis. Therefore, early detection and disease monitoring are very important for management in OSCC. Understanding of the systematic fluctuation of salivary proteins caused by tumor lesion formation helps to clarify mechanisms for tumorigenesis and monitor OSCC occurrence. For this propose, patients of OSCC and oral potentially malignant disorders (OPMDs) were recruited to collect their saliva samples in several time

points, including within 15 days before treatment, every 3~6 months after treatment, within 15 days before second treatment while OSCC recurrence or malignant transformed, every 3~6 months after second treatment. The quantitative profiling of these longitudinal samples from the same patient were analyzed using TMT protein labelling and 2D-LC-MS/MS. Among 2088 quantified salivary proteins, 54 and 37 were up- and down-regulated between first and second treatments, respectively, and 38 and 8 had trends of up- and down-regulated between before and after treatment consistently. Most of the up-regulated proteins were involved in phosphorylation, acetylation, and cell-cell adhesion; and the down-regulated proteins were involved in exosome and intermediate filament instead. They might be useful as parameters for monitoring malignant levels of oral lesions.

480

Photothermal therapy and their mechanism of Gold nanostar in breast cancer

Joo-Hee Park [1, 2], Ji-Young Yoon [1], Eun-Hee Han[1], Young-Ho Chung[1, 2]

1. Korea Basic Science Institute, Cheongju-si, South Korea

2. Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon, South Korea

Gold nanoparticles are being actively studied for photothermal therapy (PTT)-based treatment of cancers due to localized surface plasmon resonance (LSPR) effect and excellent photothermal conversion ability. In particular, Gold nanostars has a higher surface area than other gold nanoparticles, and it is highly photothermal efficiency and easy to bio-conjugation. In this study, we investigated the anticancer effect and their cell death mechanism by gold nanostar (GNS) conjugated with hyaluronic acid (HA) targeting CD44 receptor that is overexpressed in breast cancer. The characterization of GNS was confirmed by TEM, UV-visible Spectroscopy. Cell uptake test of GNS was confirmed by ICP-MS. In the cell viability test, photothermal therapy with HA-GNS showed specific cell death in breast cancer cell compared with normal cells. Flow cytometry and western blot analysis revealed that HA-GNS induced apoptosis and necroptosis in breast cancer cells. And proteomic experiments were carried out to elucidate the mechanism of necroptosis. Necroptosis is known to be a programed cell death process that induces the activation of immune cells by causing an inflammatory reaction in the tumor microenvironment. Therefore, HA-GNS photothermal therapy is proposed to induce not only the death of cancer cells but also the activation of immune cells in tumor microenvironment.

481

New insights into the function and regulation of the PAQosome chaperone complex

Philippe Cloutier[1], Christian Poitras[1], Marie-Soleil Gauthier[1], Benoit Coulombe[1]

1. Institut de recherches cliniques de Montréal (IRCM), Montreal, QUEBEC, Canada

Since its discovery a little over a decade ago, the PAQosome (also known as R2TP/PFDL) has emerged as a critical organizer in the biogenesis of several protein complexes and networks such as protein assemblies involved in transcription, mRNA maturation, translation and nutrient-sensitive signalling pathways. Unsurprisingly, evidence is mounting that this chaperone complex may be involved in tumorigenesis, consistent with a role in regulating proliferation. We present here our most recent results that led to the identification of new client protein complexes and novel post-translational modification (PTM)-based modes of regulation of the PAQosome. Similar to what has recently been reported for axonemal dynein complexes involved in cilium motility; we now identify cytoplasmic dynein complexes that are responsible for cargo transport along microtubules as new clients of the PAQosome. We have also identified a phosphorylation-dependent association of the PAQosome subunit RPAP3 with preribosome complexes. Additionally, we report for the first time the identification of a small ORF-encoded PAQosome subunit, along with a possible role in the regulation of downstream gene, asparagine synthetase (ASNS) whose expression is linked to neurological disorders and response to asparaginase, a chemotherapeutic drug used in the treatment of acute lymphoblastic leukemia (ALL). These results define novel aspects of PAQosome function and regulation, some being associated with human diseases.

482

The European Bioinformatics Community (EuBIC): who are we and what do we do?

European Bioinformatics Community (EuBIC)[1]

1. EuBIC, ,

Nowadays, mass spectrometry-based proteomics is a method of choice for studying protein quantities, structures, interactions and modifications in high-throughput setups. For these reasons, it is widely used in academia and industry performing biological, biomedical and clinical research. This is a highly interdisciplinary field that evolves at a fast pace, and for this reason, it strongly relies on experts in informatics and statistics. However, the community of bioinformaticians working on mass spectrometry data acquisition and signal processing, or integration and statistical analysis of mass spectrometry data remains scarce.

The European Bioinformatics Community (EuBIC) is an initiative by the European Proteomics Association (EuPA) - a federation of 21 European national proteomics societies including Denmark. It aims to bring together the bioinformatics community with a focus on mass spectrometry-based proteomics and improve bioinformatics through the setup of community-driven dynamics. EuBIC forms an accessible contact point for any proteomics bioinformatics questions by providing educational material online (at https://www.proteomics-academy.org/ and through a mailing list for direct support) and by organizing bioinformatics hubs at several conferences (such as HUPO). Furthermore, EuBIC organizes a dedicated yearly conference to bring together and educate researchers in computational proteomics, alternating between a general conference and a developer's meeting.

In 2020, the EuBIC developer's meeting will take place in Nyborg, Denmark (13 - 17 January 2020). Organized by and for proteomics bioinformatics researchers, it will provide a common meeting base for European scientists active in the field of bioinformatics and computational proteomics. Young scientists are explicitly encouraged to join and participate in the meeting.

483

Egg White as a Quality Control in Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry Imaging (MALDI-MSI)

Mark Condina[1], Parul Mittal[2], Matthew T Briggs[1], Manuela Klingler-Hoffmann[1], Peter Hoffmann[1]

1. Future Industries Institute, The University of South Australia, Mawson Lakes, SA, Australia

2. The University of Adelaide, Adelaide, SA, Australia

The strength of matrix-assisted laser desorption/ionisation-mass spectrometry imaging (MALDI-MSI) is to analyse and visualise spatial intensities of molecular features from an intact tissue. The distribution of the intensities can then be visualised within a single tissue section or compared in between sections, acquired consecutively. This method can be reliably used to reveal physiological structures and has the potential to identify molecular details, which correlate with biological outcomes. MALDI-MSI implementation in clinical laboratories requires the ability to ensure method qualification and validation to meet diagnostic expectations. To be able to get consistent qualitative and quantitative results, standardised sample preparation and data acquisition are of highest priority. We have previously shown that the deposition of internal calibrates onto the tissue section during sample preparation can be used to improve mass accuracy of monitored *m/z* features across the sample. Here, we present the use of external and internal controls for quality check of sample preparation and data acquisition, which is particularly relevant when either a large number of spectra are acquired during a single MALDI-MSI experiment or data from independent experiments are processed together. To monitor detector performance and sample preparation, we use egg white as an external control for peptide and *N*-glycan MALDI-MSI throughout the experiment. We have also identified endogenous peptides from cytoskeletal proteins, which can be reliably monitored in gynaecological tissue samples. Lastly, we summarize our standard quality control workflow designed to produce reliable and comparable MALDI-MSI data from single sections and tissue microarrays (TMAs).

484

Functional prediction of uPE1 proteins using a "*guilt-by-association*" approach with the CCLE dataset

Guillermo Serrano[1], Elizabeth Guruceaga[1, 2], José González-Gomariz[2], Fernando J Corrales[3], Victor Segura[1, 2]

1. Platform of Bioinformatics, Cima Universidad de Navarra, Pamplona, I am not in the U.S. or Canada, España

2. Platform of Bioinformatics, IdiSNA, Pamplona, Spain

3. Proteomics Unit, Centro Nacional de Biotecnología (CSIC), Madrid, Spain

In the last years, the Chromosome-Centric Human Proteome Project (C-HPP) has focused much of its efforts on the detection of those proteins without experimental evidence (missing proteins) using mass spectrometry-based technologies. However, new goals have now been stablished in order to improve the understanding of the protein roles in the normal cellular processes and human diseases. This is the case of uncharacterized proteins with experimental evidence (uPE1s) because these proteins lack a known function in the cell and need to be studied in detail. One of the most popular methods to predict molecular functions is the "guilt-by-association" approach, especially in the field of transcriptomics where there are a huge number of publicly available experiments. In particular, we used the CCLE dataset that contains more than 1000 RNA-Seq experiments of human cell lines to calculate the correlation of the uPE1 genes with the PE1 genes with known functions. Next, we performed a sample level functional enrichment analysis (SLEA) based on these correlations and the pathways and functions annotated in GO and KEGG. The result consisted in the set of enriched functions and pathways in the PE1 gene sets positively or negatively correlated with each uPE1 gene. This information was represented as a network to infer the functional characterization of the uPE1 proteins in an integrated view: common functions to all of them and specific functions of one or a group of uPE1s. Subsequently, these bioinformatic predictions can be used by the C-HPP research teams as a guidance to design the biological experiments needed to validate novel functions of uncharacterized proteins.

485

Plasma levels of metabolites differentiate first episode psychosis in schizophrenia and bipolar disorder patients

Alana Costa[1], Helena Joaquim[1], Leda Talib[1], Wagner Gattaz[1]

1. Laboratory of Neuroscience (LIM-27), São Paulo, SP, Brazil

Schizophrenia (SCZ) and bipolar disorder (BD) are serious psychiatric disorders and share several characteristics and the diagnosis yet is mainly clinical. The sooner they are identified, diagnosed and treated, the better the clinical prognosis. Therefore, the development of sensitive and accurate biomarkers is highly required. Lipids play an increasingly recognized role in the neuronal function and plasticity of the brain. Glycerophospholipids and molecules-like comprise 60% of the non-aqueous portion of the brain and in an even greater proportion of the dendrites and synapses. Other metabolites directly influence its functioning and remodeling, such as acylcarnitines, sphingolipids, cholesterol and other lipids. Since lipid metabolism is altered differently in

neuropsychiatric diseases, alterations in the lipid profile of the membrane can allow a discrimination between subjects in first-episode psychosis). Thus, our aim was to determine plasma levels of metabolites of subjects in FEP and controls and find cutoff values that differentiate each group. Plasma samples were analyzed for 55 drug-naïve patients (28 SCZ and 27 BD) and 30 controls. Determining the lipid profile was performed by mass spectrometry - Flow injection analysis, using AbsoluteIDQ p180® kit (Biocrates Life Sciences). Statistical analyzes were performed using a classification method - Classification And Regression Tree. We observed that the combination of four metabolites are able to differentiate the diagnoses: PCaaC26:0, PCaaC38:4, PCaaC34:3 and C16-OH. The accuracy of the method is 87,1%. Discussion: Our results suggest that the levels of some plasma metabolites differentiate subjects with SCZ, BD and controls. The levels of these metabolites can be a potential biomarker for psychosis, as well as a diagnostic marker for SCZ and BD. The findings from this study require further validation in BD and SCZ subjects, but suggest that the metabolome is a good tool to understand the pathophysiology of these disorders and presents potential diagnostic biomarkers.

486

Genomic deletion of *siat8* induces mesenchymal-to-epithelial transition in ovarian cancer cells

Cécile Cumin[1], Yen-Lin Huang[1], Charlotte Rossdam[2], Natalie Rimmer[1], Ching-Yeu Liang[1], Monica Lopez[1], André Fedier[1], Viola Heinzelmann-Schwarz[3], Falk Büttner[2], Francis Jacob[1]

1. University of Basel - University Hospital Basel, Basel, BASEL STADT, Switzerland

2. Department of Biochemistry, Institute of Clinical Biochemistry, Hannover, Germany

3. Department of Gynecology and Gynecological Oncology , Hospital for Women, Univeristy Hopsital of Basel , Basel Stadt , Basel , Switzerland

Ovarian cancer cells need to undergo reversible epithelial-to-mesenchymal transition (EMT/MET) to invade the surrounding organs of the peritoneal cavity. During this transition, cancer cells also change their glycosphingolipids (GSLs) repertoire. Previous studies in our laboratory have demonstrated that depletion of globosides (Δ*A4GALT*) induce EMT and increase gangliosides expression in ovarian cancer cells [1,2]. However, role of GSLs in MET promoting formation of primary tumor and metastic growth are still uncertain. Thus, we generated ovarian cancer cell lines homozygously deleted for *SIAT8* using the CRISPR-Cas9 technology.To the GSL content in our knockout cell lines, we used the capillary gel electrophoresis coupled to laser-induced fluorescence detection (xCGE-LIF) [3]. Here, the mean migration time (MTU) of 13.61 corresponding to the ganglioside GT3 was reduce in Δ*SIAT8* cells. Interestingly, depletion of Δ*SIAT8* cells revealed an increase of globosides Gb3 (MTU: 80.56) compared with the wild-type.Trough in vitroand in vivoassays, we already observed that deletion of SIAT8 reduces the cell capacity to invade and migrate ($p< 0.05$) but increase their competences to form spheroids in 3D cultures and tumor growth in tumor xenografts. Moreover, in the aim to understand regulation of these cellular processes we performed (phospho-) proteomic analysis and identified 11342 phosphorylation sites with 84 up-regulated and 223 down-regulated in SKOV3 Δ*SIAT8*. Bioinformatic analysis identified down-phosphorylated peptides such as MTOR, MAPK1, CTNBB1, and JUN, enzymes which are related in cancer and MAP Kinase pathway. Specifically, we confirmed decrease of Erk1/2 phosphorylation by Western blot, known to be implicated as an EMT driverwhile proteins such as ERBB2 usually involved in tight junction pathway were enriched. Here, we provide the first evidence that GSLs play an important role in EMT/MET and that the regulation of gangliosides GSLs expression is an important driver of EMT-associated pathways

1. 1. Jacob et al., 2018 Cancer Research 78(11): 2952-2965 2. Alam et al., 2017 Scientific Reports 7: 45367 3. Rossdam et al., 2019 Anal. Chem. 91(10) : 6413-6418

487

MS-CETSA for proteome-wide study of the modulation of protein interaction states

Lingyun Dai[1], Tianyun Zhao[1], Xavier Bisteau[2], Wendi Sun[1], Nayana Prabhu[1], Yan Ting Lim[1], Radoslaw Sobota[2], Philipp Kaldis[2], Pär Nordlund[2, 1, 3]

1. School of Biological Sciences, Nanyang Technological University, Singapore

*2. Institute of Molecular and Cell Biology, A*STAR, Singapore*

3. Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden

The Cellular Thermal Shift Assay (CETSA) is based on our discovery that the biophysical principle of ligand binding-induced thermal stabilization of proteins holds true for the proteins inside cells. CETSA coupled with mass spectrometry (MS-CETSA) has become a powerful discovery method for identifying drug targets and understanding drug actions on a proteome-wide scale. However, until recently, the use of CETSA for proteome-wide studies of cellular states with differential physiological interactions has not been fully explored. In a study of specific cell cycle phases in K562 cells, with a novel compact multidimensional MS-CETSA strategy, we uncovered modulations of interaction states for more than 750 proteins along the cell cycle, reflecting the biochemical processes such as protein phosphorylation, DNA binding and protein complex formation. Notably, many protein complexes are modulated in specific cell cycle phases, reflecting their roles in biological processes such as DNA replication, chromatin remodeling, transcription, translation, and nuclear membrane decomposition. Therefore, CETSA provides the first method for direct studies of the integrated modulation of protein interaction states (IMPRINTS) with physiological ligands in living

cells. We believe this approach, which we refer to as IMPRINTS-CETSA, will open novel exciting opportunities to understand the operational aspects of cellular proteomes in intact cells and tissues.

1. Dai L, Zhao T, Bisteau X, Sun W, Prabhu N, Lim YT, Sobota RM, Kaldis P, Nordlund P. Modulation of Protein-Interaction States through the Cell Cycle. Cell. 2018 May 31;173(6):1481-1494.e13.
2. Dai L, Prabhu N, Liang YY, Bacanu S, Ramos A, Nordlund P. Horizontal Cell Biology: Monitoring Global Changes of Protein Interaction States in Cells with Proteome-Wide CETSA. Annual Review of Biochemistry. 2019. June 20; 88.

488

Defining the *Campylobacter jejuni* interactome by cross-linking mass spectrometry (XL-MS)

Ashleigh L Dale[1, 2], Dylan J Harney[1, 2], Isobel J Tenison-Colins[1, 2], Joel A Cain[1, 2], Mark Larance[1, 2], Stuart J Cordwell[1, 2, 4, 3]

1. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, University of Sydney, Sydney, NSW, Australia

3. Discipline of Pathology, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

4. Sydney Mass Spectrometry, University of Sydney, Sydney, NSW, Australia

Campylobacter jejuni is the leading cause of acute bacterial gastroenteritis in the developed world and human infection is associated with consumption of contaminated poultry, in which the organism is considered an asymptomatic commensal species. Despite the prevalence of infection, the pathogenesis of *C. jejuni* remains poorly understood. The genome encodes ~1650 proteins, however greater than 50% remain functionally unknown and even less is known about their interactions, or the protein 'interactome'. Analysis of protein interactions on a global scale is invaluable in developing an understanding of the interconnectivity of biochemical pathways, and 'interactomics' facilitated by large-scale, mass spectrometry (MS)-based proteomics has become the method-of-choice for identifying protein-protein interactions (PPIs). Cross-linking mass spectrometry (XL-MS) employs MS-cleavable chemical cross-linkers, such as disuccinimidyl sulfoxide (DSSO), which act to covalently link and stabilise interacting proteins, and allow their unambiguous identification by MS/MS. As a result, XL-MS can be used to predict the function of unknown proteins, validate protein subcellular localisations, refine protein structures, and define significant interaction networks. A novel and optimised approach for XL-MS of total cell lysate and membrane protein enriched fractions using DSSO, size exclusion chromatography and a hybrid MS2-MS3 fragmentation strategy was developed and employed to globally define PPIs in *C. jejuni*. This enabled the first non-binary and comprehensive analysis of the interactome of this organism. A total of 647 proteins were identified to partake in 868 unique and significant PPIs governed by 2,588 unique Lys-Lys residue contacts. The XL-MS approach successfully covered 39.9% of the predicted proteome of *C. jejuni* and 49.7% of the proteome as previously identified by 'bottom-up' proteomics. Interrogation of the XL-MS dataset yielded known interactions and a large subset of novel interactions, and validated XL-MS as an effective approach to identify, analyse and characterise *in vivo* PPIs and protein complexes in *C. jejuni*.

489

Global proteomic and phosphoproteomic profiling to identify proteins and pathways that regulate cell survival in hypoxic conditions

Keshava K. Datta[1], Rebekah Ziegman[1], Harsha Gowda[1, 2, 3]

1. QIMR Berghofer Medical Research Institute, Herston, QLD, Australia

2. Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia

3. School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia

Hypoxia is a common feature in various solid tumors. Cancer cells in hypoxic environments are resistant to both chemotherapy and radiation. Hypoxia is also associated with immune suppression. Identification of proteins and pathways that regulate survival of cancer cells in hypoxic environments can reveal potential vulnerabilities that can be exploited to improve efficacy of anti-cancer therapy. Gene expression profiling studies have identified several hypoxia-induced genes. This includes well-known transcription factor hypoxia-inducible factor 1-alpha (HIF-1α). We carried out global proteome profiling and phosphoproteome profiling in melanoma cell lines to identify proteins and pathways that are induced by hypoxia. We used Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer for analysis and employed label free quantitation (LFQ) for global proteomic and phosphoproteomic comparison. As expected, several proteins that are known targets of hypoxia inducible factors (HIFs) were found to be overexpressed in the hypoxic models. In addition, several metabolic enzymes showed altered expression revealing metabolic reprogramming in hypoxic conditions. We also carried out global proteomic profiling of cells grown at different concentrations of oxygen. These studies revealed potential oxygen sensors that showed oxygen concentration dependent protein expression pattern. Phosphoproteomic profiling revealed kinase mediated signaling pathways that are induced in hypoxic conditions. Our data provides a comprehensive view of proteomic alterations in hypoxic conditions and reveals potential mechanisms that regulate cell survival in hypoxic environments. These mechanisms can be targeted to improve therapeutic outcomes in cancer treatment.

49

Protein glycosylation – an overlooked feature impacting Stem cell factor and Stem cell factor receptor function

Tiago Oliveira[1], Francis Jacob[2], Chi-Hung Lin[1], Natalie Rimmer[2], Andreia Almeida[1], Michaela Mischak[1], Kathirvel Alagesan[1], Mark von Itzstein[1], Tiffany Walsh[3], Daniel Kolarich[1, 4]

1. Griffith University, Southport, QLD, Australia

2. Department of Biomedicine, University Hospital Basel, University of Basel, Basel, Switzerland

3. Institute for Frontier Materials, Deakin University, Waurn Ponds, VIC, Australia

4. ARC Centre for Nanoscale BioPhotonics, Australia, Southport, QLD, Australia

c-KIT (also known as Mast/stem cell growth factor receptor Kit, tyrosine-protein kinase KIT and CD117) plays an essential role in the regulation of cell survival and proliferation, haematopoiesis, stem cell maintenance, gametogenesis, mast cell development, migration and function, and in melanogenesis. It is a class III receptor tyrosine kinase (RTK) that contains 10 known sites of *N*-glycosylation that are overall highly conserved across most mammals, as well as some predicted sites of *O*-glycosylation. Our knowledge on c-KIT and SCF glycosylation and its effect on their function, however, has to date still not been answered.

Recombinant c-KIT and SCF were expressed in mammalian cells, affinity-purified and subjected to an in-depth glycomics and glycoproteomics characterisation using porous graphitized carbon LC ESI MS/MS glycomics and Orbitrap Fusion Reversed Phase LC glycoproteomics to guide *in silico* structure modelling and function impact prediction.

HEK293 cell expressed c-KIT is glycosylated at all 10 sites with largely complex type *N*-glycans. *O*-glycosylation prediction by NetOGlyc 4.0 indicated that up to four Ser residues ($Ser^{28, 30, 35, 38}$) are potentially *O*-glycosylated. Glycoproteomics analyses confirmed that various Ser-residues are *O*-glycosylation with mostly core 1 type *O*-glycans. On the basis of available c-KIT/SCF protein structure models we have *in silico* modelled how the presence of *N*-glycans impacts the c-KIT/SCF interaction and dimerization. Our modelling data clearly shows that due to the extensive glycosylation in the vicinity of the c-KIT/SCF interface and the c-KIT extracellular dimerisation interface it is highly likely that both glycan-glycan and glycan-protein interactions will play a key role in the c-KIT/SCF interaction, and in the dimerisation of c-KIT receptors.

This is the first in-depth study of c-KIT and SCF glycosylation that shows that protein glycosylation plays a yet not satisfactorily understood role in the interaction of SCF with c-KIT, and hence, its activation.

490

High-throughput proteomic analysis of spatially distinct features of human brain tissue

Simon Davis[1, 2], Connor Scott[2], Benedikt Kessler[1], Olaf Ansorge[2], Roman Fischer[1]

1. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

2. Neuropathology, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

Background:

Near comprehensive proteomes of bulk tissue and cultured cells can now be readily generated. In the case of tissue proteomics, these approaches often lack spatial resolution, resulting in little to no knowledge about the underlying heterogeneity of the tissue. Isolating cells of a single phenotype or clusters of cells alleviates the signal averaging caused by measuring a diverse cell population. We developed a sensitive workflow for the proteomic analysis of neurons isolated from post-mortem human brain using laser-capture microdissection. We will use this methodology to investigate the spatial distribution of proteins throughout a heterogeneous brain tumour on a high-throughput proteomics platform.

Methodologies:

To test the performance of high-throughput proteomic analysis of laser-capture microdissected samples, we isolated Purkinje cells of the Cerebellum and Betz cells of the motor cortex from a post-mortem human brain. Proteins were cleaned up and digested using a modified SP3 protocol and analysed either on an Orbitrap Fusion Lumos coupled to a nano-LC using 60 minute gradients or on a Bruker TimsTOF pro coupled to an Evosep One using 11.5 minute gradients.

Findings:

We detected over 3700 proteins from both cell types using the Fusion Lumos (& long gradient LC) and over 1900 proteins from both cell types using the TimsTOF pro (& short gradient LC). The long-gradient low throughput workflow gave more reproducible results in terms of protein identification overlap, >90 % between replicates, and quantitative reproducibility with a mean Pearson Correlation Coefficient (PCC) of 0.97, when compared to the short-gradient high-throughput workflow, >80 % identification overlap and mean PCC of 0.90.

Concluding Statement:

The 10x increase in throughput increases possibilities for investigation into the spatial distribution of proteins throughout a tissue. We will next apply this methodology in order to determine spatial proteomic profiles of distinct histological features within a brain tumour.

1. Development of a Sensitive, Scalable Method for Spatial, Cell-Type-Resolved Proteomics of the Human Brain Simon Davis, Connor Scott, Olaf Ansorge, and Roman Fischer Journal of Proteome Research 2019 18 (4), 1787-1795 DOI: 10.1021/acs.jproteome.8b00981
2. A novel LC system embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics Nicolai Bache, Philipp Emanuel Geyer, Dorte B. Bekker-Jensen, Ole Hoerning, Lasse Falkenby, Peter V. Treit, Sophia Doll, Igor Paron, Johannes Bruno Müller, Florian Meier, Jesper V. Olsen, Ole Vorm, Matthias Mann Molecular & Cellular Proteomics August 13, 2018, mcp.TIR118.000853; DOI: 10.1074/mcp.TIR118.000853
3. Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer Florian Meier, Andreas-David Brunner, Scarlet Koch, Heiner Koch, Markus Lubeck, Michael Krause, Niels Goedecke, Jens Decker, Thomas Kosinski, Melvin A. Park, Nicolai Bache, Ole Hoerning, Jürgen Cox, Oliver Räther, Matthias Mann Molecular & Cellular Proteomics December 1, 2018, First published on November 1, 2018, 17 (12) 2534-2545; DOI: 10.1074/mcp.TIR118.000900

491

Controllable *in vitro* glycoengineering with an artificial Golgi column.

Nicholas J DeBono[1], Chi-Hung Lin[2], Nicolle H Packer[1, 2], Edward S.X. Moh[1]

1. ARC Centre of Excellence for Nanoscale Biophotonics, Macquarie University, Sydney, NSW, Australia

2. Institute for Glycomics, Griffith University, Gold Coast, QLD, Australia

Glycoengineering aims to generate proteins with defined glycosylation. Single monosaccharide differences such as the absence of core fucosylation to the *N*-glycan in immunoglobulin G can cause a 100-fold increase in antibody-dependent cell cytotoxicity response. Production of glycoengineered proteins can be performed *in vivo* by genetic modification of host glycosylation enzyme expression pre-purification of the target protein, or *in vitro* using glycan modifying enzymes post-purification. While *in vitro* methods can be costly from using external enzymes and require lengthy incubation times, enzymatic reactions can be controlled to a degree unattainable *in vivo*. Here, we show the promise of an artificial Golgi column by performing *in vitro* glycosylation in an on-line column format with immobilised glycosyltransferase enzymes. By immobilising recombinantly expressed β-1,4-galactosyltransferase 1 via the poly-histidine purification tag, we utilised existing liquid chromatography systems to introduce protein, nucleotide sugar substrates and essential metal ion co-factor to the enzyme, mimicking the trans-Golgi network in the cell. The flow driven process promotes the encounter between the enzyme and substrates, enabling the enzymatic transfer of galactose onto unoccupied GlcNAc of acceptor glycans on the target protein substrate within minutes. With the enzymes immobilised in a column format, protein substrates can be continuously introduced to the enzymes which allows reusability of the column, creating a more cost-effective process compared to using free enzymes. Coupling to a liquid chromatography system enables accurate optimisation of the enzymatic reaction; we have shown that adjusting physical parameters such as column internal diameter, co-factor concentration, substrate concentration and flow rate make substantial changes to the efficiency of reaction completeness in this artificial Golgi column. Multiple enzymes can also be incorporated in sequence on-line. This form of glycoengineering shows promise for future industrial applications with the potential for large-scale glycan modification of valuable therapeutics and other glycoproteins.

492

Optimizing the isolation width in Orbitrap instruments to maximize the number of label-free quantified peptides and protein.

Carmen Paschke[1], Tabiwang N. N. Arrey[1], Rosa Rakownikow Jersie-Christensen[1], David Horn[2], Martin Zeller[1], Romain Huguet[2], Bernard Delanghe[1]

1. Thermo Fisher Scientific (Bremen) GmbH, Bremen, BREMEN, Germany

2. Thermo Fisher Scientitic , San Jose, CA, USA

In recent years, the performance of the filtering quadrupole in Orbitrap instruments has been improved. Isolation widths down to 0.4 Th are available. For reporter based quantification, smaller widths should be used to avoid interference from the co-isolated peptides. For Label Free Quantification (LFQ) that might not be the case. Here we will investigate the optimal isolation width for maximizing the number of identified and quantified proteins and peptides, together with the best possible accuracy and precision in LFQ experiments on Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ and Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometers.

To optimize identification and quantification parameters for LFQ, HeLa digests have been analyzed in different concentrations using different isolation widths on a modified orbitrap tribid instrument and a modified quadrupole orbitrap.

For precision and accuracy experiments yeast digest was used in different amounts in a constant background of 200 ng HeLa digest.

Identification strategies using SequestHT with and without second peptides identification and spectral libraries have been used.

Data analysis was performed with a beta version of Thermo Scientific Proteome Discoverer™ 2.4 software.

493

Proteomic insights into the molecular mechanisms of Alzheimer's disease using two *in vitro* models: neuronal cells (SH-SY5Y) and retinal photoreceptor cells (661W)

Liting Deng[1], Kanishka Pushpitha[2], Vivek Gupta[2], Stuart Graham[2], Mehdi Mirzaei[1], Paul Haynes[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Department of Clinical Medicine, Macquarie University, Sydney, NSW, Australia

Amyloid β (Aβ) deposition is one of the hallmarks of Alzheimer's disease (AD). In AD, neurons are injured by Aβ and die throughout the brain. With the disease progression, Aβ is believed to play a central role in retinal pathology. Anti-Aβ therapy has been shown to be protective and prevent cell death. However, molecular mechanisms leading to neurodegeneration underlying Aβ toxicity have not been clearly demonstrated. Better understanding of these degenerative processes is important for developing new therapy strategies for preventing cells from damage caused by AD. Our aim is to investigate the proteome changes affected by Aβ using SH-SY5Y cells which can be differentiated to a mature neuron-like phenotype, and 661W cells which show properties of both retinal ganglion and photoreceptor cells, respectively. Aβ1-42 fragments were added to cells to induce toxicity. Two concentrations (5 and 25 µm) and two time-points (6 and 24 h) were set to mimic the AD progression. A 10-plexed proteomics approach using TMT labelling was applied and carried out on a Q Exactive Orbitrap mass spectrometer followed by functional and protein-protein interaction analyses. The results of these studies, which separately identified 7525 and 5837 proteins from SH-SY5Y and 661W cells, showed that Aβ induced regulation of proteins and pathways depends on dose and time-points. The similar, yet diverse, molecular mechanisms affected by Aβ were illustrated in two different cell lines, respectively. The exploration of the direct link between Aβ and its toxicity has revealed more than ten AD-associated pathways. It is debatable whether they all equally contribute to the disease progression, however, an alternative explanation is that different mechanisms or even the same mechanism exert specific effects at different stages of AD progression. Our research generated a data resource for future investigations in human studies, which can then be used for therapeutic targeting.

494

Potential Protein Biomarkers for the diagnosis of gastric cancer

Chris Desire[1], Mark Condina[1], Georgia Arentz[1], Matthias Ernst[2], Peter Hoffmann[1]

1. University of South Australia, Mawson Lakes, South Australia, Australia

2. Cancer and Inflammation Laboratory, Olivia Newton-John Cancer Research Institute, Cancer and Inflammation Laboratory, Olivia Newton-John Cancer Research Institute, La Trobe University School of Cancer , Heidelberg, Victoria

Gastric cancer (GC) or stomach cancer is the fourth most common cancer in the world [1] and the second leading cause of cancer related death for both genders. The onset of gastric cancer is typically without symptoms, and most stomach cancers are not diagnosed until the later stages. In accordance the survival rates are low [2]. Diagnostic medical techniques, such as those involving an endoscopy, are invasive, expensive and generally only performed when the cancer is at an advanced stage. The ability to use a simple blood test for early detection would therefore offer a non-invasive, low cost alternative that would significantly increase the survival outcomes for those suffering from this disease.

Previous work in our group identified potential protein biomarkers for the disease, including afamin, clusterin and vitamin D binding protein [3], where their levels in serum were significantly reduced in comparison to healthy or benign patients.

Here we present the results of a larger cohort analysis, developed for serum, using multiple reaction monitoring (MRM) on a triple quadrupole mass spectrometer, coupled to a high performance liquid chromatograph. The study involved serum digestion and subsequent determination of protein levels using three unique peptides per protein. The optimisation process investigated levels of in source and in sample deamidation, where the in source deamidation of the target peptides was found to be the most significant.

1. [1] Rahman et al. World J Gastroenterol, 20 (2014) 4483–4490
2. [2] Digklia and Wagner. World J Gastroenterol, 22 (2016) 2403-2414
3. [3] Humphries et al. Biochimica et Biophysica Acta, 1844 (2014) 1051-1058

495

CARP promotes breast cancer metastasis via FOXO3a down-regulation

Penchatr Diskul-Na-Ayudthaya[1], Yun-Ui Bae[1], Ngu Trinh Van[1], Sumin Seo[1], Churat Weeraphan[2], Parunya Chaiyawat[3], Kittirat Saharat[4], Daranee Chokchaichamnankit[4], Chantragan Srisomsap[4], Jisnuson Svasti[4], Sujung Song[1], Jae-won Shim[1], Seongho Ryu[1]

1. Soonchunhyang Institute of Medi-bio Science (SIMS), Cheonan-si, CHUNGCHEONGNAM-DO,, South Korea

2. Molecular Biotechnology and Bioinformatics, Prince of Songkla University, Songkla, Thailand

3. Department of Orthopedics, Laboratory and Research Network Center (OLARN), Chiang Mai University, Chiang Mai, Thailand, Thailand

4. Laboratory of Biochemistry, Chulabhorn Research Institute, Bangkok, Thailand

Triple negative breast cancer (TNBC), a heterogeneous disease of malignancies which are associated with poor prognosis and highly aggressive. Since the treatments of TNBC are still have not establish yet and under investigation and optimization. However, one of the cause of death that reduce patient survival rate is the metastasis of cancer. Aim of this study is to elucidate the mechanism of metastasis in TNBC. we propose that Ankyrin Repeat Domain 1 (ANKRD1) promotes metastasis in breast cancer. ANKRD1 is highly expressed in the malignant metastatic breast cancer cell line compared to that in the non-metastatic breast cancer cells. Likewise, high stage tumors showed higher levels of ANKRD1 compared to those in lower stage of specimens from tissue samples. Both *in vitro* and *in vivo* functional studies of ANKRD1 show that ANKRD1 is essential for cell migration and metastasis. Several functional studies have suggested that the AMPK-ACC-FOXO3a pathway plays an important role in tumor metastasis, but there is a lack of knowledge regarding upstream regulators of this axis. We found that ANKRD1 promotes metastasis in breast cancer by suppressing AMPK-ACC-FOXO3a signaling. These findings suggest that ANKRD1 could holds promise in the future for the development of targeted therapies for TNBC and diagnostic strategies for breast cancer metastasis.

496

Improved Quantitation of HER2 by Multiple Reaction Monitoring Mass Spectrometry in Formalin-Fixed Paraffin-Embedded (FFPE) Breast Cancer Tissue

Misol Do[1], Hyunsoo Kim[2], Han Suk Ryu[3], Youngsoo Kim[1, 4]

1. Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, South Korea

2. Department of Molecular Medicine, The Scripps Research Institute, San Diego, California, USA

3. Department of Pathology, Seoul National University College of Medicine, Seoul, South Korea

4. Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, South Korea

Human epidermal growth factor receptor 2 (HER2) protein is often overexpressed in breast cancer and is correlated with a worse prognosis and thus accurate detection of HER2 is crucial to provide appropriate cares for patients. However, none of the techniques are the universal gold standard to detect accurate HER2 status. In this context, we established a multiple reaction monitoring-mass spectrometry (MRM-MS) assay that improves upon the conventional methods for differentiating HER2 status in FFPE tissues. We quantified the HER2 peptides in 210 FFPE tissues including HER2 0 (n=30), 1+ (n=30), 2+FISH- (n=61), 2+FISH+ (n=59), and 3+ (n=30). We applied the ratio between the quantitative data of HER2 peptides and normalization factors as a specific value for determining HER2 status to raise the accuracy of the quantification of HER2. In order to determine whether the data generated by MRM assay matched with the data obtained by IHC and FISH scores, the quantitative data of a HER2 peptide normalized by a quantitative data of Junctional adhesion molecule A (JAM-1) peptide were used. The Mann Whitney U test determined that significant differences were found in all the HER2 and FISH groups, and especially the MRM data can distinguish between HER2 2+FISH- and HER2 2+FISH+ ($p<0.000$), which cannot be differentiated by IHC. In addition, the MRM data distinguished the HER2-negative group and the HER2-positive group that is expected to benefit from trastuzumab therapy ($p<0.000$). The novel MRM assay that we developed clearly distinguished the equivocal HER2 status that cannot be classified by the conventional method, IHC, as well as the overall HER2 classification. Our developed assay using MRM for determining HER2 status would provide clinicians with valuable diagnostic information and ensure that all patients who have HER2-positive breast cancer have the opportunity to receive proper treatment.

1. Steiner et al., Quantification of HER2 by Targeted Mass Spectrometry in Formalin Fixed Paraffin Embedded (FFPE) Breast Cancer Tissues. MCP 2015, 14:2786 2799.
2. Hembrough et al., Quantification of Anaplastic Lymphoma Kinase Protein Expression in Non Small Cell Lung Cancer Tissues from Patients Treated with Crizotinib . Clin Chem 2016 , 62:252 261.

497

Tandem Ion Mobility Coupled With Mass Spectrometry For Gas Phase Unfolding Studies

Dale Cooper-Shepherd[1], Martin Palmer[1], Lee Gethings[1], James Langridge[1], Tyren Dodgen[2]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Sydney, NSW, Australia

Over more than a decade ion mobility spectrometry (IMS) has been employed to great effect in the field of native mass spectrometry-based structural biology. In this field IMS has been used for separating protein oligomeric states, co-populated conformations and for measuring experimental collision cross-sections for use in determining protein complex architecture. In addition, from an early stage the potential of pre-IMS activation in probing gas phase protein unfolding pathways and stabilities was pursued. This experiment, called collision-induced unfolding (CIU) has spawned a whole field of research into probing protein domain architecture, protein-ligand stabilization and therapeutic antibody comparisons. Here we describe an IMS-based instrument platform with IMSn functionality which allows novel protein CIU studies to be performed in which protein sub-populations can be mobility-selected for further rounds of IMS. Furthermore, by adding an activation step in between IMS experiments, sequential rounds of unfolding can be performed on the same ion populations allowing greater insight into protein unfolding pathways.

Human TTR was prepared at a concentration of 4 micromolar in 200 mM ammonium acetate. Native ion mobility experiments were performed on a cyclic ion mobility-enabled Q-Tof (ESI-Q-cIM-Tof) system fitted with an extended time-of-flight mass analyser, a segmented quadrupole transfer ion guide and dual gain ADC. In addition to unfolding studies the ability to mobility-select after quadrupole isolation and activation allows high mobility-resolution interrogation of ligands released from protein-ligand complexes.

498

Quantitative Subcellular Proteomics of the Orbitofrontal Cortex Region of Schizophrenia Patients

Gilberto B Domont[1], Erika Velásquez[1], Gabriel RA Carneiro[1, 2], Andrea Schmitt[3], Peter Falkai[4], Daniel Martins-de-Souza[4], Fabio CS Nogueira[1, 2]

1. Federal University of Rio de Janeiro, Rio De Janeiro, RJ, Brazil

2. Laboratory of Proteomics, LADETEC, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil

3. Department of Psychiatry and Psychotherapy, Ludwig Maximilian University of Munich, Munich, Germany

4. Laboratório de Neuroproteômica, Department of Biochemistry, University of Campinas, Campinas, SP, Brazil

Background: Schizophrenia is a chronic disease characterized by the impairment of mental functions. Several studies have carried out the proteomic characterization of post-mortem brain tissue, but only a few of these, have deepened into the subcellular analysis. A subcellular characterization using iTRAQ labeling shotgun proteomics and SRM-based targeted proteomic approaches applied to the mitochondria (MIT), crude nuclear fraction (NUC) and cytoplasm (CYT), allowed the observation of dynamic cellular changes providing valuable insights about schizophrenia physiopathology.

Methodologies: Post-mortem brain tissues from the orbitofrontal cortex of 12 schizophrenia patients and eight healthy controls were prepared following Gray & Whittaker´s and Cox and Emili protocol. Extracted proteins were digested and labeled with iTRAQ 4-plex. Three reporter groups were used for patients and one for a pool of controls. The peptide mixture was injected in an EASY-1000 nanoLC coupled to LTQ Orbitrap Velos. Mass spectra were analyzed in Proteome Discoverer 2.1 against the Human Uniprot database and the statistical analysis was made in the InfernoRDN software. The SRM validation was performed in the TSQ Quantiva mass spectrometer and in the Skyline software.

Results: We identified 940 protein groups in MIT fraction, 2022 in NUC and 2433 in CYT. Among all enriched cellular fractions, 359 protein groups were dysregulated. Quantitative proteomic analysis reveals the dysregulation of biological pathways related to calcium and apoptosis in MIT. The cell signaling disruption of CREB activation and the increase of NF-kB signaling and

complement proteins C3 were the main findings in the NUC fraction. Finally, axon guidance proteins were the most dysregulated pathway in CYT.

Conclusion: The dysregulation of CREB and NF-kB signaling together with the imbalance of glutamate, calcium, and apoptosis activation are related to synaptic damage associated with behavioral and cognitive dysfunction. These findings contribute to a better understanding of the pathophysiological process in schizophrenia

499

Data-independent acquisition mass spectrometry(DIA-MS) accompanied with precursor mass assignment enhances peptide identification

Nam Dowoon[1], Mun Dong-Gi[1], Kim Hokeun[1], Pandey Akhilesh[2], Sang-Won Lee[1]

1. Korea University, Seongbukgu, SEOUL, South Korea

2. Department of Laboratory Medicine and Pathology, Mayo Clinic, MN

Data-independent acquisition (DIA) has become promising strategy for both comprehensive identification and accurate quantification due to advanced mass spectrometers and up-to-date data analysis techniques. However, DIA produces highly complex multiplexed tandem MS spectra which makes it difficult to interpret. In order to facilitate robust and large-scale quantification along with exhaustive protein identification, different strategy for analyzing DIA data has to be developed. Here, we present a novel strategy based on unambiguous precursor mass assignment through the mPE-MMR (multiplexed post-experimental monoisotopic mass refinement) procedure combining with complementary multi-stage database searching. Accurate monoisotopic masses of multiple precursor ions are determined by mPE-MMR and resulting data were processed with multi-stage search involving Spectral library search (SLS), MS-GF+ and MODa/MODi. Precise precursor mass assignment prior to SLS resulted more sensitive and accurate peptide identification compared to conventional SLS. Moreover, assignment of precursor mass data to DIA tandem MS data allowed spectrum-centric database search engines, MS-GF+ and MODa/MODi, to applied and identified sample-specific mutated peptides and post-translational-modified peptides. This novel strategy shows considerable potential for interpreting DIA data and further exploited by adopting methods that support to obtain high-dynamic range of MS data.

50

LUX-MS enables the light-controlled elucidation of ligand-receptor interactions and functional surfaceome nanoscale organization on living cells

Maik Mueller[1, 2], Stefan Vetterli[3], Milon Mondal[3], Andy Kong[4], Yannik Severin[5], Fabienne Graebnitz[6], Niculo Barandun[6], James Prudent[7], Annette Oxenius[6], John Robinson[3], Berend Snijder[5], Alexey Nesvizhskii[4], Bernd Wollscheid[1, 2]

1. Institute of Molecular Systems Biology & Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

2. Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

3. Chemistry Department, University of Zurich, Zurich, Switzerland

4. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

5. Institute of Molecular Systems Biology & Department of Biology, ETH Zurich, Zurich, Switzerland

6. Institute of Microbiology & Department of Biology, ETH Zurich, Zurich, Switzerland

7. Centrose, Madison, Wisconsin, USA

Similar to intracellular proteins cell surface proteins engage in dynamic protein interactions to confer function. However, there is limited knowledge of how cell surface residing proteins within the surfaceome of human and microbial cells are organized into functional complexes and interact with extracellular ligands due to analytical limitations. Defining functional surfaceome synapses and dynamic interactions in cis and trans would enable the development of rational strategies to manipulate cellular signaling in health and disease. Here, we developed LUX-MS which enables light-mediated proximity detection of ligand-receptor interactions and cell surface protein neighborhoods on living cells.

LUX-MS uses small molecule Singlet Oxygen Generators (SOGs) that produce spatially restricted reactive oxygen species upon illumination. SOGs can be coupled to a ligand of choice such as pathogens, antibodies or even small molecules enabling the light-controlled photo-oxidation of ligand-proximal proteins in nanometer vicinity. Using FragPipe we found bioorthogonal amino acid modifications and boosted their yield for highly-efficient chemical labeling. Combined with a tailored quantitative DDA/DIA-MS-based workflow such an optoproteomic strategy can be used to reveal acute ligand-receptor interactions and cell surface protein neighborhoods in a discovery-driven fashion.

We used LUX-MS technology within four scenarios in order to highlight its versatile application space: First, we applied LUX-MS in human peripheral blood mononuclear cells and identified the binding target of the therapeutic antibody Rituximab. Next, we decoded ligand-receptor interactions in microbes by revealing interaction of a novel class of antibiotics with an essential outer-membrane biogenesis complex. On cancer cells, we mapped surfaceome neighborhoods targeted by a small molecule ion channel inhibitor revealing candidates for extracellular-drug-conjugate targeting (EDCs). Finally, we used LUX-MS to investigate dynamic protein interactions implicated in the formation of functional immune synapses between T-cells and antigen-presenting cells.

Taken together, LUX-MS enables light-controlled proximity-detection of dynamic protein interactions to decode the extracellular interactome across organisms.

1. Bausch-Fluck, D., Milani, E. S. & Wollscheid, B. Surfaceome nanoscale organization and extracellular interaction networks. Curr. Opin. Chem. Biol. 48, 26–33 (2019)
2. Sobotzki, N. et al. HATRIC-based identification of receptors for orphan ligands. Nat. Commun. 9, 1519 (2018)
3. Frei, A. P. et al. Direct identification of ligand-receptor interactions on living cells and tissues. Nat. Biotechnol. 30, 997–1001 (2012)
4. Kuimova, M. K., Yahioglu, G. & Ogilby, P. R. Singlet oxygen in a cell: spatially dependent lifetimes and quenching rate constants. J. Am. Chem. Soc. 131, 332–340 (2009)
5. Glasgow, H. L. et al. Laminin targeting of a peripheral nerve-highlighting peptide enables degenerated nerve visualization. Proc. Natl. Acad. Sci. U. S. A. (2016). doi:10.1073/pnas.1611642113
6. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. Nat. Methods 14, 513–520 (2017)
7. Marshall, D. J. et al. Extracellular Antibody Drug Conjugates Exploiting the Proximity of Two Proteins. Mol. Ther. (2016). doi:10.1038/mt.2016.119
8. van Oostrum, M. et al. 'De novo Classification of Mouse B Cell Types using Surfaceome Proteotype Maps'. bioRxiv 620344 (2019). doi:10.1101/620344

500

In-situ Detection of Botulinum Neurotoxin A by MALDI Mass Spectrometry Using Functionalized Chips

Jiri DRESLER[1], Petr POMPACH[2], Petra DAREBNA[2], Zuzana KALANINOVA[2], Petr NOVAK[2]

1. Military Health Institute, Prague, Czech Republic, Prague, CZECH REPUBLIC, Czech Republic

2. Institute of Microbiology, v.v.i., Czech Academy of Sciences, Prague, Czech Republic

Botulinum neurotoxins (BoNTs) are bacterial proteins produced by *Clostridium* species causing lethal disease botulism. The widely used test for the identification of BoNTs in both clinical specimens and food is the mouse bioassay, which suffers by several limitations. Thus there is a need for a rapid and sensitive method for detecting of these BoNTs. Matrix assisted laser desorption (MALDI) mass spectrometry in combination with *in vitro* enzymatic assay Endopep-MS was recently demonstrated as a robust and fast technique for detection of BoNTs. In this study, we follow up this idea and use specifically designed peptide substrate biotinylated at both termini as a target for BoNT A. The products of the enzymatic reaction are peptide fragments of the original substrate that are detectable by MALDI mass spectrometry. We used MALDI chips functionalized with biotin-binding proteins streptavidin, neutravidin and avidin to enrich the biotinylated peptide fragments from crude biological matrices. These chips were prepared by modification of indium tin oxide glass using ambient ion soft landing under atmospheric pressure. One microliter of the sample after the specific BoNT A enzymatic reaction was applied on the MALDI chip. After incubation and washing the whole MALDI chips in buffer, each spot was covered by CHCA matrix. The resulting peptide fragments were measured by Autoflex MALDI mass spectrometer. The functionalized MALDI chips achieved low nonspecific interactions and efficient peptides ionization to detect BoNT A in samples. Two peptides, products of the enzymatic reaction, were observed in the spectra. The limit of detection for enriched peptides was 0.01 ng/ml of BoNT A concentration. The results indicate that detection of BoNT A using functionalized MALDI chips is sensitive, robust, fast and might be automated for general use in MALDI Biotyper system equipped laboratories.

501

Novel LC MS-based strategies to detect small changes in the plasma proteome for metaproteomics studies

Jean-Jacques Dunyach[1], Scott Peterman[1], David Sarracino[1], Jae Choi[1], Michael Belford[1], Jesse Canterbury[1], Graeme McAlister[1], Jacob Schwartz[1], Aaron Robitaille[1], Romain Huguet[1], Vlad Zabrouskov[1]

1. Thermo Fisher Scientific, San Jose, CALIFORNIA, United States

New analytical strategies continue to be developed to expand the depth of plasma proteome profiling for putative biomarker discoveries. The enormous dynamic range and vast donor variability of non-depleted plasma presents significant challenges to low-level protein expression level perturbations needed for metaproteomics. To address this, we have developed a method combining multiplexed TMT11 labeling and UHPLC separations with a new Tribrid mass spectrometer equipped with a FAIMS interface to perform SPS-MS^3 with Real-Time search. A common plasma stock is used to create the reference standard and 4 "donor" samples each with a different amount of E. *coli* digest spiked prior to labeling with the different TMT11 tags. The plasma standard is labeled with the TMT0 tag to create an MS-level normalization. The amount of *E. coli* digest spiked into each 1 µg plasma digest sample increases from 50 to 500 ng prior to labeling. A 1:1 mixture of the standard and donor sample is analyzed with and without FAIMS as well as with and without the Real-Time Search routine. The initial results demonstrate that incorporating stepped FAIMS CV settings increases the loading capacity which is critical to handling the plasma standard spiked with the TMT11 samples. Additionally, the stepped FAIMS CV settings within one experiment increases the dynamic range over 3-fold increasing the number of *E.* coli peptides and proteins accurately measured per donor sample. Lastly, incorporation of the Real-time search algorithm further increased *E. coli* peptide and protein detection as compared to both MS^2 and traditional SPS-MS^3 routines.

502

Creating unique fragmentation in a flash: small molecule structural elucidation using UVPD

Scott Peterman[1], Seema Sharma[1], Romain Huguet[1], Vlad Zabrouskov[1], Christopher Mullen[1], Jean-Jacques Dunyach[1]

1. Thermo Fisher Scientific, San Jose, CA, USA

Structural elucidation of unknown small molecules using LC-MS^n presents significant challenges due to the overwhelming structural diversity, and biotransformations. Collision based methods generally results preferential dissociation, limiting the formation of structurally informative product ions. This issue is particularly common for conjugated metabolites or fused-ring structures. Recently, 213 nm UVPD sources were made commercially available on Orbitrap Tribrid mass spectrometers and used to dissociate a mixture of small molecule standards of varying structural classes on an LC-timescale. The experiment consisted of standard data dependent acquisition with dynamic exclusion methods with both MS and UVPD MS/MS data acquired in the Orbitrap detector. The standard mixture was analyzed using the same method except activation was done using HCD and CID for comparisons. For those structures with increasing numbers of aryl rings and Pi bonds, UVPD significantly outperformed HCD/CID in both the number of product ions, and structural coverage and uniqueness. To further evaluate the benefit of UVPD for small molecule structural elucidation, a series of flavonoids were extracted from commercial fruit smoothies and analyzed using the competitive dissociation methods. The juice extract contained around 129 previously characterized flavonoids and conjugated metabolites from five different structural classes. All data processing was automated based on database matching for known flavonoids, and prediction based on in silico fragmentation assumptions.

503

Comprehensive quantitation proteomic analysis of gray and white matter from human post-mortem brain tissue.

DUC DUONG[1], Luming Yin[1], James J Lah[1], Allan I Levey[1], Chadwick M Hales[1], Nicholas T Seyfried[1]

1. EMORY UNIVERSITY, ATLANTA, GA, United States

Introduction: Mass spectrometry (MS)-based proteomic profiling of human postmortem brain tissue from healthy and disease cases have defined global molecular signatures associated with neurological disease. However, many of these studies primarily focus on the gray matter and discard the oligodendrocyte-rich white matter prior to homogenization and downstream proteomic processing. In this study, we use quantitative proteomics to analyze white and gray matter from the same post-mortem tissue, which highlight unique cell type differences and functional pathways

Methods: Post-mortem prefrontal cortex samples were carefully dissected to separate gray and white matter. Approximately, 50mg of each tissue subtype was homogenized in buffered 8M Urea. An aliquot of 100ug of total protein lysate was subjected to sequential endoproteinase LysC and trypsin digestion. The samples were then labeled with individual tandem mass tags (TMT), combined, and fractionated into 20 fractions using electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) offline fractionation. All fractions were run on an Orbitrap Fusion Lumos mass spectrometer using the synchronous precursor selection (SPS)-based MS3 method. All spectra were processed through Proteome Discoverer (version 2.1) and quantitation information was extracted for further differential expression and pathway analysis.

Conclusions

1.**Provides a comprehensive quantitative comparison of protein expression between gray and white matter highlighting different pathways of disease relevance.**

2.**Identification of differentially expressed proteins in both white and gray matter from patients with Alzheimer's disease.**

3.**Identification of various pathways with distinct relationship to white matter.**

504

In-depth TMT Quantification of native CSF by Tandem mass Tag (TMT) channel boosting.

Lingyan Ping[1], DUC M DUONG[1], James J Lah[1], Allan I Levey[1], Nicholas T Seyfried[1]

1. EMORY UNIVERSITY, ATLANTA, GA, United States

Recent developments in quantitative proteomics has enabled deep and comprehensive analysis of cultured cells and tissues samples. In combination with isobaric tandem mass tags (TMT) and technical advancements in mass spectrometry, proteome coverage beyond 10K proteins in these samples is readily accomplished. However, these pipelines and depths have not been easily transferable to biofluids such as cerebrospinal fluid (CSF). The foremost challenge of profiling CSF is that proteins can span up to 10 orders difference. In CSF, the top 14 proteins make up more than 80% of the total protein abundance. Many studies have used immunodepletion approaches to remove these proteins in an effort to achieve better depth and higher coverage of lower abundant species. These methods however add additional variability due to differences in depletion efficiency and off target depletion of interacting proteins. Other studies have utilized native non-depleted CSF in conjunction with extensive gel and/or off-line fractionation (up to 40 fractions) to achieve quantification depths beyond 2000 proteins. This however, involves an unsurmountable amount of preparation and machine time and is not practical for larger scale studies. Here, we present an efficient and innovative approach that involves both immunodepletion and off-line fractionation to achieve superior results. With this new approach, we are able to sequence native nondepleted CSF to a depth of > 3000 proteins utilizing less than 3 days of machine time. Our approach utilizes a "boosted" TMT channel comprised of depleted CSF. All other TMT channels are comprised of non-depleted CSF samples. We have applied this optimized pipeline to assess differences in 96 samples from an early stage Alzheimer's study (32 controls, 32 asymptomatic and 32 cognitively impaired).

505

Plasma thiol proteins as biomarkers of oxidative stress

Marisa Duong[1, 2], Peter Arthur[2], Scott Bringans[1]

1. Proteomics International, Nedlands, WA, Australia

2. School of Molecular Sciences, The University of Western Australia, Crawley, Western Australia, Australia

Oxidative stress refers to a systemic imbalance between oxidants, such as hydrogen peroxide, and antioxidants and is evident in many physiological perturbations (e.g. exercise) and chronic diseases (e.g. muscular dystrophies, diabetes and neurodegenerative diseases). Biomarkers of oxidative stress in blood can be used to assess the severity of oxidative stress, but the choice of the biomarker is dependent on the particular oxidant target being investigated. Thiol groups on proteins are particularly susceptible to oxidants, and one such protein, plasma albumin, containing a thiol group (cys34), has been used as a biomarker of oxidative stress. My experimental aim was to establish if there were additional proteins in plasma containing thiol groups that might provide additional information about oxidative stress in plasma. To identify thiol containing proteins, I used chemical labelling with fluorescent tags with gel electrophoresis, and stable isotopic tags with mass spectrometry. Using gel electrophoresis, thiol groups were evident on 14 protein bands. Of these, eight proteins have been identified using mass spectrometry. Data on sensitivity of these proteins, compared to plasma albumin, to various oxidants will be presented. This work provides the foundation for further investigation into the roles of thiol proteins as biomarkers in various physiological and disease conditions.

506

Plasma proteome profiling for the discovery of predictive biomarkers to identify Neurological Heterotopic Ossification (NHO)

Mriga Dutt[1], Hsu-Wen Tseng[2], Kylie Alexander[2], Jean-Pierre Levesque[2], Michelle Hill[1]

1. Precision and Systems Biomedicine, QIMR Berghofer MRI, Brisbane, QLD, Australia

2. Mater Research, The University of Queensland, Brisbane, QLD, Australia

Neurological Heterotopic Ossification (NHO) is a severe trauma-induced condition where extra-skeletal bone mineralisation occurs in soft tissues. NHO is a frequent complication of traumatic brain injury (TBI) or spinal cord injury (SCI) and often develops around joints, which causes chronic pain and joint deformation in patients; reducing the patient's quality of life. The mechanisms leading to NHO are unknown and the only effective treatment remains surgical resection. To elucidate NHO pathophysiology, we developed the first NHO model in genetically unmodified mice, which mimics most clinical features of NHO. In order to identify novel predictive biomarkers of NHO, plasma was collected from mice developing NHO. Depleted plasma from control and NHO mice (n=10 per group) was subjected to label-free, bottom-up tandem mass spectrometry on the SCIEX TripleTOF 5600, and subsequent protein similarity search and statistical analyses were performed using MaxQuant v1.5.8.0 and Perseus v1.6.2.3 respectively. Overall, 160 proteins were retrieved and of these, 13 proteins were identified to be differentially expressed between control and NHO cohorts. Extracellular matrix protein-1 (ECM1) and vitronectin (VTNC) appeared to be down-regulated in the NHO cohort, suggesting their putative role in triggering early NHO formation. These findings however, need to be validated in future studies. In conclusion, we have characterised the plasma proteome of an SCI mouse model to identify putative predictive biomarkers inducing the onset of NHO development, following a spinal-cord injury.

507

Absolute Quantification of Apolipoproteins Following Treatment with Omega-3 Carboxylic Acids and Fenofibrate

Fredrik Edfors[1], Andreas Hober[1], Maria Ryaboshapkina[2], Jonas Malmqvist[2], Louise Rosengren[2], Andrew J Percy[3], Lars Lind[4], Björn Forsström[1], Mathias Uhlen[1], Jan Oscarsson[5], Tasso Miliotis[2]

1. Department of Protein Science, Science For Life Laboratory (KTH), Solna, Sweden

2. Translational Science, Cardiovascular, Renal and Metabolism, AstraZeneca, Gothenburg, Sweden

3. Department of Applications Development, Cambridge Isotope Laboratories Inc., Tewksbury, MA, USA

4. Department of Medical Sciences, Uppsala University, Uppsala, Sweden

5. Global Medicines Development, Cardiovascular, Renal and Metabolism, AstraZeneca, Gothenburg, Sweden

Stable isotope-labeled standard (SIS) peptides used in targeted proteomics workflows provide robust protein quantification, which is required in a clinical setting. However, SIS peptides are typically added post trypsin digestion and can vary significantly between experiments and also within a protein. This can lead to uncertain quantification as both the accuracy and precision of the assay may be compromised. These drawbacks can be remedied by a new class of internal standards introduced by the Human Protein Atlas project, which are based on SIS recombinant protein fragments. These standards are added initially to the sample and SIS peptides are released upon trypsin digestion. The PrEST technology is promising for absolute quantification of protein biomarkers but has not previously been evaluated in a clinical setting. In this study, an automated and scalable sample preparation workflow was established that enabled simultaneous preparation of up to 96 samples. Robust quantification of 13 apolipoproteins was achieved by a novel multiplex SIS protein fragment-based LC-SRM/MS assay performed on non-depleted human plasma was achieved. The assay exhibited inter-day coefficients of variation between 1.5% and 14.5% (median = 3.5%) across all apolipoproteins and was subsequently utilized to investigate the effects of omega-3 carboxylic acids (OM3-CA) and fenofibrate. The proteins were quantified in a cohort consisting of a randomized placebo-controlled trial (EFFECT I, NCT02354976). No significant changes were observed in the OM3-CA arm, while treatment with fenofibrate significantly increased apoAII and reduced apoB, apoCI, apoE and apoCIV levels. The reduction in apoCIV following fenofibrate treatment is a novel finding. The study demonstrates that SIS protein fragments can facilitate the generation of robust multiplexed biomarker assays for absolute quantification of proteins in clinical studies.

508

Identification of glomerular antigens in membranous glomerulopathy using laser capture microdissection, mass spectrometry and an R based data visualization tool

Aaron J Storey[1], John M Arthur[2], Christopher P Larsen[3], Shree G Sharma[3], Tiffany N Caza[3], Daniel J Kenan[3], Christian Herzog[2], Ricky D Edmondson[4]

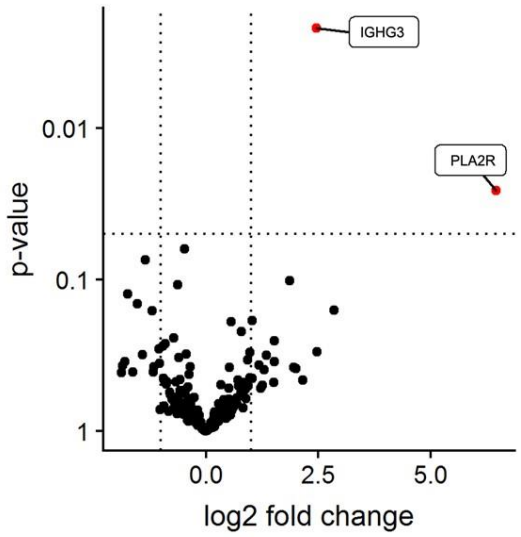
1. Biochemistry and Molecular Biology, UAMS, Little Rock, AR, United States

2. Internal Medicine, Division of Nephrology, UAMS, Little Rock, AR, USA

3. Arkana Laboratories, Little Rock, AR, United States

4. Internal Medicine, Division of Hematology/Oncology, Myeloma Center, UAMS, Little Rock, AR, United States

Membranous glomerulopathy is an autoimmune disease caused by autoantibodies directed against kidney podocyte antigens. Determination of the target autoantigen has diagnostic significance, informs prognosis, and allows for non-invasive monitoring of disease activity in serum. Many cases of membranous glomerulopathy can be classified according to the target autoantigen as either phospholipase A2 receptor (PLA2R) or thrombospondin-type-1-domain containing 7A (THSD7A)-associated disease. The autoantigenic targets for the remaining cases of membranous glomerulopathy are yet to be determined, and the goal of this project is the identification of these unknown antigens. We have developed an interactive R-based omics data visualization tool to aid in this analysis. Here we demonstrate that mass spectrometry of laser capture microdissected glomeruli can identify the antigenic etiology of membranous glomerulopathy. We utilized mass spectrometry for antigen discovery of laser capture microdissected glomeruli from FFPE and tissue IgG immunoprecipitation studies from frozen tissue. Mass spectrometric analysis of laser capture microdissected glomeruli with 23 cases of membranous glomerulopathy (MG) including 3 PLA2R-positive MG, 2 THSD7A-positive MG, and 15 cases of membranous that were negative for PLA2R and THS7DA, and 3 lupus membranous. A total of 1695 proteins were identified in this analysis. For each known membranous type, protein differential abundance was compared against the remaining groups by normalized iBAQ values. Statistical analysis was performed within the interactive R-tool, which showed meaningful differences between different subtypes of membranous glomerulopathy. PLA2R was significantly more abundant in cases of PLA2R-associated MG and showed the strongest fold change. THSD7A displayed the strongest fold changes in THSD7A-associated MG. Several proteins were seen uniquely in the PLA2R/THS7DA negative membranous or lupus membranous cases. Immunohistochemistry analysis of kidney biopsy specimens to determine their specificity is being performed.



509

Quantitative glycoproteomics provides insight into regulated enzyme activity during *N*-glycosylation stress conditions

Danila Elango[1], Ben L. Schulz[1]

1. School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD, Australia

Asparagine (*N*)- linked glycosylation is a common and important post-translational modification (PTM) present in all three domains of life - archaea, eukaryota and bacteria. It has several essential roles including facilitating protein folding, stability and function. It specifically refers to the attachment of carbohydrates onto asparagine (Asn) residues, typically within the consensus sequon N-X-S/T, where X cannot be proline. Biosynthesis of the donor glycan substrate for glycosylation is a multi-enzymatic process catalysed by the Alg (asparagine-linked glycosyltransferase) enzymes (Alg1-14). The sequential activity of the Alg enzymes results in a 14-sugar oligosaccharide ($Glc_3Man_9GlcNAc_2$) that is transferred onto a protein acceptor substrate by the oligosaccharyltransferase (OTase). Defects in the *N*-glycan biosynthetic pathway cause changes in glycan occupancy, structure, and changes in protein abundance. While the oligosaccharyltransferase (OTase) is the central enzyme in *N*-glycoprotein biosynthesis, physiological regulation of the enzyme is poorly understood. To test for the presence of a regulatory mechanism controlling OTase activity in response to glycosylation stress, we used SWATH-MS to quantify site-specific changes in glycan occupancy in a yeast model system. We compared site-specific glycosylation in yeast with defects in LLO structure (*alg6Δ*) and abundance (tunicamycin-treated cells). We identified a subset of sites that were inefficiently glycosylated in tunicamycin-treated cells but which remained efficiently glycosylated in *alg6*Δ cells. Our observations were consistent when we used the tetracycline-

repressible system to knockdown *ALG6* and *ALG7*. Our findings suggest that sequence motifs regulate site-specific OTase activity to ensure glycosylation is optimally targeted in conditions of glycosylation stress.

51

Quantitative Proteome Profiling of Stage I – IV colorectal carcinoma tissues and serum based extracellular vesicles for early onset biomarker detection

Christoph Krisp[1], Soenke Harder[1], Gerrit Wolters-Eisfeld[2], Susanne Burdak-Rothkamm[3], Ronald Simon[3], Guido Sauter[3], Hartmut Schlüter[1]

1. Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Hamburg Eppendorf, Hamburg, Germany

2. General, Visceral and Thoracic Surgery Department and Clinic, Medical Glycobiology Group, University Medical Center Haburg Eppendorf, Hamburg, Germany

3. Institute of Molecular and Cytopathology , University Medical Ceter Hamburg Eppendorf, Hamburg, Germany

Colorectal cancer (CRC) is one of the most common cancer types worldwide. It predominantly develops from benign polyps and is frequently caused by lifestyle choices. Early detection and treatment results in 5-year survival rats of greater 95% however decreases dramatically to about 10% for stage IV. Availability of reliable biomarkers for CRC is sparse. Here we use mass spectrometry to profile stage I-IV CRC specimen for early onset biomarker detection.

Fresh frozen CRC tissue specimen, with at least 10 patients per group (71 in total), were analyzed by fixed window Data Independent Acquisition (DIA) Mass spectrometry on a QExactive mass spectrometer. Pools of stage I-II and stage III-IV were fractionated by basic reversed phase (RP) chromatography for spectral library generation. Additionally, extracellular vesicle (EV) enrichment from serum samples of CRC patients and serum samples from healthy individuals were performed and analyzed by label-free data dependent acquisition on a Fusion Tribrid mass spectrometer.

Basic RP chromatography resulted in a spectral library including about 4600 proteins. Data extraction with Skyline revealed about 1900 proteins quantifiable across all 71 specimens. Statistical analysis was performed which identified stage-specific marker proteins. Especially for Stage I compared to the other stages, significant differences were observed. For later stages (III & IV) malignancy markers were significantly increased.

For the EV samples, label-free MS1 quantification was performed which quantified ~1400 proteins. The EV content from CRC patients drastically varied from the healthy individuals. With increasing stage, the number of detectable proteins associated to cancerous alteration increased. However, even the EV content from stage I CRC patients demonstrated distinct differences compared to healthy controls.

Quantitative proteome analysis of tissue and extracellular vesicle revealed potential markers for early CRC onset detection which may be used for blood-based colon cancer diagnosis.

510

GPM and SRMAtlas as novel resources for the detection and validation of missing and uncertain proteins: an update

Amr Elguoshy[1], Yoshitoshi Hirao[1], BO Xu[1], Keiko Yamamoto[1], Tadashi Yamamoto[1]

1. Niigata University, Japan, Niigata, NIIGATA, Japan

The main goal of the Chromosome -Centeric Human Proteome Project (C-HPP) is the human proteome completion by the discovery of the missing part of human proteome. According to the recent release of nextprot(2019-01), approximately 16% of all human genes have no or inadequate evidence at protein level. That are considered as missing (PE2:PE4) and uncertain proteins(PE5). in an attempt toward human proteome completion, we proposed analyzing the proteome data of other resources (GPM - SRMAtlas) side by side to peptideatlas. The results of our proposal showed that 51 missing and 4 uncertain proteins were detected in different 43 proteomic studies in GPM b greater than or equal 2 peptides. Each peptide has length of greater than or equal 9 amino acids. The newly identified missing proteins in the 43 proteomic studies may be detected due to one or more of the following proposed reasons; 1) The comprehensive analysis of specific tissue or biofluid or cell type (in specific disease or specific developmental stages) , which not investigated yet. 2) the integration of proteome study with pharmacological and other omics studies. 3) the focus on specific proteome interactome or specific family of proteins in large scale manner. 4) the discovery of new methodology or effective combinations of several methodological improvements. Then, utilizing the SRMAtlas Data To validate the newly identified missing and uncertain proteins from GPM showed that 23 missing and 1 uncertain proteins were validated by greater than or equal 1 peptide according our matching criteria .

511

Reproducibility-optimized statistical testing for proteomics studies

Tomi Suomi[1], Laura L Elo[1]

1. Turku Bioscience Centre, University of Turku and Åbo Akademi University, Turku, Finland

Differential expression analysis is one of the most common types of analyses performed on various biological and biomedical data, including mass spectrometry proteomics. A major challenge in the analysis is the choice of an appropriate test statistic, as different statistics have been shown to perform well in different datasets. To address the challenge, our reproducibility-optimized test statistic (ROTS) adjusts a modified t-statistic on the basis of the inherent properties of the data and provides a ranking of the proteins based on their statistical evidence for differential expression between sample groups. We have shown the robust performance of ROTS in multiple studies, covering both bulk and single cell measurements. The ROTS package is freely available in Bioconductor.

512

Identification and proteomic analysis of dermatophytes from human and animal dermatophytosis

Yasmin Tartor[1], Marwa Hashem[2], Shymaa Enany[3]

1. Faculty of Veterinary Medicine, Zagazig University, Zagazig, Egypt

2. Faculty of Veterinary Medicine, Suez Canal University, Ismailia, Egypt

3. Microbiology and Immunology Department, Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt

Since accurate identification of dermatophyte species is essential for epidemiological studies and implementing antifungal treatment, overcoming limitations of conventional diagnostics is a fruitful subject. In this study, we investigated real-time polymerase chain reaction (q-PCR), Matrix-Assisted Laser Desorption/ Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS), and nanoelectrospray ionization mass spectrometry (nano-ESI-MS) to detect and identify the most frequently isolated dermatophytes from human and animal dermatophytosis in comparison with conventional methods. Among 200 samples, the identified species were *M. canis* (78.22%), *Trichophyton verrucosum* (10.89%), and *T. mentagrophytes* (5.94%). Q-PCR assay displayed great execution attributes for dermatophytes detection and identification. Using MALDI-TOF MS, *M. canis*, but none of *T. violacium, T. verrucosum* or *T. mentagrophytes*, could be identified. Nano-ESI-MS accurately identify all species. The potential virulence attributes of secreted proteases were anticipated and compared between species. Secreted endoproteases belonging to families/subfamilies of metalloproteases, subtilisins, and aspartic protease were detected. The analyzed exoproteases are aminopeptidases dipeptidyl peptidases and carboxypeptidases. *M. canis* have three immunogenic proteins, siderophore iron transporter mirB, protease inhibitors, plasma membrane proteolipid 3, and annexin. In essence, q-PCR, MALDI-TOF MS and nano-ESI-MS assays are very nearly defeating difficulties of dermatophytes detection and identification, thereby, supplement or supplant conventional diagnosis of dermatophytosis.

513

Mitochondrial proteome analysis highlights Warburg effect and other carcinogenesis mechanism in cervical cancer

Sergio Encarnación-Guevara[1], Leopoldo Gómez-Caudillo[1], Roberto Jiménez-Ángel[1], Ariadna Ortega-Lozano[1], Ángel Gabriel Martínez-Batallar[1], Magdalena Hernández-Ortíz[1], María del Carmen Vargas-Lagunas[1]

1. Universidad Nacional Autónoma de México, Ciudad de México., CIUDAD DE MéXICO, Mexico

Cervical cancer incidence and mortality are rapidly growing. GLOBOCAN estimates that this neoplasia incises on 3.3% and represents 3.2% of deaths in 2018 worldwide. In women, cervical cancer ranks second in incidence and mortality behind breast cancer. Human papillomavirus (HPV) is responsible for more than 90 percent of cervical cancer cases worldwide. Besides 70% of this incidence is associated with the persistent infection with high risk human papillomavirus (HR-HPV) 16 and 18, also involved in many types of oral and anogenital cancer. HPV infection is related to alterations of cell cycle, cell death, immune system, deregulation of energetic metabolism.

Warburg effect emphasizes the energetic metabolic change observed in many types of cancer, which could be due to mitochondrial dysfunctions or structural changes. Mitochondria sense cancer metabolism as disease goes to developing. Proteomic data offers some relevant mitochondrial proteins, however solid conclusions cannot make since some mitochondrial events are supposed or difficult to correlate.

Analyzing the mitochondrial proteome in a model of HR-HPV's in cervical cancer (HaCat: control, C-33A: HPV-, SiHa: HPV-16 and CaLo: HPV-18) by means of Principal Component Analysis (PCA), followed of enrichment and PPI networks analyses; we identify a set of proteins related to different cancer HR-HPV mechanisms.

SiHa cells (HPV 16, the most frequently HPV in cervical cancer) follows a Warburg pattern with, glycolytic and viral response proteins. On the other way CaLo cells (HPV 18) interacts straightly to OXPHOS complexes, maybe inducing mitochondrial structural changes, ROS increase, HIF I stabilization, among other changes; following a different cancer mechanism.

This strategy helps to define biomarkers or molecular targets of cervical cancer in mitochondrial proteome, since it is able to detect differences between cervical cancer variants and focus in specific targets.

Part of this work was supported by CONACyT grant 220790 and DGAPA-PAPIIT grants IN207519 and IN-213216.

514

SWATH Proteomics for Human Personalized Omics Profiling (hPOP)

Kevin P Erazo Castillo[1], Sara Ahadi[2], Daniel Hornsburg[2], Tejaswini Mishra[2], Emily Higgs[2], Orit Dagan-Rosenfeld[2], Monika Avina[2], Michael P Snyder[2]

1. Chemistry Department, Stanford University, Stanford, CA, USA

2. Genetics Department, Stanford University, Stanford, CA, US

The aim of the hPOP project is to study the variance of molecular markers across a large and heterogeneous cohort. To that end, we have collected plasma, stool and urine samples from 400 participants from North America, Europe and Asia during the 2015-2019 HUPO conferences. The hPOP cohort is quite diverse, incorporating participants of different ages (average age is 43.5), sexes (30% female), ethnicities, BMIs, and various other medical markers. In collaboration with 17 labs across the globe, we aim to map the multi-molecular landscape of the participants with more than 20 omics assays covering various aspects of cellular biology. Among other molecular classes, we profiled plasma proteins via data-independent acquisition SWATH mass spectrometry using a TripleTOF 6600 System equipped with a DuoSpray Source and 25mm I.D. electrode (SCIEX). The SWATH MS data was analyzed using OpenSWATH/PyProphet and all the sample runs were aligned using TRIC software. This work complements lipidomics and metabolomics done on the same cohort and a similar longitudinal analysis in the iPOP (Integrated

Personalized Omics Profiling) project where a cohort of 100 individuals was profiled over time. In that instance, various medical risks were identified along with multiple changes in diverse biological pathways across healthy and diseased conditions. We expect to analyze hPOP data and uncover comparable trends across various segments of the world population. This study represents a large cross lab effort to explore and connect multi-omics data and a significant application of DIA SWATH mass spectrometry to broadly profile heterogeneous populations as opposed to deeply profiling individuals.

515

Analysis of the digestive proteome of the invasive golden apple snail *Pomacea canaliculata*

Sophia Escobar-Correas[2, 1], Omar Mendoza Porras[2], Israel A. Vega[1], Michelle Colgrave[2]

1. Laboratorio de Fisiología , IHEM-CONICET-UNCuyo, Mendoza, Argentina

2. CSIRO, St.Lucia, QUEENSLAND, Australia

Pomacea canaliculata, more commonly known as the golden apple snail (GAS), is a freshwater mollusc native from the Plata region comprising Argentina, Uruguay, southern Paraguay, and Brazil. Currently GAS is in the list of the top “100 world’s worst invasive alien species” since it has invaded Eastern Asia, Hawaii and Spain, where it has become a plague1. The identification of digestive enzymes in adult snails would increase the understanding of their digestive physiology and potentially generate new opportunities to eradicate and/or control this invasive species2. In this study, liquid chromatography coupled to tandem mass spectrometry was applied to define the occurrence, diversity, and origin of digestive enzymes along the digestive tract of *P. canaliculata*. Using a custom-built transcript database, a total of 3,541 proteins (≥ 95% confidence) were identified in the digestive gland and 2,518 were found in the intestine. Additionally, 920 and 801 proteins were identified in the digestive contents of crop and style sac, respectively. Later, 250 peptides derived from 81 glycosidases and 144 peptides of 55 proteases were selected and quantified using multiple reaction monitoring mass spectrometry. Through mass spectrometry, it was possible to obtain the proteome of a range of tissues specific to the survival of this important molluscan pest. This comprehensive study showed digestive enzymes involved in the digestive physiology of *P. canaliculata* which will provide the foundation for the development of strategies aiming to eradicate and/or control this invasive species

1. Lowe S, Browne M, Boudjelas S, De Poorter M. 100 of the world's worst invasive alien species: a selection from the global invasive species database: Invasive Species Specialist Group Auckland; 2000.
2. Hayes KA, Burks RL, Castro-Vazquez A, Darby PC, Heras H, Martin PR, et al. Insights from an Integrated View of the Biology of Apple Snails (Caenogastropoda: Ampullariidae). Malacologia. 2015; 58(1-2):245-302.

516

Targeted validation and quantification of cross-links from endogenous protein complexes

Marco Faini[1], Emanuela Milani[1], Charlotte Nicod[1], Alexander Leitner[1], Michael Ewing[1], Audrey Van Drogen[1], Ludovic Gillet[1], Bernd Wollscheid[1], Rudolf Aebersold[1]

1. ETH Zurich, Zurich, ZURICH, Switzerland

Background

Protein complexes represent the functional modules of the cell. Studying their assembly and structure in a close to native environment is fundamental to understanding cellular processes and their dynamics. Affinity purification mass spectrometry (AP-MS) provides bait interactors but requires multiple experiments to attribute high interaction confidence. Conversely, cross-linking mass spectrometry (XL-MS) can pinpoint protein neighbourhood and structural restraints between single amino-acid residues but requires large sample amounts and high purity. We combined the speed and throughput of AP-MS with the topological and structural information of XL-MS in a quantitative method.

Methodologies

We developed qAXL-MS, combining affinity purification from tagged human cell lines with on-beads cross-linking. After identification by spectral matching, we validate and quantify cross-linked peptides by targeted proteomics (PRM). Similarly to AP-MS, the protocol is fast, sensitive and amenable to high throughput. We implemented a data analysis pipeline in an open access R package.

Findings

First, we validated qAXL-MS on the TriC chaperonin purified from HEK293 cells. We could identify almost 3 times more cross-links than previous methods, including cross-links specific for each of the open and closed chaperonin conformation. Next, we characterized Cul4A complexes that facilitate ubiquitination. Cul4A alternates between a resting state, interacting with CAND1, and a primed state that requires neddylation. Comparing cross-links abundance between the resting and de-neddylated conditions we could confirm previous physical interactions and structural models of the two neddylation states. The accurate multi-level quantitation enabled the discrimination between compositionally and conformationally derived links. We could thus describe a set of cross-links diagnostic for several conformational changes in DDB1, CAND1 and the COP9 signalosome.

Conclusions

We developed a quantitative method to define, for a given protein complex, interaction topologies, physical interactions, distance restraints and structural markers from different cellular states from as little as ten million human cells.

517

Proteomic characteristics of gastric signet ring cell carcinoma revealed via LCM-DIA-MS and the implication on the corresponding cell lines

Yang Fan[2, 1], Yan Ren[2], Xiaomin Lou[1], Siqi Liu[2]

1. Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, Not US or Canada, China

2. BGI-Shenzhen, Shenzhen, NOT US OR CANADA, China

Signet ring cell carcinoma (SRCC) is a histological subtype of gastric tumors. It presents uniqueness in cellular morphology, epidemiology and clinicopathology as compared with other subtypes of gastric tumors, especially the adenocarcinomas (ACs). Rising in the incidence trend of gastric SRCC and lacking of practicable diagnosis for this disease prompt an urgent and valuable deed to globally portray gene expression patterns in gastric SRCC. In this project, the resected gastric tissues were collected and examined with inter-subtype matching in epidemiology and clinicopathology, resulting in 14 SRCC and 34 AC cases. Laser capture microdissection (LCM) was employed to reduce cellular heterogeneity of the tissues. Over 6,000 proteins were identified and quantified through mass spectrometry (MS) with data independent acquisition (DIA). Based on the quantitative proteomics, the differentially expressed proteins (DEPs) between tumor and adjacent tissues that were shared by SRCC and ACs were commonly enriched to extra cellular matrix and energy metabolism related pathways, while the DEPs baring the differences between SRCC and ACs were highlighted in some pathways, especially in complement cascade, which was more abundant in SRCC than ACs and its SRCC association was observed for the first time. Moreover, as an attempt to guide selection of appropriate cell models for gastric SRCC, proteomes of 15 gastric cell lines were probed and their subtype representativeness were clarified by machine learning-based classifier trained with the gastric tumor data. It turned out that all the cell lines including 3 originated from gastric SRCC, KATO-3, GCSR-1 and SNU-668, were more AC-like than SRCC-like, while MKN-7, a gastric adenocarcinoma cell line, exhibited the best representativeness for SRCC. In conclusion, this work revealed the proteomic landscape of gastric SRCC in an acceptable scale and depth in comparison with ACs and provided a suggestion on best representative cell lines for SRCC at proteomic level.

518

The role of the renin-angiotensin system in the pathophysiology of chronic lung allograft dysfunction fibrosis

Sofia Farkona[1, 2], Gregory Berra[3], Zahraa Mohammed-Ali[2], Phillip Ly[3], Liran Levy[3], Benjamin Renaud-Picard[3], Guan Zehong[3], Stephen Juvet[3], Tereza Martinu[3], Ana Konvalinka[1, 2]

1. Toronto General Research Institute, University Health Network (UHN), Toronto General Hospital, Toronto, Ontario, Canada

2. Department of Medicine, Division of Nephrology, University Health Network, Toronto, Ontario, Canada

3. Toronto Lung Transplant Program, Toronto General Hospital Research Institute, University Health Network, Toronto, Ontario, Canada

Background:

Chronic lung allograft dysfunction (CLAD) remains the leading cause of long-term mortality after lung transplant. The two main phenotypes of CLAD, bronchiolitis obliterans syndrome (BOS) and Restrictive allograft syndrome (RAS), are both characterized by chronic inflammation and fibrosis. Finding early markers of CLAD and identifying new therapeutic targets is a key goal to improve survival in this population. Angiotensin II (AngII), the main effector of the renin-angiotensin-aldosterone system (RAAS), is involved in fibrogenesis in the kidney equivalent of CLAD, interstitial fibrosis and tubular atrophy (IFTA). We have previously identified AngII-regulated proteins that are elevated in urine and kidneys of patients with IFTA. RAAS is involved in CLAD pathophysiology. We hypothesize that AngII-regulated proteins will be elevated in bronchoalveolar lavage (BAL) and lung tissue of CLAD patients, compared to stable controls.

Methods:

For the measurement of AngII-regulated proteins in BAL, we have developed parallel reaction monitoring (PRM) assays using Q-Exactive plus instrument. Two proteotypic peptides of each protein (TSP1, HMOX1, RHOB, BST1, GLNA, LAMB2, LYPA1) were selected and heavy isotope-labeled synthetic peptides were used for their absolute quantification. We performed immunohistochemical/immunofluorescence staining of TSP1 and AngII receptors in CLAD lungs and 1 control.

Results:

Using PRM assays, we detected 2 native peptides of TSP1, BST1, GLNA and LYPA1 protein and one of RHOB in a pool of BAL from controls and patients with CLAD. We constructed calibration curves and determined LOD and LOQ. In our preliminary analysis, AngII receptors and TSP1 were identified in characteristic lesions of CLAD.

Conclusion:

These preliminary data suggest that RAAS is expressed in CLAD lungs and that AngII-regulated proteins can be measured in BAL. To further these findings, we will quantify AngII-regulated proteins in BAL samples and CLAD lungs from a larger cohort of lung transplant patients.

519

Regulation of angiotensin-II fibrotic proteins in renal proximal tubular cells

Sofia Farkona[1, 2], Zahraa Mohammed-Ali[2], Parisa Shooshtari[3], Neel Sheth[2], Arun Ramani[3], Ana Konvalinka[1, 2]

1. Toronto General Research Institute, University Health Network (UHN), Toronto General Hospital, Toronto, Ontario, Canada

2. Department of Medicine, Division of Nephrology, University Health Network, Toronto, Ontario, Canada

3. Centre for Computational Medicine, Hospital for Sick Children, Toronto, Ontario, Canada

Background: Kidney fibrosis is a major clinical problem, and a final common process of injury. Angiotensin II (AngII), the main effector of the renin-angiotensin system (RAS) mediates native and allograft kidney fibrosis. Using SILAC-based proteomics, we previously identified AngII-regulated proteins in proximal tubular cells (PTECs) that may play a role in fibrosis *in vivo*. Selected Reaction Monitoring (SRM) assays targeting 6 AngII–regulated proteins demonstrated significant increase in urine excretion of

these proteins in patients with kidney graft fibrosis compared to no fibrosis. The purpose of this study is to investigate how our fibrotic proteins are regulated using chromatin accessibility profiling and transcription factor footprinting analyses.

Methods: PTECs were subjected to AngII (10^{-7}M) or vehicle and analyzed by ATAC-seq. Purified nuclei were incubated with Tn5 transposase loaded with adapters, purified and amplified with PCR. The amplified libraries were multiplexed and sequenced. The analysis of sequencing data involved aligning of ATAC-seq peaks over samples, quality checking, intensity normalization and short-listing of the replicable ATAC-seq regions differentially open between AngII and vehicle-treated PTECs.

Results: We identified 82 regions of differentially open chromatin between AngII-stimulated and vehicle-treated PTECs ($p<0.05$) that were accessible in kidney tissue. By using GTRD transcription factor database, we discovered 20 transcription factors with the highest frequency of binding to our differentially accessible ATAC-seq regions. Five transcription factors (YY1, RUNX1, SP1, CTCF, AR) have been previously linked to kidney fibrosis, AngII or chronic kidney disease. GREAT pathway analysis confirmed that these differentially accessible ATAC-seq regions were enriched in fibrotic pathways such as TGFb, SMAD2/3 signaling, and activation of the AngII receptor.

Conclusions: Using ATAC-seq, we identified open chromatin regions and transcription factors in PTECs stimulated with AngII. We confirmed that these regions were enriched in fibrotic signaling pathways. Further studies will aim to link five transcription factors to AngII-regulated proteins.

52

Phosphoproteomics of activity-dependent phospho-signalling in synaptosomes and cultured neurons

Mark E Graham[1]

1. Children's Medical Research Institute, Westmead, NSW, Australia

Neurobiological processes are modulated by linking neuronal activity to phosphorylation-based signalling to influence protein function. The identity, timing and interdependence of phosphoproteins and protein kinase activation, and mechanistic consequences was largely unresolved. We compared global activity-dependent phospho-signalling in synaptosomes with signalling in cultured hippocampal neurons, resulting in the identification of 1,917 and 7,070 unique phosphopeptides that contain activity-dependent phosphorylation sites, respectively. We monitored phospho-signalling in synaptosomes for 15 min after the stimulation, which had not been attempted at a large scale. We identified distinct patterns of presynaptic phospho-regulation across the time course that may constitute co-regulated protein networks. We also computationally predicted the protein kinases responsible. Proteins with strong signalling responses across multiple patterns were identified as potential signal integrators. The active zone scaffold protein, bassoon, was the major target of phospho-signalling, both in the number of targeted phosphorylation sites and the magnitude of change within the entire data set. These two properties of the phosphoproteomics data were combined with measures of neuronal importance from genetic data to highlight significant phospho-signalling substrates. This work constitutes a large data resource for neuronal phospho-signalling.

520

The Schwann cell secretome: a source of therapeutic targets for pancreatic cancer

Aysha Ferdoushi[1], **Xiang Li**[1], **Sam Faulkner**[1], **Phillip Jobling**[1], **Hubert Hondermarck**[1]

1. The University of Newcastle, Callaghan, NSW, Australia

Background and Significance: Pancreatic cancer is a deadly disease with no targeted therapies and a five-year patient survival rate of 5-10%. Nerves are emerging drivers in tumorigenesis. Schwann cells support nerves and are a component of the neuronal microenvironment. They have also recently been found to be involved in pancreatic cancer progression. The aim of this study was to explore the Schwann cell secretome to identify proteins that stimulate pancreatic cancer cell growth and dissemination and thus could constitute new therapeutic targets.

Methodologies: Conditioned media isolated from human primary Schwann cell cultures was tested on pancreatic cancer cells proliferation (MTT assay), migration and invasion (trans-well) assay. Secretome profiling was performed on Schwann cell conditioned media with a triple quadrupole mass spectrometer (Orbitrap) coupled to liquid chromatography (LC-MS/MS) as part of a global proteomics approach. Only proteins with at least two unique peptides identified in high confidence were considered.

Major Findings: The Schwann cell secretome was found to stimulate the growth, migration and invasion of pancreatic cancer cells. Data-dependent acquisition/discovery (DDA) search resulted in the identification of 1187 secreted proteins with a confidence corresponding to a false discovery rate (FDR) <1%. 10 candidate proteins able to stimulate pancreatic cancer cell growth and invasion were validated by western blot and functional assays.

Conclusion: The secretome of Schwann cells contain proteins that can stimulate pancreatic cancer growth. These proteins constitute novel potential therapeutic targets whose clinical value should be further considered.

521

Serine phosphorylation of STAT3 is regulated by oxidation of an adjacent methionine in Ras-driven lung adenocarcinoma

Dilanka Fernando[1, 2], Daniel Garama[1, 2], Shuai Nie[3], Daniel Gough[1, 2]

1. Department of Molecular and Translational Science, Monash University, 27-31 Wright Street, Clayton, VIC 3168, Australia

2. Centre for Cancer Research, Hudson Institute of Medical Research, 27-31 Wright Street, Clayton, VIC 3168, Australia

3. The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, VIC, Australia

Signal Transducer and Activator of Transcription (STAT) 3 is a critical signalling molecule that relays signals in response to cytokines and growth factors to regulate immune responses, metabolism and cell death. Deregulation of STAT3 activity as a transcription factor results in many diseases including cancers. Y705-phosphorylation of STAT3 is dominant for its transcriptional output which is augmented by S727-phosphorylation. Additionally we have discovered a new, non-nuclear function dependent on S727. This S727 site located within a CDK/MAPK consensus motif ie, PMSP can undergo phosphorylation by numerous serine kinases, most prominently via the MEK-ERK pathway. Mass spectrometric analyses of STAT3 post-translational modifications in A549 (Ras-driven lung cancer) cell line revealed that S727-phosphorylation was concurrent with adjacent methionine (M726) oxidation. To determine whether M726 oxidation altered S727 phosphorylation, we undertook a series of experiments. STAT3 mutants were generated that could not undergo oxidation (M726L), or mimicked oxidation (M7276Q) and stably expressed in A549 cells. We observed that the M727L blocked S727-phosphorylation whilst M727Q augmented S727-phosphorylation. Moreover we show that loss of oxidation (M727L) reduced the capacity for anchorage independent growth – indicative of reduced tumour forming potential. We next used a cell free system to directly assess the ability of Erk to phosphorylate STAT3 on S727. In these experiments we combined active, recombinant Erk2 with chemically synthesized peptides that consist of Erk2 recognition motifs in control peptides or from STAT3 and observed the effect of hydrogen peroxide-induced methionine oxidation on phosphorylation of the adjacent serine using LCMS. Our results suggest that oxidation of amino acids adjacent to Erk-catalysed phosphor-sites alters the efficacy of the kinase reaction. This has implications in cancer. Ras-driven tumours comprise ~25% of all cancers and have elevated levels of reactive oxygen species. Our data suggests that a consequence of this altered redox environment is more permissive signal transduction.

522

Time resolved assessment of nuclear protein complex dynamics via SEC-MS

Andrea Fossati[1], Fabian Frommelt[1], Li Chen[1, 2], Federico Uliana[1], Rudolf Aebersold[1, 3], Matthias Gstaiger[1], Anton Wutz[1]

1. ETH Zurich, Zurich, Switzerland

2. Monash Biomedicine Discovery Institute, Monash University, Melbourne, Australia

3. University of Zurich, Zurich, Switzerland

Introduction

X-chromosome inactivation (XCI) is a process that equalizes X-linked gene dosage in female organisms. Incomplete silencing is linked to severe phenotypes such as dysplasia and Rett syndrome while failure in XCI is lethal. While the genetic basis for XCI is well known, the mechanistic understanding on how the silencing network is organized is still missing. In this study we aimed to investigate protein complexes dynamics upon XCI in a time resolved manner with the goal of better understanding the systems level interplay between epigenetic modifiers during chromosome-wide silencing.

Methodology

Isolated nuclei from ES cells undergoing X-inactivation at different times, were lysed under mild conditions, fractionated on a size exclusion column, and measured in DIA mode. We implemented a hybrid machine learning framework to infer novel complexes from co-elution features. Following the discovery of aninteraction between Polycomb Repressive Complex 2 (PRC2) and Spen, a protein genetically linked to XCI, we proceeded to investigate the effect of Spen deletion on PRC2 activity and recruitment.

Results

We characterized > 400 differentially regulated nuclear protein complexes of which 120 are either submodules with distinct coelution features or entirely novel. Of these putative complexes, we selected one formed by PRC2 and Spen. We validated this interaction by performing endogenous immunoprecipitation of Spen. We further probed the role of Spen in PRC2 recruitment and we were able to show that ablation of Spen reduced PRC2 recruitment. Finally, Spen deletion abrogated XCI thereby pointing towards a pivotal role of this protein for epigenetic silencing.

Conclusion

We profiled nuclear protein complex dynamics in ES cells undergoing epigenetic silencing. Besides systems level insights into changes in nuclear proteome organization we identified Spen as a novel PRC2 associated protein which controls chromatin-wide epigenetic remodelling via regulation of PRC2 recruitment.

523

Ultra-sensitive LC/MS workflow for in-depth label-free analysis of single mammalian cells with nanodroplet sample processing

Khatereh Motamedchaboki[1], Yongzheng Cong[2], Yiran Liang[2], Romain Huguet[1], Yufeng Shen[3], Xuefei Sun[1], Greg Foster[1], Daniel Lopez-Ferrer[1], Andreas F Huhmer[1], Ying Zhu[4], Ryan T Kelly[2, 4]

1. Thermo Fisher Scientific, San Jose, CALIFORNIA, United States

2. Brigham Young University, Provo, UT, USA

3. CoAnn Technologies LLC, Richland, WA, USA

4. Pacific Northwest National Laboratory, Richland, WA, USA

In the last decade, single cell RNA sequencing has advanced our understanding of transcript heterogeneity. Currently, there is a strong effort to enable single cell proteomic analysis using mass spectrometry (MS)-based workflows. While the analysis of single-cell-sized aliquots from bulk-prepared tryptic digests has been demonstrated, only very recently have label-free strategies been

reported for profiling hundreds of proteins from single mammalian cells. Further development in sample processing, separations, MS and data analysis are necessary to realize single cell proteomics with greater depth of coverage and quantitative accuracy. Here we introduce an improved LC separation on new Orbitrap Eclipse™ Tribrid™ Mass Spectrometers to increase proteome coverage for single mammalian cells. Single cells were isolated via fluorescence-activated cell sorting or laser-capture microdissection. Cells were processed on Nanodroplet Processing in One Pot for Trace Samples (nanoPOTS) platform. Solid phase extraction (SPE) and analytical columns ranging from 20 to 30 µm i.d. were used for peptide trapping and separation. A Thermo Scientific™ UltiMate™ 3000 RSLCnano system coupled to Orbitrap Eclipse with a FAIMS Pro™ interface were used for this ultra-sensitive workflow. Raw data files were processed using Thermo Scientific™ Proteome Discoverer™ 2.4 software. The performance of this ultra-sensitive LC-MS workflow was evaluated and optimized using 0.2-2 ng aliquots of bulk-prepared HeLa digest. For HeLa digest loadings ranging from 0.2 ng to 2 ng, approximately 50% more peptides and 30% more proteins were identified from separations employing 20 µm LC. With the performance gains resulting from improvement of both LC separations and FAIMS-MS acquisition, ~3000 peptides and ~830 protein groups were identified by MS/MS alone from single HeLa cells, and match between runs identifications increased identifications to ~5800 peptides and ~1300 protein groups. This is the first example of >1000 proteins being identified from single mammalian cells in a label-free analysis.

524

Active Ageing – Predictive modeling and machine learning techniques uncover potential sets of metabolic markers for a longer life

Lena Fragner[1, 2], Martin Brenner[1, 2], Jana Klimashevskaia[1, 3], Simone Huber[1, 4], Bernhard Franzke[3, 5], Karl-Heinz Wagner[3, 5], Wolfram Weckwerth[1, 2]

1. Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria

2. Vienna Metabolomics Center, University of Vienna, Vienna, Austria

3. Department of Nutritional Sciences, University of Vienna, Vienna, Austria

4. University of Applied Sciences of Upper Austria, Hagenberg, Austria

5. Research Platform Active Ageing, University of Vienna, Vienna, Austria

Publish consent withheld

525

Deciphering tumoral biology by multipronged proteomics approaches: Novel therapeutic targets in chronic lymphocytic leukemia.

Manuel Fuentes[1]

1. Cancer Research Center. University of Salamanca-CSIC. IBSAL, Salamanca, SALAMANCA, Spain

B-cell chronic lymphocytic leukemia is the most prevalent hematological malignancy in Western countries, accounting for approximately 11% of all malignant hemopathies and 25-30% of leukemias in adults. It is characterized by the accumulation of mature clonal B lymphocytes in the peripheral blood and in the bone marrow, which presents a high genomic heterogeneity and alterations in intracellular signalling pathways.

Multiple genetic alterations and mutations discovered in this pathology are reflected in the proteome of this cells. Therefore, the characterization of the proteome and their post-translational modifications can be a fundamental tool as a source of identification of possible therapeutic targets. To date, no studies have investigated so far the post-translational modification profile of B-CLL tumor cells and the correlation with the therapeucti algorithm.

For this reason, to improve the knowledge of the pathogenic mechanisms at the tumor cell level, analysis of quantitative protein levels and tyrosin-phospho profiles of B-CLL cells might be crucial for better mapping the altered signalling pathways leading to inappropriate proliferation and survival signalling in tumour cells and thereby, to better understanding and classification of the disease.

The main goal of this project is to design and develop a combine proteomics strategies based on PTMs analysis of tyrosine-phosphopeptides based on immunoenrichment and further LC-MS/MS analysis, in combination with the subsequent validation by employed a customized protein array and western blot (targeting tyrosine-kinases and BCR signalling pathways). Finally, all these data sets are integrated and correlated with the clinical-biological data. The study was performed in training cohort and valiated in a different cohort (100 samples)..

Our results provide new insights into the global proteome and phosphoproteome of B-CLL, immune signalling pathways involved in tumor development and progression at the tumor cell level; which are novel targets for immunotherapy has been identified and validated in a different cohort.

526

Drug repurposing approach to target DNA gyrase from *Mycobacterium tuberculosis*

BALASUBRAMANI G L[1], RINKY RAJPUT[1]

1. Jawaharlal Nehru University, New Delhi, DELHI, India

Owing to the rise in drug resistance in tuberculosis combined with the global spread of its causative pathogen, Mycobacterium tuberculosis (Mtb), innovative anti-mycobacterial agents are urgently needed. To address this problem, we have employed a drug repurposing approach to discover novel FDA-approved drugs to inhibit Mtb growth. Here, we have used essential Mtb enzyme, DNA gyrase, a promising and potential target for novel anti-tuberculosis chemotherapeutics. High-throughput screening of compounds (using FDA-compounds library) was done against the active site of Mtb DNA gyrase, the region of ATP binding (N-

terminal domain) pocket on gyrase B subunit. Here, we identified a total of four compounds (Drug97, Drug45, Drug77, and Drug38) tightly binds to ATPase binding pocket of N-terminal domain of gyrase B (MtbGyrB47). We investigated the binding activity of identified drugs using various biophysical techniques such as thermal denaturation (CD Spectroscopy), Fluorescence titration (Fluorescence Spectroscopy) and Surface Plasmon Spectroscopy (SPR) and calculated IC50 values of these drugs using EMSA, Supercoiling, and ATPase assay. Among which, Drug97, an anthracycline antibiotic (used as an anticancer drug), was found to be a potent inhibitor of Mtb DNA gyrase. Low-μM inhibition of Mtb DNA gyrase was correlated with their low-μM minimum inhibitory concentrations for all screened FDA-drugs. Drug97 exhibited IC50 value of 0.6±0.14 μM against MtbGyrB47, kD values of 0.06±0.21 μM and MIC90 values of 0.12 μg/ml. Our results strongly suggest that the screened compounds (anthracyclines) target mycobacterial DNA gyrase, inhibits gyrase catalytic cycle.

1. Agrawal A, Roué M, Spitzfaden C, Petrella S, Aubry A, Hann M, Bax B, Mayer C. Biochem J. 2013 Dec 1;456(2):263-73. Stanger F V, Dehio C, Schirmer T. PLoS One. 2014 Sep 9;9(9):e107289. doi: 10.1371/journal.pone.0107289.

527

Investigation of Proteome Alteration in the Cerebrospinal Fluid of Glioma and Meningioma

Nikita Gahoi[1], Deeptarup Biswas[1], Saicharan Ghantasala[1], Aliasgar Moiyadi[2], Sridhar Epari[3], Mayuri N Gandhi[4], Sanjeeva Srivastava[1]

1. Indian Institute of Technology Bombay, Mumbai, MAHARASHTRA, India

2. Department of Neurosurgery, Tata Memorial Hospital, Mumbai, Maharashtra, India

3. Department of Pathology, Tata Memorial Centre, Mumbai, Maharashtra, India

4. Centre for Research in Nanotechnology and Sciences, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India

Background and Significance: Meningioma and glioma are the most prevalent primary brain tumor. Gliomas are commonly malignant and invade the normal brain tissues, whereas meningiomas are mostly indolent and benign in nature. However, higher grades of meningioma have a high recurrence risk and invasive potential. Owing to the proximity of the cerebrospinal fluid (CSF) with the central nervous system, and the presence of blood-brain barrier, makes CSF a reservoir of diagnostic and prognostic markers. In this light, an inter-grade comprehensive proteomic analysis of gliomas and meningiomas has been designed to understand the altered physiology of these tumors.

Experimental design: Differential CSF proteome profiling of meningioma grade I (n=5), meningioma grade II (n=4), LGG (n=3), GBM (n=4) and control (n=5) samples was performed by using label-based quantitative approach. The significantly altered proteins were subjected to gene ontology analysis to map their functional and biological role and a few were validated using targeted proteomic approach.

Results: Proteins associated with modulation of immune response, lipid metabolism, remodeling of extra-cellular matrix, cell surface interactors and platelet activation were found to be significantly altered. Level of several proteins including gelsolin, ENPP2, SEMA7A, CD14, osteopontin, showed significant de-regulation in the CSF of both the malignancies, and can be a putative indicator of neoplasm formation. Intriguingly, proteome trends of meningioma grade I samples was strikingly different from that of glioma and meningioma grade II samples, suggesting the differential proteome composition of benign tumors.

Conclusions: This pilot study was designed to understand the physiological modulation in the CSF of two prevalent brain malignancies to understand the underlying mechanism of tumor growth and progression. The study revealed a few candidate tumor-specific markers that can be indicative of the neoplastic growth in the CNS, which on further validation can aid in supplementing the current diagnostic and prognostic modalities.

528

Glycotorch: a glycoinformatics and molecular docking tool for improved modelling of glycosaminoglycan-protein complexes

Eric Boittier[1], Neha S Gandhi[2], Vito Ferro[1]

1. School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, QLD, Australia

2. Queensland University of Technology, Brisbane, QLD, Australia

Glycosaminoglycans (GAGs) are a family of highly sulfated amino sugars involved in an expansive range of physiological and pathological processes – including cancer, aging and the immune system. GAGs are found in all eukaryotic cells, covalently bound to proteins in the extracellular matrix. Experimental determination of protein-GAG complexes, by NMR and X-ray crystallography for example, is difficult due to high levels of natural heterogeneity and negative charge making purification from biological sources challenging.[1] Therefore, computational approaches, to complement experimental techniques, are highly desirable.

Molecular docking is an attractive computational tool that aims to provide insight into how GAGs and proteins interact at the atomic level. The reliability of docking, when predicting the complexes of carbohydrates, is limited by the parameterisation of these methods, the majority of which were originally intended for small drug-like molecules. Naturally occurring GAGs contain many charged moieties and rotatable bonds – features uncommon to most druglike molecules. To overcome this limitation, novel scoring functions based on molecular mechanics principles were developed to better model these highly charged and flexible molecules, parameterised to Kohn and Sham DFT energy calculations.

Glycotorch, a glycoinformatics and molecular docking tool, aims to improve the modelling and analysis of GAG-protein complexes. *Glycotorch Vina*, an extension of the popular program, Autodock Vina,[2] provides an implementation of these novel scoring functions and boasts a greatly improved ability to reproduce known crystal structures of GAG-protein complexes. A companion website (glycotorch.com) is also in development and provides tools to help users prepare input files and analyse

output during docking experiments. *Glycotorch Vina* also allows users to enforce constraints or additional scoring functions. This feature allows *Glycotorch Vina* to complement experimental data, such as 1D and 2D NMR observables, to create better models of biological systems and enhance carbohydrate-based drug design.

1. Gandhi and Mancera, Chem Biol Drug Des. 2008; 72(6):455-82.
2. Trott and Olson, J Comput Chem. 2010; 31(2):455-61.

529

Elucidating cross-talk between *O*-GlcNAc and phosphorylation in Tau peptides

Neha S Gandhi[1], François-Xavier Cantrelle[2], Xavier Trivelli[2], Clément Despres[2], Oliver Reimann[2], Christian P Hackenberger[3], Isabelle Landrieu[2], Caroline Smet-Nocca[2]

1. Queensland University of Technology, Brisbane, QLD, Australia

2. Unité de Glycobiologie Structurale et Fonctionnelle, Univ. Lille, CNRS UMR8576,, Lille, France

3. Leibniz-Institut für Molekulare Pharmakologie, Campus Berlin-Buch Robert-Roessle-Str. 10, Berlin, Germany

Tau protein is a key player in Alzheimer's disease as the main constituent of intracellular neurofibrillary tangles.[1] Tau protein is modified by multiple post-translational modifications (PTMs) and these PTMs can affect each other in what is generally termed "PTM crosstalk".[2] Among several post-translational modifications of Tau, both phosphorylation and *O*-GlcNAcylation are known to occur on at least 12 Ser/Thr sites in particular *O*-GlcNAcylation is well-studied in the C-terminal region of the Tau (residues 392-411 from the longest isoform of Tau; PHF-1 epitope).[3] *O*-GlcNAcylation has been involved in the modulation of tau phosphorylation levels and inhibition of tau aggregation properties while a decrease of *O*-GlcNAcylation could be involved in tau hyperphosphorylation.[4] However, the molecular mechanisms at the basis of these observations is ill-defined. Our study aims to decipher the role of *O*-GlcNAcylation in the regulation of tau phosphorylation and conformation and thereby modulate aggregation. We used advanced molecular dynamics (MD) simulations in conjunction with high-resolution NMR to describe the direct *O*-GlcNAcylation/phosphorylation crosstalk around the PHF-1 epitope and investigate their effect on peptide conformation. Our results from the all-atom explicit solvent MD simulations that the PHF-1 peptide in the absence of post-translational modifications forms transient helix from residues Ser404-Gln410, in particular, the hydrogen bond between backbone atoms of Leu408-Ser404 and Ser404-Val399 was found to be present in these helical conformations. Our advanced MD simulations around the PHF-1 phospho-epitope show that phosphorylation and *O*-GlcNAcylation both disrupt the turn-like conformation. The structural changes correlated well with the NMR studies on these Tau fragments. Our findings refute the general norm in the field that phosphorylation and *O*-GlcNAcylation of Tau are reciprocally antagonistic.

1. Li, X. et al. Eur. J. Neurosci. 2006, 23 (8), 2078–2086.
2. Lim, S. et al. Int J Mol Sci 2015, 16 (9), 20212–20224.
3. Bourre, G. et al. Front Endocrinol 2018, 9, 595.
4. Yuzwa, S. A. et al. Nature chemical biology 2008, 4 (8), 483–490.

53

SurfaceGenie: A web-based application for prioritizing cell-type specific marker candidates

Matthew Waas[1], Shana T Snarrenberg[1], Jack Littrell[1], Rachel A Jones Lipinski[1], Polly A Hansen[1], John A Corbett[1], Rebekah L Gundry[1, 2]

1. Medical College of Wisconsin, Milwaukee, WI, United States

2. Center for Biomedical Mass Spectrometry Research, Medical College of Wisconsin, Milwaukee, Wisconsin, United States

Cell surface proteins play critical roles in a wide range of biological functions and disease processes through mediation of contacts and signals between a cell and its environment. Owing to their biological significance and accessibility, cell surface proteins are attractive targets for developing tools and strategies to identify, study, and manipulate specific cell types of interest. Applications ranging from immunophenotyping and immunotherapy to targeted drug delivery and *in vivo* imaging are enabled by exploitation of cell-type specific surface proteins. Despite their utility and relevance, the unique combination of molecules present at the cell surface are not yet described for most cell types. While modern mass spectrometry approaches have proven invaluable for generating discovery-driven, empirically derived snapshot views of surface proteins, significant challenges remain when analyzing these often-large datasets for the purpose of identifying candidate markers that are most applicable for downstream applications. To overcome these challenges, we developed *GenieScore*, a prioritization metric that integrates a consensus-based prediction of cell surface localization with user-input data to rank-order candidate cell-type specific surface markers. We have demonstrated its utility for analyzing human and rodent data from proteomic and transcriptomic experiments in the areas of cancer, stem cell, and islet biology. The calculation of *GenieScores*, as well as additional scoring algorithm permutations that enable prioritization of co-expressed and intracellular cell-type specific candidate markers, is made accessible via the freely available SurfaceGenie web-application at www.cellsurfer.net/surfacegenie.

530

Associating H_2O_2 and NO related changes in the proteome of m*ycobacterium smegmatis* with enhanced survival in macrophages

Naadir Ganief[1], Jessica Sjoureman[1], Claudia Albeldas[1], Kehilwe C Nakedi[1], Clemens Hermann[1], Bridget Calder[1], Jonathan M Blackburn[2, 1], Nelson C Soares[1]

1. Division of Chemical and Systems Biology, Department of Integrative Biomedical Sciences, University of Cape Town, Cape Town, Western Cape, South Africa

2. Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, Western Cape, South Africa

Background

Mycobacterium tuberculosis is typically spread through aerosolised droplets containing the bacilli[1]. These bacilli are thought to originate from ruptured granulomas in the lung where they may multiply for extended periods, while experiencing sub-lethal stresses including oxidative and nitrosative insult[1,2,3]. Transmitted bacilli may, therefore, exist in a biochemical state typically not observed during routine culture conditions. This work[4] aims to understand the potential role of these stresses on the infectivity and survival of *Mycobacteria* by correlating the proteomic response and survival of *Mycobacteria* following infection in-vitro.

Methods

Mycobacterium smegmatis was cultured with sub-lethal doses of H_2O_2, DETA-NO or without stress and the proteome was measured using label-free shotgun LC-MS/MS, at three time points corresponding to: onset of, recovery from, and post recovery from stress. Similarly, bacteria from these time points were used to infect Raw264.7 murine macrophage like cells; uptake and survival were quantified 3 hours and 24 hours post infection, respectively. Proteomic data was processed and analysed using MaxQuant, Perseus, StringDB, KEGG and Cytoscape.

Results

We observed that sub-lethal doses of H_2O_2 or DETA-NO increased Mycobacterial survival in macrophages by up to 9-fold with DETA-NO able to induce the most resistant phenotype. Proteomic analysis identified and quantified 3336 protein. Pathway analysis revealed, that sub-lethal challenge with DETA-NO or H_2O_2, induced expression of the DevR regulon, and other known virulence factors. Furthermore, DETA-NO altered regulation in lipid metabolism, which may implicate the cell wall in increased survival.

Conclusions

These findings indicate a large overlap between the response to sub-lethal doses of H_2O_2 or DETA-NO, which may account for the increased survival observed. Additionally, we correlated an altered lipid metabolism with the higher rates of survival observed for bacteria exposed to DETA-NO challenge. This work highlights the need to study *Mycobacteria* under physiological conditions and may highlight potential therapies.

1. Gengenbacher, M., & Kaufmann, S. H. E. (2012). Mycobacterium tuberculosis: Success through dormancy. FEMS Microbiology Reviews, 36, 514–532. https://doi.org/10.1111/j.1574-6976.2012.00331.
2. (Ehrt & Schnappinger, 2009; Ganief et al., 2018; Gengenbacher & Kaufmann, 2012; Gouzy, Poquet, & Neyrolles, 2014)Ehrt, S., & Schnappinger, D. (2009). Mycobacterial survival strategies in the phagosome: Defence against host stresses. Cellular Microbiology, 11(May), 1170–1178. https://doi.org/10.1111/j.1462-5822.2009.01335.x
3. Gouzy, A., Poquet, Y., & Neyrolles, O. (2014). Amino acid capture and utilization within the Mycobacterium tuberculosis phagosome. Future Microbiology, 9(MAY 2014), 631–637. https://doi.org/10.2217/fmb.14.28
4. Ganief, N., Sjouerman, J., Albeldas, C., Nakedi, K. C., Hermann, C., Calder, B., … Soares, N. C. (2018). Associating H 2 O 2- and NO-related changes in the proteome of Mycobacterium smegmatis with enhanced survival in macrophage. Emerging Microbes & Infections. https://doi.org/10.1038/s41426-018-0210-2

531

RNF115 as a negative regulator of endosomal trafficking

Megan S Gant[1, 2], Heunis Tiaan[1], Bilkei-Gorzó Orsolya[1], Lamoliatte Frédéric[1], Härtlova Anetta[2], Trost Matthias[1]

1. Newcastle University, Newcastle, TYNE AND WEAR, United Kingdom

2. Immunology and Microbiology, Göteborgs Universitet, Gothenburg, Sweden

Introduction

Macrophages are key immune cells present in every tissue in our bodies. They initiate the innate and adaptive immune systems and clear harmful particles by phagocytosis. Particles are recognized by pathogen recognition receptors (PRRs) that trigger internalization of the bound particle into a membranous organelle called the phagosome. Phagosomes mature by fusing with early and late endosomes and finally the lysosome where the internalized cargo is degraded. Toll Like Receptors (TLRs) are the most widely understood group of PRRs. They are integral to the initiation of the innate immune response by recognizing microbial pattern associated molecular patterns (PAMPs) and danger associated molecular patterns (DAMPs) and initiating the pro-inflammatory and type I interferon pathways from the plasma membrane and the endosome, respectively. We have previously shown that IFN-γ stimulation greatly impacts phagosome functions. Here, we wanted to investigate the role of ubiquitin in phagosomal maturation and endosomal TLR signalling in IFN-γ activated macrophages.

Methods

Phagosomes were isolated from bone marrow-derived macrophages (BMDMs) and analysed using mass spectrometry, qPCR, western blot, microscopy and flow cytometry.

Results

We have identified the ubiquitin E3 ligase RNF115 as a key regulator of phagosome maturation. RNF115 is recruited to the phagosome and upregulated upon activation with IFN-γ. We show that loss of RNF115 promotes phagosome maturation, indicating that RNF115 is a negative regulator of vesicular trafficking to the lysosome. Furthermore, loss of RNF115 induces an increased pro-inflammatory response to Toll-like receptor (TLR) agonists. We demonstrate increased TLR signalling from the phagosome in RNF115 KO cells and confirm that endosomal and phagosomal maturation is important for the initiation of TLR2 induction of type I IFN from the endo-lysosomal system. Taken together, we have shown that RNF115 plays a role in the regulation of endosomal TLR initiation of the innate immune response in murine macrophages.

532

An informatics map for understanding rare mitochondrial disease symptomology

Anders O Garlid[1], Vladimir Guevara-Gonzalez[1], Calvin T Schaffer[1], Ding Wang[1], Keith D Garlid[1], Peipei Ping[1, 2]

1. University of California, Los Angeles, Los Angeles, CA, United States

2. Cardiovascular Data Science Training Program at UCLA, Los Angeles, CA, United States

Introduction: Mitochondrial diseases are complex, rare, and fatal, frequently leading to disruption of mitochondrial proteomes and function. Our limited understanding of these diseases leads to delayed diagnoses and a dearth of treatment options. Advancing our ability to provide timely and accurate diagnosis as well as effective treatment depends on FAIR (Findable, Accessible, Interoperable, Reusable) data resources to decode clinical narratives, interfacing with proteomics knowledgebases to facilitate molecular phenotyping of disease.

Methods: To impose structure on unstructured clinical information on rare mitochondrial diseases (RMDs), we created a metadata template for clinical case reports (CCRs) and aggregated over 400 CCRs on 9 RMDs, including deficiencies in respiratory complexes I-V, carnitine deficiency, Charcot-Marie-Tooth disease, MDCMC, and Barth syndrome. A digital map of ICD-10 codes is constructed to gain systematic understanding of symptoms in RMDs. For example, among 52 CCRs on 111 patients with Barth syndrome, we extracted 1,051 instances of 211 unique ICD-10 codes, along with detailed metadata.

Applications: The metadata and ICD-10 codes are housed on the MitoCases platform (www.mitocases.org/), providing a highly indexed and searchable interface through which to acquire CCRs of relevance for RMD clinicians and researchers. Searching sets of symptoms returns relevant CCRs categorized by disease, aiding in literature curation, case review, and diagnosis. Beginning with an input list of proteins from PRIDE using UniProtKB IDs, users are presented with a collection of CCRs involving the proteins of interest. Use cases are provided with sample Jupyter notebooks to assist in downstream analysis of demographics, symptoms, and biomolecules.

Conclusion: The landscape of the MitoCases digital map highlights shared and common symptoms as well as rare and unique characteristics, revealing pathogenesis and mechanistic insights underlying RMDs. Text data standardization and integration with existing protein resources as well as ontologies render metadata FAIR, enabling elevated understanding and improved patient care.

533

Establishing a Rapid, Sensitive QC protocol Utilizing Trapped Ion Mobility Mass Spectrometry and Parallel Accumulation Serial Fragmentation

Michael Krawitsky[1], Conor Mullens[2], Christopher Adams[1], Melvin Gay[3], Shourjo Ghose[2], Matt Willetts[2], Gary Kruppa[2]

1. Bruker Daltonics Inc, San Jose, CA, USA

2. Bruker Daltonics Inc, Billerica, MA, USA

3. BRUKER Pty. LTD, Preston, Australia

Introduction

In clinical and omics workflows, the quality control sample is relied upon to monitor instrument health and performance. In this work we assess isotopologues of known quantity and spiking them into a human leukemia cell line. We explore low sample concentrations run on short gradients (<30 min) to simultaneously optimize qualitative and quantitative qualification in a short period of time.

Methods

Human myelogenous leukemia cell line (K562 - Promega) was reconstituted to 22.5-180 ng/uL where a series of 35 peptide isotopologues (Pierce 7x5) at 5 orders of magnitude different concentrations, between 100 fmol (heaviest) to 8 amol (lightest), were spiked into the cell lysate. Separation was performed on a 25 cm Ionopticks Aurora column Coupled to a timsTOF Pro mass spectrometer (Bruker) operated in PASEF mode. Data analysis was performed with PEAKS Studio (BSI)

Results

To date we tested and optimized experimental conditions to achieve optimal instrument performance, which can be applied to multi-omic and clinical workflows. Peptide digest concentrations (12.5 ng to 180ng), isotopologue loads (8 amol-100 fmol on column), gradient length (8 & 18 min), column length (10 & 25 cm) and tims settings (PASEF scans, cycle time, accumulation and CCS range) were tested to develop a fast, reproducible and robust LC-MS method.

Optimized timsTOF Pro settings are as followed: Capillary voltage was 1600V, gas flow 3 l/min, gas temp: 180C, imex values were between 0.7 and 1.50 V • s/cm2 (End set to 2.10 V • s/cm2), accumulation time was 100 ms, cycle time approximately 73.0 ms (with 100% duty cycle enabled), 10 PASEF MS/MS scans and a total cycle time of 0.87sec.

Conclusions

When ion mobility and CCS merge with the complexity of a digested lysate, spiked with known levels of isotopologues the result is the most complete approach to monitor instrument sensitivity and performance.

534

High speed untargeted 4D-lipidomics LC-MS/MS workflows with Parallel Accumulation Serial Fragmentation (PASEF)

Aiko Barsch[1], Pierre-Olivier Schmit[2], Melvin Gay[3], Sven W Meyer[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Bruker France, Wissembourg, France

3. BRUKER Pty. LTD, Preston, Australia

The reliable acquisition of MS/MS data is a crucial step in lipidomics and metabolomics studies to identify analytes. This requires MS/MS fragmentation fast enough to generate high quality spectra over a sharp chromatographic peak. High sample throughput is a prerequisite for measuring large, statistically relevant sample cohorts such as those required in longitudinal clinical research studies. In this presentation, lipid extracts from the NIST reference SRM1950 plasma were used to investigate the number of lipids identified using three different reversed phase LC gradient times (6, 11, 20 minutes).

The MS data was acquired in ESI positive mode using a timsTOF Pro instrument (Bruker) in TIMS on and off modes. The acquisition of MS/MS data from the lipid extract was supported by the TIMS technology as it adds a complementary ion mobility dimension in addition to LC separation. This facilitated very fast MS/MS acquisition using a serial fragmentation of co-eluting precursors. The so-called Parallel Accumulation Serial Fragmentation (PASEF) mode allowed measuring MS/MS spectra at speeds up to 100 Hz. Resulting data were processed using the MetaboScape (Bruker) software considering all four dimensions: m/z, RT, intensity and CCS value. The open source LipidBlast MS/MS library (https://fiehnlab.ucdavis.edu/projects/LipidBlast) was used to assign lipid identities.

An identification of more than 200 lipids in 6 minute runs highlighted that PASEF MS/MS spectral quality was compatible with fast LC separations enabling up to 3x throughput compared to typical 20 minute LC runs. This demonstrates that complementary to an in-depth “ID as many as possible” approach, PASEF allows for a very fast lipid profiling based on MS/MS spectra. Additionally, we will show that even at reduced LC run times, the ion mobility separates co-eluting isobaric or isomeric compounds and provides accurate and reproducible CCS values for high confident lipid ID.

535

Genome Annotation of a Model Diatom Phaeodactylum tricornutum Using an Integrated Proteogenomic Pipeline

feng ge[1]

1. Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, HUBEI, China

Diatoms comprise a diverse and ecologically important group of eukaryotic phytoplankton that significantly contributes to marine primary production and global carbon cycling. Phaeodactylum tricornutum is commonly used as a model organism for studying diatom biology. Although its genome was sequenced in 2008, a high-quality genome annotation is still not available for this diatom. Here we report the development of an integrated proteogenomic pipeline and its application for improved annotation of P. tricornutum genome using mass spectrometry (MS)-based proteomics data. Our proteogenomic analysis unambiguously identified approximately 8300 genes and revealed 606 novel proteins, 506 revised genes, 94 splice variants, 58 single amino acid variants, and a holistic view of post-translational modifications in P. tricornutum. We experimentally confirmed a subset of novel events and obtained MS evidence for more than 200 micropeptides in P. tricornutum. These findings expand the genomic landscape of P. tricornutum and provide a rich resource for the study of diatom biology. The proteogenomic pipeline we developed in this study is applicable to any sequenced eukaryote and thus represents a significant contribution to the toolset for eukaryotic proteogenomic analysis. The pipeline and its source code are freely available at https://sourceforge.net/projects/gapeproteogenomic.

536

Quantitative redox proteomics in FLT3-ITD acute myeloid leukaemia patient samples revealed reactive oxygen species regulated oncogenic signalling, a novel treatment target.

Zacary P Germon[1, 2], Jonathan R Sillar[1, 2, 3], Honggang Huang[4], Heather C Murray[1, 2], Ryan Duchatel[1, 2], Anoop K Enjeti[1, 2, 3], Martin R Larsen[4], Nicole M Verrills[1, 2], Matthew D Dun[1, 2]

1. Faculty of Health and Medicine , University of Newcastle, Newcastle, NSW, Australia

2. Cancer Research Program, Hunter Medical Research Institute, Newcastle, NSW, Australia

3. Haematology Department, Calvary Mater Hospital, Newcastle, NSW, Australia

4. Department of Molecular Biology and Biochemistry, Protein Research Group, University of Southern Denmark, Odense, Denmark

Reactive oxygen species (ROS) are a heterogeneous group of molecules including superoxide anion and hydrogen peroxide, once considered a benign by-product of cellular metabolism. It is now widely accepted that ROS play an intricate role in regulating cellular signalling, implicit in haematopoietic stem cell growth, differentiation and self-renewal. However, under situations of oxidative stress, ROS promote tumourigeneses in all cancers, particularly haematopoietic malignancies. The radical nature of ROS, promote rapid oxidative modifications to proteins via the oxidation of critical cysteines of signalling protein when cellular defences (antioxidants) become depleted. To determine the role ROS play in leukaemogenesis we assessed the oxidative status of cysteines known to be redox sensitive. Blast cells from 14 acute myeloid leukaemia (AML) patients, either harbouring FLT3-ITD mutations or wild-type FLT3 were subjected to high-resolution quantitative phosphoproteomic profiling and protein cysteine

oxidation analysis using Cysteine-specific Phosphonate Adaptable Tags (CysPAT) to selectively label cysteine-containing peptides. These peptides were then enriched with titanium dioxide (TiO2) and subjected to mass spectrometry analysis. FLT3-ITD+ patients showed significantly increased expression of NADPH oxidase 2 (NOX2) and associated subunits, directly responsible for the production of ROS. Oxidation and inactivation of key FLT3 tumour suppressor proteins were seen in FLT3-ITD+ compared to wild-type patients. Proteins important in maintaining cellular homeostasis, such as antioxidants were differentially oxidised between patient subtypes supporting the notion of Redox dysfunction in FLT3-ITD+ AML. Targeting NOX2 in combination with FLT3 inhibitors currently in clinical use, significantly reduced cytoplasmic ROS and synergistically triggered cytotoxic cell death. Reducing oxidative stress switched off oncogenic signalling in key growth and survival signalling pathways, to activate apoptotic pathways. Our investigations shows, cooperation between oncogenic kinases, metabolism and oxidative stress, revealing a novel treatment paradigm currently under preclinical evaluation.

537

Investigating the effects of high fat diet on the testicular proteome using an ion mobility enabled data independent approach

S Jarvis[1], L Gethings[2], L Samanta[3], S Pedroni[4], D Withers[4], N Gray[1], R Plumb[5], R Winston[1], C Williamson[6], C Bevan[1]

1. Imperial College, London, UK

2. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

3. American Center for Reproductive Medicine, Cleveland, OH, USA

4. MRC London Institute of Medical Sciences, London, United Kingdom

5. Waters Corporation, Milford, MA, USA

6. Kings College, London, United Kingdom

Diet can have a significant impact on normal physiology, with high calorific foods and sedentary lifestyles contributing to the development of obesity. Globally, there is an obesity epidemic with 1 in 3 people overweight (WHO, 2015) and obesity occurring at a younger age. It is also increasingly recognised that being overweight or obese may have a deleterious effects on fertility. The association between obesity and impaired male reproductive function is multifactorial, involving alterations at the level of the hypothalamic-pituitary-gonadal (HPG) axis, as well as direct effects on the testis, including spermatogenesis and crucially somatic cell function (Sermondade et al., 2013). Here, we used a large scale discovery approach to study the effects of diet-induced obesity on the testis, generating a valuable proteomic dataset of potential 'hits' which provided insight into the biological pathways and protein networks associated with high fat (HF) diet.

Tissue samples originating from mice fed a HF or chow diet were homogenized and digested with trypsin overnight. Peptides were separated over a 90-minute gradient of 3-40% acetonitrile/0.1% formic acid. MS data were acquired using a Synapt G2-Si operating in IMS-DIA (HDMS^E). Samples were acquired in a random fashion and as technical triplicates. The data were processed using Progenesis QI for Proteomics and searched against a Uniprot *Mus musculus* database, with 1% FDR. Curated results were interrogated using Ingenuity to derive pathways of biological significance.

A total of 4960 proteins were identified across the whole study with 920 those being unique to HF. A number of statistically significant proteins with differential regulation were exhibited and found to be located within potential pathways which are identified in mediating the effects of HF diet on the testis. For example, pathways related to the blood testis barrier were readily identified as statistically significant and disrupted in animals exposed to HF diet.

538

Technical Reproducibility and Analytical Precision for a Multi-omic Study utilising Data Independent Strategies

Christopher Hughes[1], Lisa Reid[1], Adam King[1], Sarah Lennon[1], Lee Gethings[1], Ian Wilson[2], Robert Plumb[3]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Imperial College, London, UK

3. Waters Corporation, Milford, MA, USA

OMIC based studies often consist of large scale cohorts in order to allow for more confidence in statistically relevant findings. The LC-MS system and associated methodologies used to perform these experiments need to demonstrate robustness, reproducibility and stability over the entire analysis. These characteristics then allow for the identification of statistically relevant markers and reliable quantification. In this study, we show demonstrate these features for a multi-omic study using data independent approaches over multiple platforms between different laboratories.

Samples consisted of liver tissue and plasma extracts derived from mice which had been administered statins. Appropriate sample preparation protocols provided proteomic, lipidomic and metabolomic sample sets. LC-MS data were collected using standard flow chromatographic conditions coupled to a Synapt XS mass spectrometer, operating in DIA (SONAR or HDMS^E). For ion mobility (IMS) based acquisitions, collision cross section (CCS) values were generated. The IMS was calibrated using a mixture of compounds, covering a range of m/z values and selected for use in either positive or negative ESI. Data were processed using either Progenesis QI for Proteomics or QI for proteomic and lipidomic/metabolomic data respectively. Identifications resulting from Uniprot (proteomics), LipidMaps (lipids) and HMDB (metabolites) were appended to the processed data. Multivariate statistical analysis was conducted using SIMCA P. In order to assess the reproducibility aspect, data was collected for the same sample sets using replica chromatographic conditions over multiple platforms in different laboratories.

Assessment of the data across all OMIC experiments from the multiple platforms, resulted in low %CV's for normalised abundance and hence provided high quantitative accuracy. High mass accuracy was also maintained across the various datasets with a significant amount of the collected data being less than 2 ppm. CCS measurements were also shown to be highly reproducibility for both day-to-day (based on the same platform) and laboratory-to-laboratory.

539

Analysis of histones from HEK293T cells using a QTOF with trapped ion mobility and PASEF workflows

Shourjo Ghose[1], Matthew Willetts[1], Miranda gardner[2], Michael Freitas[2], Gary Kruppa[1], Adam Rainczuk[1]

1. Bruker Daltonics, Billerica, MA, United States

2. Department of Cancer Biology and Genetics,, The Ohio State University , Columbus , Ohio, USA

Introduction

Histones form the fundamental unit of the eukaryotic nucleosome consisting of 146 bp of DNA wrapped around protein octamer units. Bottom-up approaches to studying histones result in a mixture of isobaric species, making them challenging to analyze via routine LC-MS/MS. Here, we have utilized trapped ion mobility on a QTOF, to separate isobaric species for the detection of various modforms. Utilizing PASEF workflows, we detect over 2000 peptide sequences and achieve increased coverage of histone modforms without specialized chromatography and data interpretation.

Methods

Histones were acid-extracted from Human embryonic kidney 293T cells and derivatized with propionic anhydride/ACN (1:3) pre- and post-digestion with trypsin. Peptides were separated on a 1.6µm C18 25cm x 75µm column (Ionoptiks) using a nanoElute nano LC (Bruker Daltonics) coupled to a trapped ion mobility equipped Q-TOF mass spectrometer (timsTOF Pro). Acquired data were analyzed via Data Analysis (Bruker Daltonics) and PEAKS (Bioinformatics Solutions, Inc) softwares.

Preliminary Data or Plenary Speakers Abstract

Due to the abundance of arginine and lysine amino acids, tryptic digests of histones often produce many fragments of less than 5.Therefore, propionic anhydride is utilized to derivatize the lysine residues resulting in Arg-C like peptides fragments. However, this process coupled with the presence of multiple post translational modifications on histones, produces a complex mixture of isobaric and near-isobaric species. Here, we have utilized standard reverse phase chromatography coupled with PASEF workflows to identify over 200 histone protein groups, 2000 peptide sequences and 6000 peptide spectrum matches (PSM) (at 0.1% FDR as estimated by PEAKS using decoy fusion) from 270ng of protein. Additionally, the orthogonal separation afforded by trapped ion mobility has allowed us to generate more PSMs per modform. This subsequently translates into more modforms identified for a given peptide sequence resulting in improved depth and coverage of PTMs on a particular histone.

540

Integrative proteomic and phosphoproteomic study illustrates role of extracellular matrix in blast disease signaling and host immunity

Sudip Ghosh[1], Arunima Sinha[1], Kanika Narula[1], Eman Elagamey[1], Niranjan Chakraborty[1], Subhra Chakraborty[1]

1. National Institute of Plant Genome Research (NIPGR), India, New Delhi, DELHI, India

Background

Plant-pathogen interaction is a multifaceted process that involves two distinct forms of chemical communication referred as recognition and defense. Plant extracellular matrix (ECM) dictates early interaction between plant and pathogen and acts as regulatory hub of many biological responses and signaling networks that perceives and transmits patho-stress signal. Blast disease caused by hemibiotrophic fungus *Magnoporthe oryzae* is a major impediment for global crop productivity. To elucidate the role of ECM in imparting resistance against blast disease, temporal changes of ECM proteome and phosphoproteome was studied in a resistant rice cultivar upon *M. oryzae* infection.

Method

Patho-stress was imposed on three-leaf-stage seedlings and tissues were harvested at early, middle and later phases of *Magnaporthe* attack. Temporal proteome and phosphoproteome were developed with ECM enriched fraction using TMT labeling and peptide fractionation followed by TiO2-based enrichment of phosphopeptides. Proteins were identified using LC-MS/MS analysis. Integrated global network was built to identify immunity related pathways.

Results

Multiplex quantification of patho-stressed rice ECM proteoforms led to the identification of 336 immune-responsive proteins (IRPs) and 23 phosphopeptides mapped to 17 immune-responsive phosphoproteins (IRPPs) related to cell wall remodeling, extension, hydration and acidification linked to ROS and extracellular calcium signaling. Furthermore, the data highlighted protein phosphorylation as a critical regulatory mechanism controlling anisotropic growth and remodeling of the wall imparting immunity. Cluster and network analyses identified significant proteoform groups and hubs pointing towards the onset and context of disease signaling.

Conclusions

Integrated proteome and phosphoproteome analyses for the first time provides an useful insight into the complex regulatory network operating in the ECM. A comprehensive analysis not only helps to unravel the mechanism of blast resistance, but also enlist novel biomarkers for targeted genetic manipulation for food and nutrition security.

541

Defining the mechanism of action of ozonide antimalarials using untargeted metabolomics and proteomics

Carlo Giannangelo[1], Ghizal Siddiqui[1], Amanda De Paoli[1], Bethany M Anderson[1, 2], Laura E Edgington-Mitchell[1, 2, 3], Susan A Charman[1], Darren J Creek[1]

1. Monash Institute of Pharmaceutical Sciences, Monash University, Melbourne, Victoria, Australia

2. Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, Victoria, Australia

3. Department of Maxillofacial Surgery, College of Dentistry, New York University, New York, USA

Malaria caused approximately 435 000 deaths in 2017 and resistance has now emerged to the current first-line artemisinin-based antimalarials, highlighting the urgent need for new malaria therapeutics. The ozonides are a novel class of synthetic antimalarial drugs with potent activity against the deadliest malaria parasite species, *Plasmodium falciparum*, but their mechanism of action is poorly defined. In this study, the mode of action of OZ277 (marketed in India) and OZ439 (in Phase IIb clinical trials) were examined in *P. falciparum* infected red blood cells using an untargeted multi-omics approach consisting of proteomics, peptidomics and time-dependent metabolomics.

Untargeted metabolomic profiling using LC-MS with high resolution accurate mass spectrometry revealed a rapid depletion of short chain haemoglobin-derived peptides and the formation of ozonide-alkylated haem adducts in drug-treated parasite cultures. A dedicated peptidomics method was also developed and revealed drug-induced accumulation of long chain haemoglobin-derived peptides. These findings agree with the proposal that ozonides are activated by free haem in the food vacuole of the parasite and initially perturb haemoglobin catabolism. Extended ozonide exposure disrupted secondary biochemical pathways, including lipid metabolism and nucleotide biosynthesis. Quantitative proteomic analysis confirmed ozonide treatment perturbs these pathways, but also revealed an upregulation of proteins involved in translation and the ubiquitin-proteasome system, suggestive of the parasite engaging a proteostatic stress response to mitigate ozonide-induced damage. Furthermore, untargeted chemical proteomic studies showed that ozonides alkylate multiple proteins within the parasite, which likely contributes to ozonide-induced protein stress.

This unbiased multi-omics approach revealed ozonides initially impact haemoglobin digestion, followed by secondary effects on additional biochemical pathways that are critical for parasite survival. These findings provide new insight into the mode of action of ozonides, and facilitates new opportunities for interventions to enhance their antimalarial efficacy and reduce the potential for developing resistance.

542

Proteins present in saliva participate in tooth enamel hardening in pigs

Ana Gil de Bona[1, 2], Fabian Schulte[1, 2], Maren Teichmann[1, 2], Markus Hardt[1, 2], Felicitas B Bidlack[1, 2]

1. The Forsyth Institute, Cambridge, MA, United States

2. Department of Developmental Biology, Harvard School of Dental Medicine, Boston, MA, United States

Background: In mammals, tooth crown formation completes before eruption, but enamel can harden or decay after eruption with saliva playing a critical role in these processes. However, the role of salivary proteases and peptides in post-eruptive enamel hardening is unresolved. The mechanism of enamel formation is highly conserved in mammals, although speed and eruption timing vary between species. Compared to humans, pig teeth form very fast and erupt with less mineralized enamel that hardens further after eruption to withstand chewing forces. Our working hypothesis is that post-eruptive porcine enamel mineralization is mediated by saliva proteases and pellicle components that facilitate removal of enamel matrix proteins and influx of mineral ions to grow enamel crystals. The goal of this work is to characterize the protein composition of porcine whole saliva, plaque/pellicle and enamel at different stages after eruption. The purpose of this work is to resolve and leverage the mechanism of this rapid and marked enamel hardening to avoid dental decay in pigs and use the pig model for human dental research.

Methods: Saliva samples were collected from pigs 2, 4 and 16 weeks old using ropes, pigs sacrificed, and dentitions collected. Trypsin digestion and analysis by LC-MS/MS were performed after protein extraction. PEAKS software was used to identify protein composition and compare protein abundances.

Findings: Preliminary saliva data from 2-, 4- and 16-week-old pigs (n=1 pig per age) showed differences in protein profiles and abundances. Mmp9 and Klk peptides were identified with proteases, e.g. Cathepsin D, Alpha-1-antitrypsin, Serpin, Lipocalin (a binding protein), Serum albumin and IgG.

Conclusions: The developmental differences of pigs are also observed in proteomic profiles in saliva. Analyses of additional saliva and enamel samples are ongoing to validate these results and test their participation in enamel mineralization processes.

543

Identification of immunogenic β-lactam modified ligands in T-cell meditated penicillin hypersensitivity

Shawn Goh[1], Patricia Illing[1], Nicole Mifsud[1], Katherine Scull[1], Robert Puy[2], Robyn O'Hehir[2], Anthony Purcell[1]

1. Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC, Australia

2. Department of Allergy, Immunology and Respiratory Medicine, Monash University, Central Clinical School, The Alfred Hospital, Melbourne, VIC, Australia

The Penicillins are a group of β-lactam antibiotics that provide a first line of defence against microbial infections. However, 11.5% of individuals prescribed penicillin develop allergic responses to the drug, which if left unchecked can have serious implications on patient health. Traditionally, these allergies were considered to be predominantly IgE mediated, although new evidence points towards the involvement of T-cells. The ability of penicillins to haptenate serum proteins (through covalent binding) can lead to

the generation of neoantigens, thereby triggering an immune response. These haptenated neoantigens, following processing and presentation by antigen presenting cells, may be displayed in the form of peptides by the human leukocyte antigen (HLA) molecules for recognition by T-cells. Activation of these drug-specific T-cells can mediate a range of pathologies, ranging from mild cutaneous reactions to more severe clinical conditions such as drug-induced liver injury.

We have recruited 11 penicillin-induced hypersensitivity patients, with only 40% having elevated serum IgE against the causative drug. This suggests that other factors/cellular populations (*e.g.* T-cells) may contribute to adverse reactions observed in these individuals. Here, we describe that penicillin-responsive T-cells isolated from a hypersensitive patient are activated upon exposure to penicillin and restricted to the common Caucasian HLA-A*02:01 allomorph. We hypothesised that a HLA-A*02:01-restricted penicillin-haptenated neoepitope was presented to these drug-specific T-cells and this contributed to the adverse drug reaction.

Applying an immunopurification LC-MS/MS workflow, we analysed the immunopeptidome of HLA-A*02:01 from patient-derived B-LCLs treated with or without penicillin. In combination with a robust set of criteria for the identification of haptenated peptides, based on haptenated human serum albumin, we identified three HLA-A*02:01 presented penicillin-haptenated peptides.

Together, these results will allow us to utilise the fragmentation patterns of penicillin-modified peptides to conduct further, more targeted MS analysis that will provide novel molecular insights into this common drug reaction.

544

Pro-inflammatory cytokines induce immunopeptidome plasticity within Triple Negative Breast Cancer cells.

Gabriel Goncalves[1], Divya Duscharla[1], Ritchlynn Aranha[1], Rochelle Ayala Perez[1], Nicole Mifsud[1], Nathan Croft[1], Ethan Passantino [1], Anthony Purcell[1], Pouya Faridi[1]

1. Department of Biochemistry and Molecular Biology, Monash University , Clayton, Victoria , Australia

The heterogeneous nature of Triple-Negative Breast Cancer (TNBC) has highlighted the need for novel approaches in treatment. Recent studies demonstrate the feasibility of a peptide vaccination approach to induce T-cell mediated killing of these cancerous cells.

To identify potential T-cell targets we examined how the pro-inflammatory cytokine interferon-γ (IFNγ), which is highly expressed within the microenvironment of infiltrated tumours, modulated the peptide repertoire of HLA-I and HLA-II alleles. We isolated native HLA-I and HLA-II-peptide complexes from a TNBC model cell line (MDA-MB-231) under mild lysis conditions using immunoaffinity chromatography, isolated their peptide cargo by RP-HPLC, prior to analysis of peptide containing fractions by high resolution mass spectrometry for peptide identification.

We identified 46,128 unique peptides presented across HLA-I and HLA-II. In general, IFNγ treatment caused a 2.2-fold increase in peptide abundance and increased the diversity of the peptide repertoire from 14,270 peptides (derived from 4892 source proteins) to 26,135 peptides (derived from 6668 source proteins) in the HLA-I immunopeptidome. We see a 66% overlap between peptides presented across HLA-I in IFNγ treated and untreated samples. In contrast, HLA-II was not expressed in untreated cells, yet was expressed upon IFNγ treatment and resulted in the detection of 25,441 peptides following isolation of HLA-II molecules from the treated cancer cells. In total, 8044 unique source proteins were represented via the TNBC peptidome presented by HLA-I and -II molecules, with 3152 being unique to IFNγ treatment and only 505 being unique to untreated cells.

These results highlight the high degree of plasticity of the TNBC immunopeptidome in response to changes in the tumour microenvironment. IFNg increased the diversity and abundance of the peptide repertoire suggesting that under pro-inflammatory conditions a greater variety of vaccine targets are unveiled to the immune system. This has important implications for the development of personalised cancer vaccination strategies.

545

A proteomic landscape of triple-negative breast cancer

Tianqi Gong[1, 2], Jun Qin[2], Zhimin Shao[3], Pengyuan Yang[1], Weimin Zhu[2]

1. Fudan University, Shanghai, SHANGHAI, China

2. National Center for Protein Sciences (Beijing), Beijing, China

3. Fudan University Shanghai Cancer Center, Shanghai, China

Publish consent withheld

546

An efficient workflow for identification and monitoring of host cell proteins during monoclonal antibody bioprocessing

Paul N Goulding[1], Catalin E Doneanu[2], Malcolm Anderson[3], Alex Xenopoulos[4], Romas Skudas[5], Ying Qing Yu[2], Weibin Chen[2]

1. Nonlinear Dynamics, Waters Corporation, Newcastle, UK

2. Waters Corporation, Milford, MA, USA

3. Waters Corporation , Wilmslow, UK

4. EMD Millipore Corporation , Bedford, MA, USA

5. Merck KGaA , Darmstadt, Germany

Even after sophisticated purifications steps, low levels (1-100 ppm) of host cell proteins (HCPs) remain in the final purified drug substance. Some of the HCPs may cause immunogenic reaction in humans, therefore it is critical for patient safety that HCPs be

identified and quantified. The analytical methods typically used for HCP quantification are based on immunoassays (ELISA), but ELISA cannot guarantee proteome-wide coverage.

In recent years, LC/MS-based assays have been adopted as orthogonal techniques to ELISA for HCP analysis due to their flexibility and potential for full proteome-wide applications. Here we describe an efficient analytical scale LC/MS workflow that allows the identification and quantification of HCPs during mAb purification in a CHO cell line.

The first step of the HCP identification and quantification workflow is the HCP discovery assay employing data-independent MS^E acquisition using 90 min peptide separations. Following data processing with Progenesis QI for proteomics 4.2, HCPs are identified by a proteome-wide database search. In addition, LC/MS data can be assembled into spectral libraries, containing peptide precursors, charge states and retention times. In the second step of the HCP workflow, additional HCP samples derived from the purification of the same biopharmaceutical are analyzed by higher–throughput HCP monitoring assays using MS^E acquisitions with 30 min peptide separations. Each purified sample can be analyzed by LC/MSE followed by searching against the spectral library for HCP identification and quantification. The HCP workflow described above was tested for identification of HCPs from the NIST mAb and then applied to identification and monitoring of HCPs from a different monoclonal antibody [1]. Our results show that HCPs can be confidently identified, quantified and monitored in biopharmaceutical samples using the 1D LC-MS HCP workflow.

1. 1.Huang L, Wang N, Mitchell CE, Brownlee TJ, Maple SR, De Filippis MR Anal Chem 2017, 89, 5436.

547

Assessment of hyperglycaemia-induced modifications in the kidney mitochondrial proteome of a streptozotocin-induced diabetic rat model

Cesare Granata[1], Vicki Thallas-Bonke[2], Matthew Snelson[1], Adrienne Laskowski[1], Georg Ramm[1], Mark E Cooper[1], David Stroud[3], Rebecca H Ritchie[2], Melinda T Coughlan[1]

1. Monash University, Melbourne, VIC, Australia

2. Baker Heart & Diabetes Institute, Melbourne, VIC, Australia

3. The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, VIC, Australia

The kidney is the second most mitochondria-rich organ, and effective energy production is essential for maintaining a healthy kidney function. Dysfunctional mitochondrial bioenergetics such as defective respiratory chain function, and compromised mitochondrial morphology are thought to be central to the development of diabetic kidney disease (DKD). However, the role of systemic glucose concentrations in the development of this disease remains unknown. The effects of different blood glucose concentrations in the kidney of a rodent model of streptozotocin-induced diabetes, including respiratory chain function, mitochondrial dynamics and the mitoproteomic landscape were investigated. Diabetic rats were treated with either an intensive or a conventional insulin therapy for 8 weeks, resulting in blood glucose levels of ~20mmol/l (moderate hyperglycaemia, MHG) or ~30mmol/l (severe hyperglycaemia, SHG), respectively, and were compared to controls. Albuminuria and glomerulosclerosis, representing hallmarks of diabetic kidney disease, were induced in both MHG and SHG. However, intensive insulin therapy (MHG) afforded renoprotection, leading to improved glomerular hyperfiltration, and decreased albuminuria, glomerular injury, and renal fibrosis, compared to rats with SHG. Kidney mitochondrial bioenergetics were altered in both groups, and a decline in complex I activity, increased citrate synthase activity and reactive oxygen species generation, and greater mitochondrial fragmentation within the proximal tubular epithelial cells, were reported. Lowering blood glucose to MHG improved the mitochondrial phenotype. Quantitative mitochondrial proteomics revealed a clear change in the mitochondrial signature induced by hyperglycaemia. Distinct differential landscapes were observed in SHG rats compared with control and MHG; gene ontology analysis showed upregulation of ketone, fatty acid, and glutathione metabolic processes, and downregulation of transmembrane transport and protein translation pathways. These data demonstrate that hyperglycemia induces alterations in mitochondrial bioenergetics and the mitoproteomic landscape, and correlate with the severity of renal lesion, suggesting that severe hyperglycaemia is a key determinant of mitochondrial homeostasis in DKD.

548

Quantitative characterisation of changes in the proteome of head & neck squamous cell carcinomas under hypoxia

Yongchuan Gu[1], George Guo[1], Leo Payne[1], Frederik Pruijn[1], Bill Wilson[1]

1. University of Auckland, Grafton, AUCKLAND, New Zealand

Hypoxic tumour tissue is known to be resistant to treatment and is associated with a poor clinical prognosis. There are several reasons why this might be, including the capacity of hypoxia to drive genomic instability and alter DNA damage repair pathways. Head and neck squamous cell carcinoma (HNSCC) has been found to be a complex group of malignancies, with notable features that include a marked effect of hypoxia on treatment outcomes in standard-of-care chemoradiation and evidence for hypoxia-induced immunosuppression. UT-SCC-54C, a HNSCC line, was used as a representative to characterise the proteome-wide changes under hypoxic culture.

In order to generate a comprehensive peptide mass spectral library, whole-cell lysates from log-phase cultures growing under standard (aerobic) culture conditions were collected, digested with trypsin, and fractionated by strong cation exchange column for deep protein coverage to facilitate protein identification. Each fraction was spiked with iRT standards (retention time calibrant) and analysed in a data-dependent acquisition mode. Cells cultured under several chronic hypoxic conditions were then analysed together with those cultured under standard conditions using a TripleTOF system in data-independent acquisition mode (SWATH) to obtain quantitative data.

More than 6,000 individual proteins were identified and this formed the basis of the mass spectral library for the cell line. This spectral library was further refined to a peptide ion library for SWATH data extraction (using Skyline). The SWATH-MS data identified about 3,700 proteins and 12,800 peptides in each sample (condition), suggesting good protein coverage depth.

Statistical analysis using Tukey Median Polish analysis (MSStats) revealed some key protein changes under chronic hypoxia (compared to air). As expected, proteins involved in HIF-1 signalling pathways (e.g. ALDOA, PDK1, PGK1 and SLC2A1) were significantly increased. Interestingly, several proteins with potential roles in genomic stability (c7orf49 and UHRF1) were downregulated under hypoxia and these are currently being investigated.

549

A novel strategy for deep N-glycomics

Yudong Guan[1], Christoph Krisp[1], Morten Thaysen-Andersen[2], Nicolle Packer[2], Pengyuan Yang[3], Hartmut Schlüter [1]

1. Universitätsklinikum Hamburg-Eppendorf (UKE), Hamburg, Germany

2. Macquarie University, Sydney, Australia

3. Fudan University, Shanghai, China

Glycan biosynthesis is affected by disease states more significantly than protein, which shows greater potential to develop as biomarker. Therefore, qualitative and quantitative analysis is indispensable for glycomics and mass spectrometry (MS) has been the most powerful analytical tool in this hot field. However, the native glycan has lower ionization efficiency and produces more complex fragments than peptide. To solve those problems, permethylation has been developed as the most efficient derivative approach with improved ionization efficiency and simple fragments in MS analysis. In this study, we further optimize the solid-phase permethylation by different parameters and develop a novel strategy for N-glycomics to match the experimental data with theoretical database by R-scripts, which increase significantly the number of identifiable N-glycans with isotope-based data quality control. Furthermore, a novel bundled sequencing algorithm is designed to identify the N-glycoforms at MS2 levels. By this strategy, 57 N-glycan species (133 N-glycoforms) from ovalbumin, 90 N-glycan species (162 N-glycoforms) from etanercept, 245 N-glycan species (395 N-glycoforms) from human acute promyelocytic leukemia cells and 343 N-glycan species (833 N-glycoforms) from corpus callosum of adult mouse are identified. The identified N-glycans are verified by pGlyco software. This strategy is also applicative on O-glycomics. Besides, stable isotopic labeling and label-free quantification are performed for N-glycomics, in which we prove that glycan is more sensitive than its related protein after bioinformatic analysis. Finally, this study provides a novel pathway for N-glycomics to realize deep identification and biomarker discovery.

55

Combined glycan and extracellular matrix protein imaging mass spectrometry workflows for FFPE prostate cancer tissues

Richard R Drake[1], Connor West[1], Kim Norris-Caneda[1], Fred David[1], Jennifer Bethard[1], Anand S Mehta[1], Peggi M Angel[1]

1. Medical University of South Carolina, Charleston, SC, United States

Malignant transformation of the extracellular matrix (ECM) disrupts the homeostatic tissue microenvironment, altering tissue processes of adhesion, cell death, migration, and proliferation to promote tumor growth. The cellular glycocalyx, comprised of multiple types of glycans represented by N-linked and O-linked glycoproteins, glycosphingolipids, and glycosaminoglycans, in turn interacts with multiple ECM proteins in the stroma, as well as immune system components. We hypothesize that characterization of co-localized glycosylation and ECM protein changes in the prostate tumor microenvironment could lead to novel diagnostic and therapeutic biomarkers of PCa across the clinical spectrum of the disease. A combined MALDI imaging mass spectrometry analysis workflow has been used to characterize FFPE prostate tissue slices representative of the clinical spectrum of PCa progression. Tissues are processed initially for release of N-glycans by spraying a molecular coating of peptide N-glycosidase F (PNGase F) enzyme via solvent sprayer, followed by analysis on a 7T MALDI-FTICR mass spectrometer. Analytes and matrix are removed, followed by digestion of the tissue with sprayed bacterial collagenase, and additional analysis by MALDI-FTICR MS. Tissue images are processed using FlexImaging and SCiLS Lab software. Detected glycan masses are cross-referenced with an in-house N-glycan structural database. Total numbers of N-glycans detected for PCa tissues are generally n = 60-70. Additional glycan analysis of adjacent tissue slides is feasible to assess sialic acid N-glycan isomers using chemical amidation, and fucosylated N-glycan isomers using a core fucose specific endoglycosidase, Endo F3. Collagenase-derived ECM peptides are compared to an in-house database of prostate ECM peptides, determined separately by LC-MS, which currently represents 87 individual PCa tissue proteins (68% ECM proteins with 18 collagen sub-types, 10% membrane proteins). Multiple MS images can be overlayed with H&E stained tissue to histopathology co-localize N-glycans and ECM proteins associated with tumor, stroma, hyperplasia and normal gland regions.

550

NQO1 is a determinant for cellular sensitivity to anti-tumor agent napabucasin

Gaigai Guo[1], Zhouyong Gao[2], Mengsha Tong[1, 3], Dongdong Zhan[1, 4], Guangshun Wang[2], Yi Wang[1, 2], Jun Qin[1, 2]

1. State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Changping District, BEIJING, China

2. Joint Center for Translational Medicine, Baodi Hospital, Tianjin, China

3. School of Life Sciences, Tsinghua University, Beijing, China

4. Department of Life Sciences, East China Normal University, Shanghai, China

Napabucasin (NAPA) has been shown as a potent cancer stemness inhibitor that has demonstrated promising activity towards cancer of different types in early phase clinical trials despite satisfactory results in phase III trials. Unfortunately, two recent phase III NAPA clinical trials failed to meet the primary endpoint of overall survival. The reason for the failure is not clear, but possible

ways to rescue the clinical trial are to stratify patients with biomarkers that could predict NAPA response, and/or to seek alternative strategies to increase the NAPA efficacy. Here, we report the identification of NAD(P)H dehydrogenase 1 (NQO1) as a major determinant for NAPA efficacy by a proteomic approach. In vitro evidence showed that NAPA is a substrate for NQO1, which mediates the generation of ROS for cell death. We also demonstrate that inducing NQO1 expression by pre-exposure to IR or activation of its transcription factor NRF2 by drugs, including one approved FDA, can increase the cytotoxicity of NAPA. Our findings suggest the use of a companion diagnostic test to identify tumors with high NQO1 expression in future NAPA trials and the strategies for expanding the application of NAPA-based regimens for human cancer therapy.

551

Insights into the proteomics of abiotic stresses in rice

Fatemeh Habibpourmehraban[1], Sara Hamzelou[1], Brian Atwell[2], Paul Haynes[1]

1. Department of Molecular Science, Macquarie university, Sydney, NSW, Australia

2. Department of Biology, Macquarie university, Sydney, NSW, Australia

Rice as a cereal crop species is a significant part of the staple diet for half of the world's population and is grown in every continent apart from Antarctica. Reduced rice crop productivity is mostly attributed to various abiotic stresses, which are a major area of concern when we are faced with increasing food requirements. The major abiotic stresses include drought, oscillating temperature and high salinity, all of which negatively influence the yield of crops.

In nature it is unlikely that plants are exposed to abiotic stresses in isolation. Plants respond to multiple stresses differently from how they respond to individual stresses, activating a specific program of gene and protein regulation relating to the exact environmental stress. Rather than being additive, the presence of an abiotic stress can have the effect of reducing or enhancing the susceptibility to other abiotic stresses. In this project, we are examining the combined effects of drought, temperature and salt stress in different permutations.

As a baseline study, we have investigated the proteomic response to drought stress in eight different rice varieties; Nipponbare, Doongara, IAC1131, Mahsuri, Reiziq, N22, IR2006-P12 and IDSA77. Plants were grown in a temperature controlled greenhouse to late vegetative stage and then exposed to drought stress, with leaf samples collected at the point of severe stress, and also following recovery. For initial studies of combination of multiple abiotic stresses, Nipponbare plants were subjected to simultaneous salt, drought and temperature oscillation stresses. Proteins were extracted from leaf tissue, with in-gel and in-solution digested peptides separated and identified using nanoflow reversed-phase liquid chromatography – tandem orbitrap mass spectrometry on a Thermo Q-exactive. Peptides and proteins were identified using Proteome Discoverer and GPM software. Our results will be useful for the design of agronomically relevant strategies for the development of broad spectrum stress tolerant crops.

552

Effect of mAb carbohydrate composition on Fc:CD16a interaction

Joanna Hajduk[1], Sebastian Malik[2], Jana Bangerter[1], Marco Thomann[2], Dietmar Reusch[2], Renato Zenobi[1]

1. Department of Chemistry and Applied Biosciences, ETH Zurich, Zürich, Switzerland

2. Pharma Technical Development Penzberg, Roche Diagnostics GmbH, Penzberg , Germany

Minor changes in the quality of biologically manufactured monoclonal antibodies (mAbs) can affect their bioactivity, efficacy and safety. One of the most important variation concerns the N-glycosylation pattern, which has direct impact on an anti-tumor mechanism called antibody-dependent cell cytotoxicity (ADCC). Thus, careful engineering and production of the mAbs is expected to enhance protein-receptor binding and ADCC. The specific aim of this study is to evaluate the influence of terminal carbohydrates within the Ab Fc fragment on the interaction with CD16a receptor in native and label free conditions.

Here, we apply native electrospray ionization mass spectrometry (ESI-MS) to study mAb binding to a recombinant CD16a extracellular domain. The mAb molecules comprise the variants with minimal and maximal galactosylation as well as α2,3 and α2,6-sialic acid isomers. Direct ESI-MS titration allows us to determine the solution-phase antibody-receptor equilibria and by using a temperature-controlled nanoelectrospray source the thermal stability of the complex is examined. Additionally, the system is validated with a protein-cell assay using genetically modified human natural killer (NK) cell line expressing high affinity CD16a.

Based on these, we prove that galactosylation of the fucosylated Fc fragment increased the CD16a binding affinity by 1.5-fold when compared with non-galactosylated variant. The α2,6-sialylation has no significant effect on binding affinity, whereas the α2,3-sialylation decreases it by 1.72-fold. In line with expectation, the galactoslylated and α2,6-sialylated mAb:CD16a complex exhibit higher thermal stability when measured using a temperature gradient from 20 to 50˚C. A similar binding pattern is observed based on indirect staining approach using NK cells.

The results of our study deliver a profound insight into Ab:CD16a equilibria and shed new light on the Fc fragment N-glycosylation chemistry. Moreover, a new perspective for native ESI-MS approach in analysis of highly heterogeneous proteins is proposed.

553

Plasma protein signatures in dilated cardiomyopathy patients

Elke Hammer[1], Marcus Dörr[1], Vishnu M Dhople[1], Kristin Lehnert[1], Sabine Ameling[1], Kerstin Weitmann[1], Stephan B Felix[1], Uwe Völker[1]

1. University Medicine Greifswald, Greifswald, MV, Germany

Background: Dilated cardiomyopathy (DCM) is an important cause of heart failure. Comprehensive data on blood plasma protein profiles associated with this disease are still rare. Therefore, we profiled samples of a DCM patient cohort to derive molecular markers that are related to specific cardiovascular phenotypes and mortality.

Methods: A retrospective analysis of a registry of 610 patients with non-familial DCM enrolled at the University Medicine Greifswald from 2004-2012 was performed. Quantitative protein intensities obtained by tandem mass spectrometry (MS) after depletion of six high abundant plasma proteins were used for the analysis of associations between circulating protein levels and phenotypes (e.g. body mass index (BMI), disease duration, mortality, left ventricular ejection fraction (LVEF) and left ventricle diameter at diastole corrected for body surface area (LVEDDI)).

Results: The analysis of plasma samples allowed the collection of quantitative data of 341 circulating proteins. The reproducibility of MS data acquisition was monitored using a quality control plasma pool displaying a MS analysis variance of 12% and total technical variance of 22%. Results for proteins being associated with age, sex, and BMI correlate with already published data and confirm the plausibility of the data obtained. In addition, linear regression analyses revealed only two proteins being associated with LVEDDI and a robust protein signature associated with systolic heart function (LVEF) as well as mortality. Results for associations with self-reported disease duration were less significant, probably due to the large range of values (0-30 months) and the unequal distribution among patients (90% with disease duration ≥6 months).

Conclusions: Proteomic profiling of plasma from a cardiovascular DCM patient cohort allowed generation of high quality protein data which are used for integrative analysis with other OMICS data and can be utilised for the identification of biomarkers for prediction of outcome in longitudinal studies.

554

Comparison of proteomes of three species of rice from different ecological backgrounds in drought stress

Sara Hamzelou[1], Karthik Kamath2[2], Farhad Masoomi Aladizgeh[3], Mehdi Mirzaei[2], Brian J Atwell[3], Paul A Haynes[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Australian Proteome Analysis Facility, Macquarie University, Sydney, NSW, Australia

3. Department of Biological Sciences , Macquarie University, Sydney, NSW, Australia

Drought is a major hazard for world rice production. Substantial genetic variation in drought tolerance has been reported within rice ecotypes. Proteomic analysis provides insights into the expression of key genes by contrasting diverse germplasm that has been exposed to the steady-state drought. In this study, we analysed the proteome of leaves of three species of rice grown principally in Asia, Africa and Oceania. Commercial Asian cultivar (*Oryza sativa;* Nipponbare), Australian (*Oryza australiensis*), and African (*Oryza glaberrima*) rice plants were exposed to moderate drought stress and leaves from stressed and untreated control plants were harvested for protein extraction, followed by label-free quantitative shotgun proteomics analysis.

Leaf water potential was measured in plants that were well-watered throughout or exposed to constant but moderate water deficits. Proteins were extracted from leaf tissue using trichloroacetic acid – acetone extraction and precipitation. In-solution digested peptides were separated and identified using nanoflow reversed-phase liquid chromatography – tandem orbitrap mass spectrometry on a Thermo Q-Exactive. Peptides and proteins were identified and quantified using MaxQuant.

Oryza australiensis is tolerant to variable water supply in nature and in our controlled experiments. However, this response was not observed in *O. glaberrima* which is also known as a stress tolerant species. Stress-responsive proteins were up-regulated in all three species, a number of which were unique to *O. glaberrima*. Interestingly, preliminary analysis has shown that the identified differentially expressed proteins were assigned to variant biochemical pathways showing the different strategies of each species to modulate molecular responses to overcome water stress.

555

Proteogenomics analysis of EGFR mutations in non-small-cell lung cancer

Chi-Ting Lai[1], Sung-Liang Yu[2], Ya-Hsuan Chang[3], Hsuan-Yu Chen[3], Yu-Ju Chen[4], Chia-Li Han[5]

1. Genome and Systems Biology Degree Program, National Taiwan University, Taipei, Taiwan

2. Department of Clinical Laboratory Sciences and Medical Biotechnology, National Taiwan University, Taipei, Taiwan

3. Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

4. Institute of Chemistry, Academia Sinica, Taipei, Taiwan

5. Master Program in Clinical Pharmacogenomics and Pharmacoproteomics, Taipei Medical University, Taipei, TAIPEI, Taiwan

Oncogenic *EGFR* mutations, in particular those in tyrosine kinase domain, have been confirmed to associate with sensitivity to tyrosine kinase inhibitors (TKIs) and are indicators for prescribing TKIs in non-small-lung cancer (NSCLC). However, patients received the TKIs showed diverse responses, indicating the functional restriction of genetic test to guide the TKI prescription. Several studies reported that the expression levels of mutations in DNA, RNA and proteins levels are not always quantitatively correlated. Thus, we aimed to develop a mass spectrometry (MS)-based proteogenomics strategy for multiplexed screening of somatic EGFR mutations in protein level. The proteogenomics strategy integrated bioinformatics analysis of mutant EGFR protein sequences, affinity purification of EGFR protein complex, parallel enzymatic gel-assisted digestions, LC-DDA-MS/MS and LC-

PRM-MS analyses, customized database searching using multiple engines, and construction of reference intervals using synthesized isotopic standard peptides, for unambiguous identification and absolute quantification of wild-type and mutant EGFR proteins. The strategy offered unambiguous identification of peptides covering 34 mutated and 33 wild-type sites in EGFR as well as EGFR-interacting proteins. The quantitation results revealed concomitant and heterogeneous expressions of mutated and wild-type EGFR proteins in a series of NSCLC cells and xenograft tumors that harbored different *EGFR* genotypes, suggesting diverse and heterogenous expressions of the same EGFR mutations in individual tumors. Our developed proteogenomics strategy can robustly determine the multiple somatic mutations in EGFR at protein level which may provide a better prediction on adaption to TKI treatments and improve the understanding towards the molecular impact of *EGFR* mutations during cancer progression.

556

Quantifying biological change in shotgun proteomics experiments using newly developed software

David CL Handler[1], Abdulrahman Mansour M Shathili[1], Paul A Haynes[1]

1. Molecular Sciences, Macquarie University, Sydney, NSW, Australia

Discovery shotgun proteomics often yields large, noisy data for the researcher to organise and filter to produce quality results. Our laboratory employs a method of replicate summing with spectral counting to filter out low quality and non-reproducible peptide matches to improve the quality of downstream quantitative analysis. This is especially important for our research where our samples range in complexity from ancient skin samples to grape cell culture to mice retinal tissue. To facilitate these data handling processes, we have developed several pieces of software that are freely available for use.

Same-Same analysis helps to correct overzealous Multiple Testing Corrections. Using permutation analysis, Same-Same will take six replicates of PSM-searched data and calculate a modified Benjamini-Hochberg cut-off value. Using this method, researchers can make more informed decisions at the discovery stage about which proteins are suited to quantitation.

PeptideWitch produces high stringency data from PSM-searched results; in addition to conducting the Same-Same process, Peptide Witch will perform control vs treatment analysis on input data. Data is converted to highly stringent and reproducible results before being quantified, with data presented in Heatmaps, PCA charts, Venn Diagrams and Excel outputs.

PeptideMind helps validate control vs treatment quantitation using machine-learning assisted protocols. Users will upload two states with six replicates each and PeptideMind will create a 400-strong permutation analysis for differentially regulated protein IDs. The program is able to detect outliers and highlight reproducible results using a consensus of four different machine learning algorithms, thus acting as a form of statistical validation of quantitation procedures.

557

Evaluation and Minimization of Over-alkylation in Proteomic Sample Preparation

Zhaomei Shi[1], Chengqian Zhang[1], Ying Han[1], Yan Ren[2], Piliang Hao[1]

1. Shanghai Tech University, Shanghai, SHANGHAI, China

2. BGI-Shenzhen, Shenzhen, China

Reduction and alkylation are essential steps in shotgun proteomic sample preparation, but over-alkylation can occur on peptide N-terminus and amino acid residues other than cysteine, which adversely affects protein identification and quantification. To date, different reduction and alkylation conditions are used in different laboratories, but a systematic evaluation of the extent of over-alkylation from these different protocols has not been done yet. In this study, we comprehensively evaluated various sample preparation protocols that used different denaturants, different digestion buffers and different concentrations of reduction and alkylation reagents. N-term carbamidomethylation occurred on 2-11% of the identified peptides in some commonly used protocols, and it increased over 40 times as the concentration of dithiothreitol and iodoacetamide increased from 1mM/4mM to 20mM/80mM. The remaining excessive iodoacetamide was proven to be the major cause of over-alkylation during trypsin digestion. The use of digestion buffers at pH 6 reduces over-alkylation, but greatly introduces some other artificial modifications. In conclusion, it is feasible to minimize over-alkylation by using a few mM of dithiothreitol and iodoacetamide at the ratio of 1:2 for reduction and alkylation or quenching the excessive iodoacetamide with dithiothreitol before digestion. This study greatly improves proteomic sample preparation in the aspect of reduction and alkylation.

558

Determining the specificity of an anti-PD-1 antibody using a new bead array format covering about 80% of the human proteome

Matthias Harbers[1], Shusei Sugiyama[1], Miwako Denda[1], Satoshi Ozawa[1], Ryo Morishita[2]

1. CellFree Sciences Co., Ltd., Yokohama, Kanagawa , Japan

2. CellFree Sciences Co., Ltd., Matsuyama, Ehime, Japan

Anti-programmed cell death protein 1 (PD-1/PDCD1/CD279) is a suppressor of the immune system playing an important role in preventing autoimmunity. Anti-PD-1 antibodies are therefore used to enable immune responses during tumor therapy. We used a commercial monoclonal anti-PD-1 antibody against the human PD-1 protein as an example to establish routines for proteome-wide antibody specificity testing applying our new bead array format representing ~80% of the human proteome. This format allows parallel in-solution testing of many more proteins than possible in any individual cell line experiment where proteins on the array are presented in a similar concentration range. For testing the specificity of a given antibody, our process comprises three steps: First, we confirm the binding conditions and detection method for the antibody using a set of positive and negative controls. Second, the specificity of the antibody is analyzed on the bead array in a special 1536-well plate format. In each well 14 different

proteins are pooled allowing for quick overviews on the specificity of the antibody. This bead array format may also be used for further monitoring the specificity of different antibody batches or testing different secondary antibodies and detection methods. Third, to identify the actual targets that possibly cross-react with the antibody, we individually test each protein from the wells found positive in the initial screen on the 1536-well plate. This process led to the identification of three unrelated proteins that cross-reacted with the anti-PD-1 antibody. Alignments of the amino acid sequences from those cross-reacting proteins revealed a common epitope that is shared with human PD-1 protein thus demonstrating that our process and the new bead array platform are providing very accurate conditions for testing antibody specificity on a proteome-wide scale. This is an essential process when further characterizing antibodies for use in research, diagnostics or even therapy.

559

Small-protein enrichment assay enables the rapid, unbiased quantification of over 100 low abundance factors from human plasma

Dylan J Harney[1], Amy T Hutchison[2], Zhiduan Su[1], Luke Hatchwell[1], Leonie K Heilbronn[2], Samantha Hocking[3], David E James[1], Mark Larance[1]

1. School of Life and Environmental Sciences, University of Sydney, Camperdown, NSW, Australia

2. Discipline of Medicine, University of Adelaide, Adelaide, South Australia, Australia

3. Central Clinical School, Faculty of Medicine and Health, University of Sydney, Camperdown, NSW, Australia

Unbiased and sensitive quantification of low abundance small proteins in human plasma (e.g. hormones, immune factors, metabolic regulators) remains an unmet need. These small protein factors are typically analysed individually and using antibodies that can lack specificity. Mass spectrometry (MS)-based proteomics has the potential to address these problems, however the analysis of plasma by MS is plagued by the extremely large dynamic range of this body fluid, with protein abundances spanning at least 13 orders of magnitude. Here we describe an enrichment assay (SPEA), that greatly simplifies the plasma dynamic range problem by enriching small-proteins of 2-10 kDa, enabling the rapid, specific and sensitive quantification of >100 small-protein factors in a single untargeted LC-MS/MS acquisition. Applying this method to perform deep-proteome profiling of human plasma we identify C5ORF46 as a previously uncharacterized human plasma protein. C5ORF46 is expressed in either major blood vessels, or the heart, and whose role may be positively correlated with HDL particle abundance and reverse cholesterol transport. We further demonstrate the reproducibility of our workflow for low abundance protein analysis using a stable-isotope labelled protein standard of insulin spiked into human plasma. We applied this method to study the intermittent fasting response in an overweight human female cohort and observed several unexpected changes including; decreased plasma abundance of the iron homeostasis regulator hepcidin, decreased VLDL-associated protein APOC4 and increased plasma osteopontin. Thus, SPEA provides the ability to study numerous important, low-abundance hormones in a single rapid assay by overcoming much of the dynamic range problem in plasma proteomics, which makes it a promising tool for plasma protein discovery and quantitative assay development.

56

Imaging age-induced perturbations of mitochondrial function, neurotransmission and lipid signaling in specific brain structures

Per E. Andren[1], Theodosia Vallianatou[1], Elva Fridjonsdottir[1], Mohammadreza Shariatgorji[1], Anna Nilsson[1], Maria Karlgren[2], Heather Hulme[1], Per Svenningsson[3]

1. Dept. of Pharmaceutical Biosciences, Medical Mass Spectrometry Imaging, Uppsala University, Uppsala, Sweden

2. Dept. of Pharmacy, Uppsala University, Uppsala, SE, Sweden

3. Dept. of Neurology, Karolinska Institutet, Stockholm, Sweden

Aging is considered as the major risk factor for neurodegenerative diseases, such as Alzheimer's (AD) and Parkinson's disease (PD), constituting also the non-pathological (normal) aging process a substantial and very important part of neurological studies.

Here, we simultaneously imaged multiple metabolic pathways, including the dopaminergic, noradrenergic, serotonergic, histaminergic, and GABAergic systems, significantly altered by normal aging in detailed structures of mouse brain tissue sections by using MALDI-mass spectrometry imaging (MSI). We examined the interplay between aging and the response to tacrine-induced acetylcholinesterase (AChE) inhibition. The cholinergic system has a central role related to normal cognition and age-related cognitive decline, including dementias such as AD, where a progressive loss of limbic and neocortical cholinergic innervation is observed. Indeed, treatment with AChE increases the availability of acetylcholine (ACh) in the brain. In addition, we studied the neuropeptide and neurotransmitter system changes in animal models of PD.

Carnitine and acetylcholine/choline metabolism were the major pathways affected by aging and acetylcholinesterase inhibition in striatum, hippocampus and cortical areas. Age-related disrupted lipid metabolism was observed in striatum, cortex and specific hippocampal substructures. An age-induced increase of endogenous antioxidants, such as α-tocopherol in the hippocampus, was detected. Structure specific age-related changes were found in multiple monoaminergic pathways, e.g., dopamine and 3-methoxytyramine were found decreased in ventral pallidum, 3,4-dihydroxyphenylacetaldehyde levels were increased in caudate-putamen, and norepinephrine was decreased in hippocampus and somatosensory cortex. Using a unilateral 6-hydroxydopamine rat model of PD we simultaneously imaged >20 neuropeptides and found changes in enkephalin, dynorphin, tachykinin and neurotensin neuropeptides caused by the dopamine denervation and L-DOPA treatment in multiple brain regions.

Our comprehensive imaging approach showed significant and novel age-induced and region-specific metabolic perturbations in mitochondrial function, neurotransmission, and lipid signaling, not always attenuated by the acetylcholinesterase inhibition.

560

Benchmarking PEA-Olink technology with mass spectrometry-based protein quantifications in blood plasma

Stefanie M Hauck[1], Agnese Petrera[1], Marianne Sandin[2], Cornelia Huth[1], Barbara Thorand[1], Fabian Metzger[1], Christian Gieger[1], Christine von Toerne[1]

1. Helmholtz Center Munich, Munich, GERMANY, Germany

2. Olink Proteomics, Uppsala, Sweden

The latest breakthrough in biomarker discovery is the proximity extension assay (PEA) technology developed by Olink (Sweden). This antibody-based technology allows for relative quantification of 1165 human protein biomarkers in plasma and other biological samples. Helmholtz Center Munich hosts the first platform certified for Olink technology in Germany. Upon integration of this platform in our proteomics facility, we aim to correlate PEA-based profiling with mass spectrometry-based measurements (MS) in human plasma.

We used 370 plasma samples from the Cooperative Health Research in the Region of Augsburg (KORA) study, a population-based longitudinal study. These samples are analyzed by PEA (10 different Olink panels), data-independent acquisition (DIA)-MS, targeted SRM-MS, and SOMAscanAssay V3.2 (SomaLogic).

We have accumulated a DIA-MS dataset from 370 non-depleted human plasma samples quantified by matching to a comprehensive in-house spectral library covering 1924 proteins (based on 16 918 peptides). When comparing the proteins quantified by DIA-MS (direct identification above threshold in >75% of samples, 1% protein FDR filtered) to the complete set of markers measured by PEA (valid data in >95% of samples), we observe an overlap of 30 proteins. As expected, the overlapping proteins are mainly contained in Olink panels enriched for high abundant proteins. We find convincing correlations between PEA and DIA-MS in the overlapping proteins. Further, correlation to overlapping SRM-MS measurements are excellent, exemplified by MBL2, which correlates at 0.87 and 0.82 (Spearman) when comparing PEA to DIA-MS and SRM-MS, respectively. Additional correlations between technologies, including SomaLogic will be presented. Further, correlations between genetics and protein abundances (pWAS) measured by the different technologies will be discussed.

PEA provides a highly complementary technology to mass spectrometry-based proteomics, overcoming the notorious challenge of low identification and quantification rates in plasma samples. The high correlation of proteins detected by both technologies mutually confirms the methods for biomarker discovery.

561

The lung cancer “breathalyser”: using non-invasive breath testing for diagnosis of lung cancer

Sarah A Hayes[1, 2], Ben Harris[2, 3], Hannah Kamitakahara [1, 2], Ramyashree Prasanna Kumar[1, 2], Pascal Steffen[2, 4], Christoph Krisp [5], Stephen J Clarke[1, 2, 6], Mark P Molloy[2, 4], Nick Pavlakis[1, 2, 6], Viive M Howell[1, 2]

1. Bill Walsh Translational Cancer Research Laboratory, Hormones and Cancer, Kolling Institute, St Leonards, Sydney, NSW, Australia

2. Sydney Medical School, Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia

3. Department of Respiratory Medicine, Royal North Shore Hospital, St Leonards, Sydney, NSW, Australia

4. Bowel Cancer & Biomarker Laboratory, Hormones and Cancer, Kolling Institute, St Leonards, Sydney, NSW, Australia

5. Australian Proteome Analysis Facility, Macquarie University, North Ryde, Sydney, NSW, Australia

6. Department of Medical Oncology, Royal North Shore Hospital, St Leonards, Sydney, NSW, Australia

Early diagnosis of cancer provides the greatest chance of a cure or favourable Analysis of exhaled breath condensate (EBC) has recently been proposed as a non-invasive method to diagnose early-stage lung cancer. Since this sample type has not been comprehensively profiled before; here, we investigated variables that may affect protein yield in EBC, prior to constructing protein profiles of human breath from a lung cancer patient and a healthy control.

Using EBC from healthy volunteers, we assessed the effect on total EBC protein of storage tube, method of sample concentration, cooling temperature, use of protease inhibitors and filters/nose-clips and the effect of delaying processing for up to 24hrs. Total protein was quantitated using the Protein Assay Kit on a Qubit 3.0 Fluorometer. Finally, protein mapping of EBC from a patient with NSCLC adenocarcinoma and from an age- and gender-matched control volunteer was performed by Information Dependent Acquisition MS on a Sciex 6600 TripleTOF.

We determined that EBC should be collected at lower cooling temperatures (-80°C) for greatest EBC volume and protein yield. EBC should be stored in plastic with the addition of protease inhibitors to ensure stability during sample processing as well as long-term stability of the sample (assessed up to 2 years post collection). In a preliminary MS profile, 57 proteins were identified, including keratins, mucins and inflammatory proteins. Differentially expressed proteins (n=21, FC>1.5) between lung cancer and control EBC include several serpins (A3/B4/B7/B12), S100 proteins (A7/A11/A14), as well as proteases and lipocalins known for their involvement in lung cancer development and progression.

Understanding the factors that affect EBC yield and quality is crucial for the generation of reproducible and accurate EBC proteomic profiles. Improving our knowledge how expression of these lung proteins change during carcinogenesis is crucial to provide an “exhaled biomarker fingerprint” of lung cancer.

562

The dynamic mitochondrial complexome

Juliana Heidler[1], Heiko Giese[2], Jana Meisterknecht[1], Sarah Stenton[3, 4], Holger Prokisch[3, 4], Ilka Wittig[1]

1. ZBC - Functional Proteomics, Goethe University Frankfurt, Frankfurt/M., Germany

2. Molecular Bioinformatics, Goethe University Frankfurt, Frankfurt/M., Germany

3. Institute of Human Genetics, Technische Universität München, München, Germany

4. Institute of Human Genetics, Helmholtz Zentrum München, München, Germany

Many patients with mitochondrial disorders suffer from impaired assembly of mitochondrial protein complexes due to defects in genes encoding assembly factors. The sequence of protein complex assembly was intensively studied in proliferating cells. These studies mostly reflect the de-novo assembly and give only limited information of the protein complex dynamics in differentiated cells and tissues. The combination of blue native electrophoresis with quantitative mass spectrometry identifies even scarce sub-complexes, assembly intermediates, and complex remodeling. In this study, we combined complexome profiling and pulse stable isotope labeling of amino acids in cell culture (Pulsed-SILAC) to study the turnover and half-life of single proteins within protein complexes in differentiated post-mitotic C2C12-myotubes. The results represent a comprehensive data collection of dynamics in all mitochondrial protein complexes. Complete replacement e.g. for all protein complexes from the oxidative phosphorylation system requires about a month. We detected higher turnover rates between interface subunits of dimeric and oligomeric ATP synthase and supercomplexes suggesting more dynamics between as within complexes. In detail, we identified subunits of complex I with higher turnover rates in the parts of electron transport modules, suggesting quality control and repairing mechanisms within assembled complexes to ensure full bioenergetics function in differentiated cells. Application of the developed method to patients with a mitochondrial disorder discovered a novel factor involved in quality control and repair of complex I.

563

Towards personalized diagnostics via longitudinal study of the human plasma *N*-glycome

René Hennig[1], Samanta Cajic[2], Udo Reichl[2, 3], Erdmann Rapp[1]

1. glyXera GmbH, Magdeburg, SACHSEN-ANHALT, Germany

2. Bioprocess Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Sachsen-Anhalt, Germany

3. Chair of Bioprocess Engineering, Otto-von-Guericke University, Magdeburg, Sachsen-Anhalt, Germany

Facilitated by substantial advances in analytical methods, plasma *N*-glycans have emerged as potential candidates for biomarkers. In the recent years, several investigations could link aberrant plasma *N*-glycosylation to numerous diseases. However, due to often limited specificity and sensitivity, only a very limited number of glycan biomarkers were approved by the authorities up to now.

The inter-individual heterogeneity of the plasma *N*-glycomes often mask small disease related changes in conventional large cross-sectional cohort studies, with a one-time per person sampling approach. This problem could be overcome by a longitudinal sampling approach, detecting already small changes during disease progression by monitoring the plasma *N*-glycome of individuals over time.

To evaluate this, we collected blood plasma samples of five healthy donors over a time period of up to six years. The plasma N-glycome was analyzed by utilizing multiplexed capillary gel electrophoresis with laser induced fluorescence detection (xCGE-LIF) as core technology [1]. By using this miniaturized electrokinetic separation technique with up to 96 capillaries in parallel, assisted by a 96-well format sample preparation and an automated data analysis by glyXtool™, a massive reduction of analysis time and costs per sample could be achieved.

We could demonstrate that the individual's *N*-glycome is remarkably stable over a period of several years and free of seasonal changes. Furthermore, we could demonstrate that the inter-individual differences of the *N*-glycome are enormous, but by looking at the progression of the plasma *N*-glycome of a single person, small changes could be detected and linked to lifestyle and environmental factors. Consequently, we could show that a longitudinal sampling approach (taking time-series from individuals) can be beneficial compared to large-scale studies, where small disease related changes in the *N*-glycome are often hidden within the inter-individual variation of the *N*-glycome [2].

1. [1] Hennig, R.; Rapp, E.; Kottler, R.; Cajic, S.; Borowiak, M.; Reichl, U.; N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF. Methods Molecular Biology (2015) 1331, 123-143.
2. [2] Hennig, R.; Cajic, S.; Borowiak, M.; Hoffmann, M.; Kottler, R.; Reichl, U.; Rapp, E.; Towards personalized diagnostics via longitudinal study of the human plasma N-glycome. Biochimica et Biophysica Acta - General Subjects (2016) 1860, 1728-1738.

564

Proteomics and immunoproteomics characterization of *Staphylococcus aureus* strains causing infective endocarditis

Christian Hentschker[1], Kristin Surmann[1], Stephan Michalik[1], Jawad Iqbal[2], Larissa Busch[1], Manuela Gesell Salazar[1], Tanja Meyer[1], Ulrike Mäder[1], Eeva Ruotsalainen[3], Vesa P. Kontinen[4], Stephan Fuchs[5], Barbara M. Bröker[2], Silva Holtfreter[2], Uwe Völker[1]

1. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany

2. Department of Immunology, University Medicine Greifswald, Greifswald, Germany

3. Division of Infectious Diseases, Inflammation Center, Helsinki University Central Hospital, Helsinki, Finland

4. Department of Bacteriology and Immunology, The Haartman Institute, University of Helsinki, Helsinki, Finland

5. Department of Infectious Diseases, Robert-Koch-Institute, Wernigerode, Germany

Background

Staphylococcus aureus is the leading cause of infective endocarditis (IE), a steadily increasing disease. IE is often associated with age and heart valve transplantation, but also affects intravenous (i.v.) drug users. The mortality rate of IE is high and therapy is complicated by the variety of *S. aureus* strains including resistances to several antibiotics. Aims of this study were to decipher bacterial proteins required to establish IE and screening for differences that might explain the varying outcome of normal IE patients (NDU) vs. i.v. drug users (IDU). Here, we used a multi-omics approach to study whether bacterial factors (genomics, proteomics) or the host's antibody response (immunoproteomics) impact the IE outcome.

Methods

The secretome and intracellular proteome of 25 *S. aureus* strains from IE patients (11 IDU, 14 NDU) and 25 strains from healthy carriers, which were matched for clonal clusters, were analyzed by nanoLC-MS/MS after cultivation in RPMI medium until stationary phase. Data analysis was done with a multi-strain database. The plasma of IE patients (28 days vs. disease onset) was analyzed by immunoproteomics using the xMAP® technology with a panel of 80 recombinant *S. aureus* antigens.

Results

Differences between IE causing strains and controls occurred in metabolism, cell surface proteins, proteins encoded by mobile genetic elements, and in the virulence factor repertoire. IE strains also displayed increased production of proteins involved in attachment to e.g., heart valves. Immunoproteomics revealed an increase in antibody titers against secreted toxins and immune evasion factors during the course of IE as well as differences between responses of IDU and NDU.

Conclusion

IE strains differ from colonizing strains in abundances of particular proteins. Moreover, higher anti-*S. aureus* antibody titers at disease onset as well as during infection might contribute to the better outcome of IDU patients.

565

Immunopeptidome analysis for metastatic cancer immunotherapy

SunHee Heo[1], HeeJin Lee[1], Gyungyub Gong[1]

1. Asan Medical Center, Seoul, SEOUL, South Korea

Immunotherapy has been considered a promising anticancer therapy of multiple tumors. Nevertheless, the really is that there is a lack of considerations and studies for metastatic cancers.

We have tried to establish the technology platforms for metastatic cancer immunotherapies. First, we had to know importance of the immune microenvironment with metastatic tumors. For this, we have tried to analysis for tumor-infiltrating lymphocytes and tertiary lymphoid structures in the primary breast tumors and metastatic sites in order to understand about tumor immune environments at various metastatic sites or between primary breast tumors and metastatic sites. Three hundred and thirty-five patients with metastatic breast cancers in the lung, liver, brain, or ovary, with available

hematoxylin and eosin-stained slides of metastatic sites (biopsy, n = 172; excision, n = 163), who were treated at the Asan Medical Center from January 2000 to March 2017, were included. Among the 335 cases, hematoxylin and eosin-stained slides of primary breast cancer were also available for evaluating tumor-infiltrating lymphocytes and tertiary lymphoid structures for 245 cases (biopsy, n = 37; excision, n = 208). Our results showed that metastatic breast tumors in the lung had more tumor-infiltrating lymphocytes than did tumors at other sites and matched primary tumors. In addition, the presence of tertiary lymphoid structures in metastatic sites is a critical factor for the level of tumor-infiltrating lymphocytes. These results suggested that immunotherapy would be useful for metastatic cancers.

Based on these results, we have planned to establish the technology platforms for identification of tumor specific antigen using metastatic tumor tissues. For this, we have collected the metastatic cancer tissues and matched adjacent normal tissue for analysis of somatic mutations and immunopeptidome. At the same time, we have tried to develop tumor specific antigen analysis method using data of whole genome sequencing and whole transcriptome sequencing with human leukocyte antigen peptidome data obtained from Liquid chromatography–mass spectrometry.

Ultimately, this study could provide a useful strategy for identifying tumor specific antigen which can be used cancer vaccine or/and tumor specific antigen responsive T-cell production for metastatic cancer immunotherapy.

566

Proteome changes that drive acute myeloid leukemia patients to relapse

Maria Hernandez-Valladares[2, 1], Elise Aasebø[1], Frode Berven[3], Frode Selheim[3], Øystein Bruserud[1]

1. Clinical Sciences, University of Bergen, Bergen, Hordaland, Norway

2. University of Sydney, Camperdown, NSW, Australia

3. Biomedicine, University of Bergen, Bergen, Hordaland, Norway

Acute myeloid leukemia (AML) is a hematological cancer that affects mainly the elderly and infants. Although complete remission (CR) is achieved for the majority of the patients after induction and consolidation therapies, nearly two-thirds relapse within a short interval. Therefore, understanding of the biological factors that determine relapse has become a major interest in clinical AML. In order to identify the proteins and their phosphorylated modifications involved in AML relapse, we performed a global proteome and phosphoproteome study by liquid chromatography mass spectrometry (LC-MS) with primary cells from 41 AML patients at time of diagnosis that were defined as RELAPSE or REL_FREE according to their relapse status after a 5-year clinical follow-up post diagnosis. Our findings show that RNA processing and V-ATP-ases proteins along with CDKs and CSK2 activities are increased in patients that will relapse. MS-based results were further validated with cell proliferation assays using bafilomycin A1, CX-4945, abemaciclib and SNS-032 inhibitors. Moreover, proteomics studies with eight matched diagnosis-relapse AML samples showed the persistent activities of RNA processing and DNA repair factors during relapse and uncovered the drug targetability of mitochondrial protein synthesis in resistant AML blasts. Our study presents molecules that could predict AML relapse and direct new therapeutic strategies that might circumvent more aggressive AML episodes.

567

Quantifying the lipidome for respiratory disease: A rapid and comprehensive targeted approach

Nyasha Munjoma[1], Giorgis Isaac[2], David Heywood[1], Lee Gethings[1], Robert Plumb[2]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Milford, MA, USA

Respiratory linked conditions associated with chronic obstructive pulmonary disease (COPD), asthma and infection are increasing with associated socio-economic costs. Recent reports have shown costs to exceed £11 billion per year for cases recorded in the UK. COPD in particular is a heterogeneous disease which is a major cause of illness and death worldwide. The combination of genetic and lifestyle factors are known to contribute towards increasing the probability of encountering the condition. Here, we describe a lipidomic approach to reveal molecular factors that may be involved in these biomolecular processes.

The analyses of plasma samples from three biological states of varying phenotype (control, COPD and asthma patients) were conducted. Lipid analysis was performed using LipidQuan, which is a streamlined and integrated lipidomic workflow. This platform consists of highly specific MRM transitions based on the fatty acyl chain fragments which were used for identification and quantification of multiple lipid species. Chromatographic conditions allowed for the separation of individual lipid classes with a complete analysis time of 8 minutes per sample. Data were processed using both TargetLynx and Skyline. Statistical analysis was conducted using SIMCA and additional data visualisation provided by MetaboAnalyst.

Biological significance of the results was established by merging the data from all experiments and performing pathway analysis. Statistical analysis of the data revealed clear separation between the various cohorts. Unsupervised PCA resulted in separation of healthy controls, COPD and asthma patients. Application of the metadata also revealed significant differences between smoking status, with subsets readily observed within the COPD population. Loadings analysis revealed FFA, LPC, PC and SM classes to be the main contributors to sample type clustering. Pathway analysis revealed a number of components related to inflammation, oxidative and immunity processes were identified as significant and associated with signalling, metabolic and regulatory pathways.

568

Characterization of cancer/testis antigens in non-small-cell lung cancer using transcriptomics and antibody-based proteomics

Feria Hikmet Noraddin[1], Max Backman[2], Patrick Micke[2], Cecilia Lindskog[1]

1. Department of Immunology, Genetics and Pathology, Human Protein Atlas, Uppsala University, Uppsala, Sweden

2. Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden

Cancer/testis antigens (CTAs) are a group of proteins ectopically expressed in a wide range of cancer types. The highest frequency of CTAs is found in non-small-cell lung cancer (NSCLC), ovarian cancer and melanoma. Because of their cancer specificity and immunogenic features, CTAs serve as potential targets for highly specific immunotherapy and cancer elimination.

In a previous study[1], 90 CTAs were identified based on a systematic approach using quantitative transcriptomics of tissues from 199 NSCLC patients compared to 142 normal tissues corresponding to 32 different organ sites. The CTAs were defined as having at least 5 times higher gene expression in NSCLC than any other tissue (excluding testis and placenta), and expression in at least in 2% of NSCLC cases. Thirty-five of the CTAs have previously been reported in the CTdatabase resource, while the remaining 55 were defined as novel CTAs. Here, we performed an in-depth evaluation of the CTA landscape of NSCLC by further analysis of the 90 CTAs on the protein level. The CTAs were categorized based on previous literature, level of protein existence (PE status) and cell type-specific expression using Human Protein Atlas immunohistochemistry data[2]. Ten CTA targets were chosen for further characterization by immunohistochemistry on a well-characterized NSCLC cohort to link the CTA protein profiles to both clinical and survival data, mutation pattern (82 genes) as well as nine different *in situ* immune profiles (PD-L1,

CD4, CD8, FOXP3, CD45Ro, CD20, CD138, CD163 and Npk146). This in-depth characterization provides a unique opportunity for extending the knowledge of CTAs in lung cancer and identification of novel potential targets for immunotherapy in NSCLC.

1. Djureinovic D et al. Profiling cancer testis antigens in non-small-cell lung cancer. JCI Insight. 2016 Jul 7;1(10):e86837. doi: 10.1172/jci.insight.86837.
2. Uhlén M et al. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419. doi: 10.1126/science.1260419.

569

Peptidomics workflow for urine biomarker discovery

Yoshitoshi Hirao[1], Amr Elguoshy[1], Suguru Saito[1], Bo Xu[1], Hidehiko Fujinaka[2], Shigeru Miyazaki[3], Keiko Yamamoto[1], Tadashi Yamamoto[1, 3]

1. Biofluid Biomarker Center, Niigata University, Niigata, Japan

2. National Hospital Organization Niigata Hospital, Kashiwazaki, Japan

3. Shinrakuen Hospital, Niigata, Japan

Many trials for biomarker discovery have been going on by mass spectrometry (MS) in plasma, tissue, urine or others. Now a day, technologies of MS are still developing to provide new machines and softwares. In fact, a number of proteins and peptides identified with high confidence by MS have increased a lot during these several decades.

The urine remains as one of the preferable biological samples for clinical examinations for diagnosis or monitoring of diseases, due to its non-invasive nature of collection. Some proteins and peptides, which are smaller than albumin, may be secreted in the urine from plasma through the glomerulus in the kidney. To find biomarkers for systemic organs in the urine, we aimed to develop a workflow to identify the urinary small proteins and native peptides, which had been processed in the body and thought to be novel biomarkers for many diseases. However, there are several difficulties for this such as less amounts of the small proteins and native peptides, no established protocol for collection and purification of these proteins and peptide, no qualitative and quantitative methods by MS.

In this study, we focused on mass spectrometry of the urinary small proteins and native peptides and established new protocols to propose a workflow for biomarker discovery by peptidomics. By our workflow more than 10,000 distinct peptides have been identified. These peptides were grouped as 500 peptide clusters; C-peptide cluster includes its intact form and also peptides forms truncated at the N- or C-terminal.

In conclusion, identification of the small proteins and peptides is still challenging due to no standard protocols for sample preparation, MS analysis methods. We proposed a new workflow for MS-based identification of native peptides in the urine to promote biomarker discovery in the near future.

57

Visualising the physiological biochemistry of human ocular lens transparency and cataract with imaging mass spectrometry

Gus Grey[1], Rebecca Perwick[1], George Guo[1], Nicholas Demarais[1], Mitchell Nye-Wood[1], Ivy Li[1], Julie Lim[1], Paul Donaldson[1]

1. University of Auckland, Auckland, New Zealand

To function as an effective optical element, the avascular lens exhibits several specialisations that include an ordered cellular structure, a gradient in refractive index, and a circulating current to deliver nutrients and remove waste products from the lens nucleus. The aging lens undergoes many changes to metabolites and proteins that alter these physiological and optical functions in specific lens regions, leading to the formation of age-related nuclear (ARN) cataract. A combination of human and laboratory-aged bovine lenses have been analysed by advanced mass spectrometry techniques to spatially map these changes and understand how they contribute to ARN cataract formation at the whole organ level.

A range of human lens ages (29y-82y) or bovine lenses, either laboratory-aged by hyperbaric oxygen treatment or organ cultured in artificial aqueous humour containing stable isotopically-labelled metabolites, were analyzed. Axial cryosections (20um) of lenses were collected on MALDI targets, and matrix applied by a TM-Sprayer. Positive ion mode MALDI-TOF (for proteins) or negative ion mode MALDI-FT-ICR (for metabolites) imaging mass spectrometry (IMS) was used to map the distribution of protein and metabolite distributions in each lens at 150um spatial resolution. SCiLS lab software was used to visualise and quantify age-related changes to mass spectral signals. GC-MS and LC-MS/MS approaches were used to validate distributions detected by IMS.

Several age-related modifications to proteins were observed in specific regions of the lens. For example, glutathione modified beta and gamma crystallins at specific sites, while a corresponding decrease in reduced glutathione levels was detected in the lens core. Additional alterations in lipid and UV filter levels and distributions were spatially mapped. Initial experiments to assess the role of the lens circulating current in lens metabolite transport established IMS as an imaging approach that could be utilised in developing future therapeutic interventions to delay the onset of lens cataract.

570

A resource for improved antibody validation

Andreas Hober[1], Kalle von Feilitzen[1], Anna Sjöberg[2], Tove Boström[2], Marie Utterbäck[2], Björn Forsström[1, 2], Fredrik Edfors[1], Mathias Uhlen[1]

1. Department of Protein Science, Science for Life Laboratory, KTH - Royal Institute of Technology, Stockholm, Sweden

2. Atlas Antibodies, Stockholm, Sweden

Antibodies are of great importance within many different areas of proteomics research as well as within clinical diagnostics. Using affinity reagents as the mean of detection puts a high demand on the reagents when it comes to selectivity and specificity. This has been illuminated by the International Working Group for Antibody Validation (IWGAV), which suggest that an antibody must be evaluated in an application-specific manner. When an antibody is to be used within a Western blot assay, the initial measure for evaluating its specificity is the theoretical molecular weight of its target protein, or rather the migration of a set of reference proteins during gel electrophoresis. It is well known that the migration of proteins during gel electrophoresis can differ from the standards used within the ladder. A dataset of migration patterns for more than 8,000 proteins across nine cell lines and tissues has been established by characterizing the respective proteins migration pattern by mass spectrometry. The dataset has been benchmarked against more than 200 siRNA validated antibodies and more than 6,000 antibodies from the Human Protein Atlas have been evaluated using the data. This strategy provides a more accurate way of evaluating antibodies based on the actual migration of their target protein rather than a theoretical molecular weight and also highlights the importance of validating antibodies in an application-specific manner.

571

Development of quantitative assay platform for multi-phosphorylated peptides using multiple reaction monitoring (MRM) method

Ji Hye Hong[1], Jonghwa Jin[1]

1. KBIO HEALTH, Cheongji-si, CHUNGBUK, South Korea

Protein phosphorylation is one of the post-translational modifications. Obtaining quantitative information on post-translational modification patterns of proteins is crucial to understanding the signaling and regulatory processes in many diseases. Despite many studies to quantify phosphorylation levels in biological systems, it is still difficult to measure the absolute phosphorylation level due to the lack of reliable quantitative methods. In this study, we developed the quantitative assay platform for multi-phosphorylated peptide consisting of three steps using multiple reaction monitoring (MRM). We first selected and synthesized single & multi-phosphopeptide to be phosphorylated by a serine, threonine and tyrosine; DRV*S**IHPF, DRV*Y**IHPF, DRV*T**IHPF, GADDSY*Y**TAR, GADDS*Y***Y**TAR, TRDI*Y**ETD*Y***Y**RK (* means phosphorylation). In the second step, we analyzed the mixture of phospho/non-phosphopeptides using Q-Exactive mass spectrometry to confirm the synthetic peptides and make a spectral library. In order to find the best method of enrichment efficiency in four different enrichment techniques; Fe-NTA (Thermo), TiO_2 (Thermo), magnetic TiO_2 (Pierce), TiO_2 Mag Sepharose (GE Healthcare), we compared the efficiency of phosphopeptide enrichment. The third step was to assess whether our assay platform showed the quantitative accuracy, stability and reproducibility in measuring the phosphorylation level; Response curves were generated, reproducibility was verified, and six phosphopeptides were evaluated for selectivity. Moreover, we followed the “CPTAC Assay Guidelines”, which provides a list of experiments that will help users obtain more reliable data for more accurate and stable quantitative assays. In conclusion, our developed the quantitative assay platform can be applied to more complicated samples such as tissue or cell lysates to analyze new phosphorylation sites and more phosphorylated peptides.

572

Kojak 2.0: New features for the analysis of cross-linked proteins

Michael R Hoopmann[1], Alex Zelter[2], Michael Riffle[2], Jimmy K Eng[2], Trisha N Davis[2], Robert L Moritz[1]

1. Institute for Systems Biology, Seattle, WA, United States

2. University of Washington, Seattle, WA, United States

Shotgun MS analysis of cross-linked proteins is a versatile tool in proteomics. Data analysis of cross-linked proteins has unique challenges for which specialized algorithms are required. Kojak was initially released in 2015 and designed to perform database searching on MS/MS spectra of cross-linked peptides. Designed to be computationally efficient, Kojak is highly customizable and allows for analysis with many different cross-linkers on both small and large datasets. Its simple interface, combined with adherence to open data standards, enabled Kojak’s use with diverse experimental conditions and allowed integration into analytical pipelines. Development of the algorithm continues to build upon these core features. Here we present Kojak version 2.0, a major update to the original Kojak algorithm.

Algorithm improvements include an optimized two stage search strategy that prioritizes identification of the larger peptide in the cross-link in the first pass. In the second pass, only those peptides that can link to the best candidates in the first pass are searched, providing a significant cost savings in computation time as database searches become larger and include increasing numbers of modifications in the parameters. The scoring functions were updated to include calculation of E-values, including individually for each peptide in the cross-link, enabling assessment of the cross-linked PSM using the E-value of its lowest scoring peptide, an invaluable parameter for downstream validation algorithms such as PeptideProphet and Percolator. Kojak now includes a feature that makes use of ^{15}N-labeled proteins mixed with their natural abundance counterpart, to enable accurate identification of inter-protein and intra-protein cross-links from homomultimers. Pipeline improvements include more open data standards for input (mzML, mzXML, MGF, Thermo RAW) and output (pepXML and mzIdentML), allowing integration into any workflow using these highly ubiquitous formats. Kojak 2.0 remains open-source and multi-platform.

573

Quantitative proteomics analysis highlights the dynamic changes of signalling events associated with neuronal death in excitotoxicity

Ashfaqul Hoque[1, 3, 2], Nicholas A. Williamson[3], Ching-Seng Ang[3], Jonathan S. Oakhill[1], Heung-Chin Cheng[3, 2]

1. St Vincent's Institute of Medical Research, Fitzroy, VIC, Australia

2. Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria, Australia

3. Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, Australia

Excitotoxicity, caused by over-stimulation or dysregulation of ionotropic glutamate receptors (iGluRs), is a major pathological process directing neuronal death in both acute and chronic neurological disorders. The aberrantly stimulated iGluRs direct massive influx of calcium ions into the affected neurons, leading to changes in expression and phosphorylation levels of specific neuronal proteins to modulate their functions and direct their participation in signalling pathways that induce excitotoxic neuronal death. To define these pathways, herein we utilised quantitative proteomics and phosphoproteomics analysis and identified over 150 neuronal proteins with significant dynamic temporal changes in abundance and/or phosphorylation levels at different time points (5 min to 4 h) after glutamate over-stimulation. Bioinformatic analysis predict that many of them are components of signalling networks directing defective neuronal morphology and functions, also Akt, Gsk3, Cdk5, JNK, CK2, SGK1 are predicted as the potential upstream kinases phosphorylating some of these perturbed proteins. Biochemical analysis confirmed the findings of the proteomic analysis for Erk1/2, Gsk3 and Tau. We also defined >40 neuronal protein and phosphoprotein molecules including CK2 and AMPK that are downstream of GluN-2B containing extra-synaptic NMDA receptors. Our predicted signalling networks and signalling dynamics of neuronal protein kinases form the conceptual framework for future investigation to define the spatial and temporal organisation of cell signalling pathways governing neuronal death in excitotoxicity.

574

Source induced dissociation (SID) with target MS2 for high throughput nanoLC-ESI-MS/MS mapping of multiple glycotopes

Cheng-Te Hsiao[1], Po-Wei Wang[1], Hua-Chien Chang[1], Kay-Hooi Khoo[1]

1. Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei 115, Taiwan

Protein glycosylation affects and modulates the structure and functions of membrane and secreted proteins. The glycans are attached to either asparagine or serine/threonine via a few common core structures, which can be extended either in linear or branched form before terminally capped by sialic acid, fucose, and/or sulfate at different positions to constitute the critical terminal Glyco-epitopes, or glycotopes. We have recently developed an advanced glycomic workflow that homes in on identifying these glycotopes by comprehensive nanoLC-MS^2-product dependent(pd)-MS^3 analysis of permethylated glycans. The coupled MS^3 step is required to distinguish the various isomeric variants of each glycotope, as defined by a specific MS^2 diagnostic ion, via the further induced linkage-specific cleavages. We now introduce an alternative implementation of similar glycotope mapping by coupling source induced dissociation (SID) with target MS^2 at tailored collision energy (CE), which also simultaneously allows selective data dependent MS^2 acquisition of sodiated species at higher CE setting. Within a 3 sec top speed mode cycle, a full MS survey scan acquired in the Orbitrap is followed, as first priority, by target MS^2 of a list of anticipated source induced fragment ions of interest, akin to a pseudoMS^3. A 2nd priority HCD-MS^2 acquisition at 50% NCE was then triggered on the sodiated form whenever a delta mass difference of 21.98 Da is detected to provide complementary sequence-informative fragmentation. A comparison with previously acquired pd-MS^3 data showed that current target MS^2 analysis produced similar results for high level glycomic mapping of the expressed glycotopes, including ± sialyl $Le^{x/a}$, $Le^{y/b}$, and disialylated Hex-3/4HexNAc glycotopes. Importantly, it yielded much better quality MS^2 spectra, instead of otherwise required pd-MS^3, to allow higher confident ID and relative quantification, as well as being able to target more glycotopes to be verified at MS^2 level in a single run.

575

Quantitative pleural effusion proteomics reveals the drug resistance-associated and prognostic biomarkers in lung adenocarcinoma

Ting-Feng Hsiao[1], Chih-Liang Wang[2], Ien-Chiuam Chiou[1], Hao-Yu Lin[1], Chia-Jung Yu[2, 1, 4, 3]

1. Graduate Institute of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan

2. Department of Thoracic Medicine, Chang Gung Memorial Hospital, Linkou, Taoyuan, , Taiwan

3. Department of Cell and Molecular Biology, College of Medicine, Chang Gung University, Taoyuan, Taiwan

4. Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

Introduction: Lung adenocarcinoma (ADC) is the most common subtype of non-small cell lung cancer. Epidermal growth factor receptor (EGFR) mutation is a predictor for initial treatment efficacy of inhibitors (EGFR-TKIs). Unfortunately, it can't be used to evaluate the outcomes of cancer patients. In addition, the drug-resistance against EGFR-TKIs and the following chemotherapy eventually lead to the poor survival of cancer patients. It is therefore important to search the drug susceptibility-associated biomarkers to benefit lung ADC patients. Pleural effusion (PE), a tumor-proximal body fluid, is associated with lung malignancy and serves as a promising source for biomarker discovery.

Methods: We herein applied high-abundant serum protein removal system and iTRAQ-based quantitative proteomics approach to generate the drug resistance-associated PE proteome with 561 quantified proteins from lung ADC patients who were sensitive or resistant to EGFR-TKI, EGFR-TKI/chemotherapy, and benign pulmonary disease. Both of the PE and serum levels of potential biomarkers were verified by ELISA. The protein expression in lung tissues was detected by Western blotting. Progression free survival (PFS) and overall survival (OS) were included to evaluate the prognostic value of potential maker.

Results: We selected seven proteins as potential markers of EGFR-TKI resistance by integration of our PE proteome with Gene Expression Profilling Interactive Analysis. We confirmed that the PE level of CDH3 was increased in patient with EGFR-TKI resistance as compared to patients with treatment naïve. The protein level of CDH3 was significantly increased in lung ADC tissue compared to adjacent normal parts. The serum level of CDH3 was indeed correlated to treatment efficacy of EGFR-TKI. Importantly, the CDH3 levels at baseline were associated with PFS and OS of ADC patients.

Conclusions: Our results collectively provide the useful databases for drug resistance study and prognosis of advanced ADC patients with targeted therapy.

576

Assessment of candidate biomarkers in paired saliva and plasma samples from oral cancer patients by targeted mass spectrometry

Yung-Chin Hsiao[1], Lang-Ming Chi[2], Kun-Yi Chien[3], Jau-Song Yu[1], Yu-Sun Chang[1]

1. Molecular Medicine Research Center, Chang Gung University, Tao-Yuan, Taiwan

2. Clinical Proteomics Core Laboratory, Chang Gung Memorial Hospital, Tao-Yuan, Taiwan

3. Department of Biochemistry and Molecular Biology, Chang Gung University, Tao-Yuan, Taiwan

For oral cancer, numerous saliva- and plasma-derived protein biomarker candidates have been discovered and/or verified; however, it is unclear about the behavior of these biomarker candidates as saliva or plasma biomarkers. In this study, we developed two targeted, multiplexed quantification assays, MRM and SISCAPA-MRM, to quantify 30 potential protein biomarkers in both plasma and saliva samples collected from 30 healthy controls and 30 oral cancer patients. In comparison with MRM assay, SISCAPA-MRM applying the anti-peptide antibodies for an additional immuno-enrichment effectively improved (>1.5 fold) the detection sensitivity of 11 and 21 targets in measurement of saliva and plasma samples, respectively. The integrated results revealed that the salivary levels of these 30 selected biomarkers weakly correlated ($r < 0.2$) to their plasma levels. Five candidate biomarkers (MMP1, MMP3, TNC, CSTA and PADI1) exhibited significant alternations and disease-discriminating powers (AUC >0.7) in saliva sample; nevertheless, no such targets could be found in plasma samples. Our data support the notion that saliva may be more suitable for the protein biomarker-based detection of oral cancer, and the newly developed SISCAPA-MRM assay could be applied to further verify the clinical utility of multiple oral cancer biomarker candidates in large cohort of saliva samples.

577

A novel LysargiNase- and chemical derivatization-based strategy LAACTer greatly facilitating in-depth profiling of C-terminome

Hao Hu[1], Wensi Zao[1], Mengdi Zhu[1], Linhui Zhai[1], Minjia Tan[1]

1. Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Pudong, SHANGHAI, China

Protein C-termini impact protein functions in multifaceted way, including localization, stability, and activity, etc. Among proteome-wide C-termini enrichment strategies, polyallylamine enrichment-based C-terminomics (PECTer), such as C-TAILS, represents a promising one due to the relatively simple workflow. Nonetheless, it still suffers from trypsin-related sequence bias and low identification rates. We sought to leverage the N-terminal cleavage specificity of LysargiNase to address these issues. By incorporation of LysargiNase digestion, neo-N-terminal acetylation, and a-ion-aided peptide matching into PECTer (termed LAACTer), we achieved to increase the average identification number by ~150% and total coverage by ~160%. LAACTer, together with PECTer, identified a total of ~ 1000 unique C-termini, which represented the most prominent improvement of PECTer so far. As to the quantitative performance, LAACTer showed even higher orthogonality to PECTer as revealed by identification for GluC cleavage products where only 2% of the total peptides were commonly identified by two methods. Finally, by combinational use of LAACTer and the PECTer, we revealed several novel functional insights for the C-termini of 293T cell, including sequence features related to protein stability, function alteration of proteins due to C-terminal processing, and possible truncation mechanisms. In conclusion, the newly established LAACTer will assist to deep C-terminome profiling and find wide applications in exploration for biological roles of protein C-termini.

578

An outer membrane protease OmpT based strategy for facilitated analysis of histone post translational modifications

Yajun Hu[1], Hong Jin[1], Pengyuan Yang[1]

1. Institutes of Biomedical Sciences, Fudan University, ShangHai, China

Histone modifications play an important role in regulating transcriptional gene expression and chromatin processes in eukaryotes. Increasing researches proved that aberrant post-translational modifications (PTMs) of histones is associated with many diseases. However, MS-based identification and quantification of histone PTMs is still challenging. Although classic chemical derivatization in conjunction with trypsin digestion protocol is a widely used in histone PTMs analysis in bottom-up strategy, several side reactions have been observed in practice. In this work, outer membrane protease T OmpT was utilized as a protease for direct histone proteolysis and generated convenient histone H3 peptide lengths for retention on reversed-phase chromatography. The powerful and unique tolerance of OmpT for modified lysine and arginine was demonstrated and quantitatively described for the first time, making it useful for detecting naturally modifications. Using the optimized digestion conditions, we succeeded to identified most of abundant modifications on histone from HEK293T cells. Additionally, histone H3 PTMs were quantitatively profiled in KMS11 cell line with selective knock out of translocated NSD2 allele (TKO) and the original parental KMS11 cell lines

(PAR), revealed the NSD2 was of high specificity on H3K36 dimethylation. Histone chemical derivatizations are not required in our strategy, showing a remarkable strength over the conventional trypsin-based workflow.

579

Phosphoproteomic Characterization of Signaling Networks Resulting from Activation of Chemokine Receptor CCR2

Cheng Huang[1, 2], Simon R Foster[1], Anup D Shah[2], Ralf B Schittenhelm[2], Martin J Stone[1]

1. Infection and Immunity Program, Monash Biomedicine Discovery Institute , Monash University, Clayton, VIC, Australia

2. Monash Proteomics and Metabolomics Facility, Monash University, Clayton, VIC, Australia

Inflammation is the body's response to injury or infection. A hallmark of inflammation is the accumulation of leukocytes, which remove pathogens and necrotic tissue by phagocytosis and proteolytic degradation. Leukocytes/monocytes are mainly recruited by chemokine activation of chemokine receptors, resulting in leukocyte morphological changes, extravasation into the inflamed tissue and chemotaxis along the chemokine gradient to the site of injury or infection. while monocyte chemoattractant proteins (MCPs) and chemokine receptor 2 (CCR2) is the major pair, involving this process and contributing to the pathogenesis of atherosclerosis, obesity and type 2 diabetes. CCR2 is known to signal via G protein and β-arrestin-mediated pathways, the downstream signalling pathways have not been thoroughly explored.

Protein phosphorylation and dephosphorylation are crucial for cellular signal transduction. Dynamic regulation of reversible, site-specific protein phosphorylation is critical to the signalling networks. Here, we performed a data-independent acquisition (DIA) based proteome and phosphoproteome quantification workflow to investigate signal transduction and regulation in MCP-1-activated CCR2-expressing cells. This workflow showed excellent reproducibility and quantification accuracy. More importantly, in addition to some canonical signalling pathways, such as MAPK, JAK/STAT and Akt/mTOR, we have mapped and manually curated other signalling networks, including Rho guanine nucleotide exchange factors (ARHGEFs), nuclear pore complex (NPC) proteins and actin cytoskeleton. Most of the characterised and quantified phosphopeptides in these networks have never been linked to MCP/CCR2 signalling. Phosphorylation kinetic study confirmed the dynamics phosphorylation regulation of the canonical networks; furthermore, it provides the phosphorylation kinetics information of ARHGEFs, NPC and actin cytoskeleton, the consequence of activating which matches the biological function of MCP/CCR2 signalling- guiding cell migration. In light of the accurate quantification and high reproducibility provided by DIA, this study provides new insights into MCP/CCR2 signalling and may guide the identification of potential therapeutic targets.

58

Mass Spectrometry Imaging in Cancer Diagnostics and Cancer Spheroids as Drug Testing Model

Peter Hoffmann[1], M.R Condina[1], P Mittal[2], M Klinger-Hoffmann[1], M.K Oehler[3]

1. University of South Australia, Adelaide, SA, Australia

2. Adelaide Proteomics Centre, School of BiologicalSciences , The University of Adelaide, Adelaide, Australia

3. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, Australia

Mass Spectrometry Imaging (MSI) is typically used to determine the distribution of proteins, lipids or metabolites in fresh frozen tissue. Formalin-fixed paraffin embedded (FFPE) tissue can be analyzed after antigen retrieval and enzymatic digest with trypsin or PNGase F on tissue for imaging tryptic peptides and N-glycans. Imaging tryptic peptides and N-glycans in FFPE tissue has multiple advantages over imaging of intact proteins in fresh frozen tissue. These includes the identification of peptides and N-glycans by matching accurate m/z from the MSI experiment with in situ MS/MS on tissue and high quality LC-MS/MS data obtained through in solution digestion of relevant laser dissected tissue. A novel method for investigating tissue-specific N-linked glycans was recently developed by our group on FFPE tissue.
Here we present the latest developments within our group, including up-to-date methods for analysis of FFPE tissue (e.g. tryptic peptide and PNGase F released glycans) and the use of tissue micro arrays. We present that MSI can spatially profile glycoforms in tissue-specific regions, while through LC-ESI-MS/MS the corresponding glycol compositions are structurally characterized. These methods are applied to endometrial and ovarian cancer FFPE tissues to potentially make diagnostic decisions in order to improve treatment of cancer patients.
We also show the testing of cancer drugs (olaparib, CDK 4/6 and CDK 9 Inhibitors) using 3 dimensional ovarian cancer spheroids and show distribution of the drugs and their efficacy by using embedding, sectioning and MSI analysis.

580

Antibody array-based quantitative proteomics platforms in biomarker discovery and development

Yingqing M Mao[1, 2], Zhizhou K Kuang[1, 2], Zhiqiang L Lv[2], Wei H Huang[2], Huihua Z Zhang[2], Ruopan H Huang[1, 2, 3, 4]

1. RayBiotech Life Inc, Norcross, GA, USA

2. RayBiotech Inc., Guangzhou, China

3. South China Biochip Research Center, Guangzhou, China

4. Affiliated Cancer Hospital & Institute of Guangzhou Medical University,Guangzhou Medical University, Guangzhou, China

Protein biomarkers are critical for drug target identification, early disease detection and precision medicine. In the biomarker discovery phase, it is important to capture as many candidate markers as possible in a biological sample. To meet this demand, we have developed both sandwich-based and label-based antibody array platforms for candidate biomarker screening. Currently,

the expression levels of more than 1000 proteins can be quantitatively measured using our sandwich-based platform. Most proteins can be detected at pg/ml levels; an additional 2000 proteins can be captured by the label-based platform. Therefore, in total, 3000 proteins can be profiled by combining both platforms. Furthermore, potential biomarkers with well-documented importance from the literature can be collected. Additional antibody arrays can then be generated targeting those candidates, allowing better biomarkers to ultimately be optimized and identified. Finally, to translate the biomarkers to clinical applications, we have developed an automated solution to detect multiple proteins in array format. Using this platform, we have analyzed hundreds of serum samples from different diseases including hepatoma, kidney cancer, endometriosis, glioblastoma and other diseases. A number of valuable biomarker panels which show high specificity and sensitivity have been identified. The results demonstrate the great promise of our platform for identifying and developing candidate biomarkers with potential clinical significance which may be broadly applicable to cancer and other human diseases.

581

Using Reverse Phase Protein Array as a Discovery Platform for Drug Selection in Triple Negative Breast Cancer Preclinical Animal Models

Shixia Huang[1], Lacey E Dobrolecki[1], Kimal Pajapakshe[1], Chad Shaw[1], Hsin-Yi Cindy Lu[1], Ramos Carlos [1], Dimuthu Perera[1], Coarfa Cristian[1], Michael T Lewis[1], Dean P Edwards[1]

1. Baylor College of Medicine, Houston, TEXAS, United States

Clinically, breast cancers are divided into three distinct groups: those that express the estrogen hormone receptor (ER+) (typically also express the progesterone receptor (PR+)), those that overexpress the ErbB2 (HER2) oncogene (HER2+), and those that express none of these three markers (termed “triple negative” breast cancer TNBC).

Unlike ER+ and HER2+ breast cancers; there are currently no targeted therapies against TNBC. Treatment of TNBC entails surgery coupled with radio- or chemotherapy, or both. The most commonly used chemotherapies are Taxanes (e.g. Docetaxel, Paclitaxel) and more recently, platinum-based agents (e.g. Cisplatin, Carboplatin). However, other than BRCA1/2 mutation status correlating with increased efficacy of platinum-based agents, there are currently no clinically useful predictors of differential treatment response among these commonly used chemotherapeutics.

We hypothesized that individual PDX may respond differentially to each chemotherapeutic, and as a consequence, a molecular predictor of differential chemotherapy response could be developed that might be useful clinically to predict benefit from one chemotherapy over another.

Using Reverse Phase Protein Array (RPPA) as a discovery platform, we analyzed a series of TNBC PDX models to identify potential protein pathways associated with drug resistance. We have identified MEK1, EZH2, and HDAC6 which are functionally validated in a 16 arms preclinical trial with single or double anti-cancer agents to overcome chemoresistance.

In this study, differential expressed proteins were identified among triple negative breast cancer PDX models upon Docetaxel and Carboplatin treatments. Inhibitors to two proteins (MEK1, EZH2) from the common up-regulated list as well as other targets of interest were used to design a 16 arms preclinical trial with single or double anti-cancer agents to overcome chemoresistance for triple negative breast cancers. Preliminary data for drugs to MEK1, EZH2, or HDAC in combination of Chemo-drugs showed that combined anti-cancer agents are more effective than the single agent.

582

Signalling Networks in the Analysis of Phospho/Proteomic Data

Hannah Huckstep[1], Jarrod Sandow[1], Andrew Webb[1], Liam Fearnley[1], Melissa Davis[1, 2]

1. Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

2. Department of Biochemistry and Molecular Biology, Faculty of Medicine, The University of Melbourne , Parkville, VIC, Australia

Signalling networks have the potential to provide useful insight into mechanisms driving disease progression. Proteomics and phosphoproteomics represent attractive datasets for this work as they directly measure the changes involved in signal propagation. Phosphorylation is one of the most common PTMs involved in regulating biological processes and the phosphorylation (or dephosphorylation) of a protein can have an activating or inhibiting effect. It has been estimated that as much as a third of the eukaryotic proteome is phosphorylated at one time indicating the significance of phosphorylation in modulating cell behaviour. Nevertheless, the simple identification and quantification of phosphoproteins from different conditions is not sufficient to reconstruct the mechanisms underpinning the observed differences. Understanding how altered proteins cooperate to modulate function requires exploration of the molecular interaction networks underpinning these changes.

An important first step in this process is the derivation of a network capturing known interactions. Most knowledgebases today organise such information into pathways, which do not properly capture the global flow of information across the entire signalling system. Here we conduct an overlap analysis consisting of a multi-factor comparison of several widely used knowledgebases in order to characterise their coverage of the global signalling network. Using these public signalling data, we designed a customised network more amenable to mapping of phospho/proteomic data using the flexible graphical database, neo4J. We explore how peptide level phosphoproteomics measurements can be mapped to this framework to interpret the functional consequences of the observed changes in protein phosphorylation. To facilitate this, we have developed a suite of methods to interrogate the global signalling network and extract informative subnetworks. This approach will enable a more unbiased and complete analysis to be performed over networks encompassing specific proteins and phosphoproteins of interest.

583

Scanning SWATH® Acquisition – The Next Step in the Data Independent Acquisition (DIA) Evolution

Christie Hunter[1], Stephen Tate[2], Nic Bloomfield[2]

1. SCIEX, Redwood City, CA, USA

2. SCIEX, Toronto, ON, Canada

The use of microflow LC in combination with SWATH® Acquisition has been growing steadily for quantitative proteomics as it provides better robustness and higher throughput when measuring larger sample cohorts. While variable window SWATH provides very good quantitative results on faster gradients, more narrow peak widths require faster cycle times and therefore present an opportunity for further acquisition improvements. With Scanning SWATH acquisition, a narrow Q1 window can be very rapidly scanned across a Q1 mass range and MS/MS is acquired at every small step. This adds an extra dimension to the DIA data that can be used to improve data quality.

Microflow LC was performed on the SCIEX TripleTOF® 6600 System with the OptiFlow® Source using the nanoLC™ 425 system. A trap-elute workflow was used and a range of gradient lengths from 5 – 45 mins was explored, with key acquisition parameters for Scanning SWATH Acquisition varied to optimize and compare to variable window SWATH acquisition using a research prototype Analyst® Software. Data was processed with research prototype SWATH software and results were analyzed using the SWATH Replicates Template.

Using a variety of complex matrices, both single species and mixtures of digests, acquisition parameters for Scanning SWATH Acquisition was explored. Not surprisingly, using a more narrow Q1 window provided improvements in # of peptides quantified over variable window SWATH for the microflow gradients tested. Optimization results will be presented.

584

Accelerating DIA Studies to Extend Workflow Utility, Using Fast Microflow LC Gradients

Christie Hunter[1], Nick Morrice[2]

1. SCIEX, Redwood City, CA, USA

2. SCIEX, Phoenix House Lakeside , Warrington, Cheshire, UK

Proteomics has typically been done using nanoflow LC for sensitivity but the time to results slow. With higher flow rates, sample can be loaded faster, trap/column can be washed and equilibrated faster, and gradients are formed faster, allowing much faster run times to be achieved. Microflow LC has been used increasing in quantitative proteomics in combination with SWATH® Acquisition, to provide better robustness and higher throughput when measuring larger sample cohorts. Here, the impact of gradient length on protein identification and quantitation results with DIA was explored.

Microflow LC was performed on the SCIEX TripleTOF® 6600 System with OptiFlow® Source using the nanoLC™ 425 system plumbed in microflow mode. Trap-elute workflow was used and a range of gradient lengths from 5 – 45 mins was explored, and key acquisition parameters for SWATH Acquisition were varied to optimize for the much faster run times. Data was processed with SWATH 2.0 microapp in PeakView® Software 2.2 and OneOmics™ in SCIEX Cloud.

Using complex digested cell lysates, SWATH Acquisition experiments were performed using gradient lengths ranging from 5-45 mins and protein quantitation results were assessed. Fast MS/MS acquisition rates were found to be critical because this enabled more, smaller variable Q1 windows to improve S/N for quantitation. Even with the fastest gradients, methods with 60-100 windows with very fast accumulation times of 15 msec improved results. Optimization results will be presented. Optimized methods were then used to compare quality of quantitation between long and shortened gradients and similar fold change values were measured confirming accelerated gradients can be used for industrialized quantitative proteomics.

585

SAAVpedia: identification, functional annotation, and retrieval of single amino acid variants for proteogenomic interpretation

Heeyoun Hwang[1], Soo Youn Lee[1], Young-Mook Kang[1], Hyejin Kim[1, 2], Dong Geun Kim[1, 2], Ji Eun Jeong[1, 2], Jin Young Kim[1], Jong Shin Yoo[1, 2]

1. Korea Basic Science Institute, Cheongju, CHUNGBUK, Rep. of Korea

2. Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon, Rep. of Korea

Next-generation genome sequencing has enabled the discovery of numerous disease/drug-associated non-synonymous single nucleotide variants (nsSNVs) that alter the amino acid sequences of a protein. Although several studies have attempted to characterize pathogenic nsSNVs, few have been confirmed as single amino acid variants (SAAVs) at the protein level. Here, we developed the SAAVpedia platform to identify, annotate, and retrieve pathogenic SAAV candidates from proteomic and genomic data. The platform consists of four modules: SAAVidentifier, SAAVannotator, SNV/SAAVretriever, and SAAVvisualizer. The SAAVidentifier provides a reference database containing 18,206,090 SAAVs, and performs identification and quality assessment of SAAVs. The SAAVannotator provides functional annotation with biological, clinical and pharmacological information for interpretation of condition specific SAAVs. The SNV/SAAVretriever module enables bi-directional navigation between relevant SAAVs and nsSNVs with diverse genomic and proteomic data. SAAVvisualizer provides various statistical plots based on functional annotations of detected SAAVs. To demonstrate utility of SAAVpedia, the proteogenomic pipeline with protein-protein interaction network analysis was applied to proteomic data from breast cancer and glioblastoma patients. We will extend the SAAV validation database from a variety of proteomic data to further biomedical research. SAAVpedia will play a key role in pathogenic biomarker discovery based on massive proteogenomic data interpretation. The SAAVpedia is available at https://www.SAAVpedia.org/.

586

Quantitative proteomics analysis using label-free DARTS and LC-MS/MS method reveals a target protein of cancer therapeutic small molecule

Hui-Yun Hwang[1], Yoon Sun Cho[1], Jin Young Kim[2], Ki Na Yun[2], Jong Shin Yoo[2], Eun hyeong Lee[3], In june Kim[3], György Marko-Varga[1, 4], Ho Jeong Kwon[1]

1. Chemical Genomics Global Research Laboratory, Department of Biotechnology, Yonsei University, Seoul, South Korea

2. Biomedical Omics Group, Korea Basic Science Institute, Chungbuk, South Korea

3. Biomedical Science and Engineering Interdisciplinary Program, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

4. Clinical Protein Science & Imaging, Department of Biomedical Engineering, Lund University, Lund, Sweden

Manipulating autophagy is a promising strategy for treating cancer with several autophagy inhibitors shown to induce autophagic cell death. One of these, autophagonizer (APZ), induces apoptosis-independent cell death by binding an unknown target via an unknown mechanism. To identify APZ targets we developed a label-free drug affinity responsive target stability (DARTS) approach with a liquid chromatography/tandem mass spectrometry (LC-MS/MS) readout. Of 35 protein interactors, we identified APZBP as a key target protein of unmodified APZ in autophagy. Additionally, sequence coverage analysis in DARTS LC-MS proteome revealed that specific peptide fragments in the nucleotide binding domain (NBD, ATPase domain) of APZBP were protected by APZ in the pronase-treated proteome. APZBP inhibition through siRNA knockdown or by a specific inhibitor, inhibited autophagy in HeLa cells similar to APZ treatment, whereas APZBP overexpression overcame inhibition induced by APZ. Additionally, APZ exhibited significant synergism with TMZ (temozolomide), which has been used for chemotherapeutic agent of glioblastoma, *in vitro* and orthotopic mouse model. These findings demonstrate the potential of APZ to induce autophagic cell death and its development as a potent anticancer drug. Our study showed that combined DARTS and LC-MS/MS-based target identification of non-tagging bioactive small molecule of interest is effective for novel protein target identification and revealed a novel function for APZBP as cancer therapeutic target.

587

Lessons of protein functional annotation history

Ekaterina V Ilgisonis[1], Pavel V Pogodin[1], Elena A Ponomarenko[1]

1. Institute of Biomedical Chemistry (IBMC), Moscow, MOSCOW, Russia

Proteins serve a variety of functions within cells. Some are involved in structural support and movement, others in enzymatic activity, and still others in interaction with the outside world.

The gap between the number of known proteins and their biological function annotation is continually increasing. Protein function is a multifaceted and complex phenomenon, and there is currently no single algorithm to determine it. There are a lot of algorithms of functional prediction today based on sequence homology, combining data from multiple different sources and implementing advanced machine learning technique. But how one can validate protein function prediction?

At the first stage we made retrospective analysis of Nextprot database (Gaudet et al., 2017) to reveal the most popular way of experiments for protein function validation (for different functions-different experimental approaches). After that we decided to focus our efforts on the functional annotation of chromosome 18 upe1 proteins (upe1 protein – proteins without known function).We decided to perform text-mining and meta-analysis. Search queries - the names of this protein in the PubMed does not give results. PRIDE contained 23 datasets with this protein. For the further analysis we have chosen 16 datasets created after 2016 (when HPP Data Interpretation Guidelines version 2.0 were published). This datasets were described in 12 articles respectively. Analysis of their MeSH-terms allowed us to form primary hypothesis about the Q68DL7 protein functional role.

At the next stage we analized co-occurance of this protein with other proteins in the same articles and experimental datasets. We used COFACTOR (Zhang et al., 2017) and I-TASSER (Yang et al., 2015) algorithms for protein function prediction based on protein structure. Basing on the principle “guily-by-association” the hypothesis about the role of this protein in different metabolic pathways was formulated.

588

Utilising LC-MS/MS to identify novel T cell targets in influenza infection

Patricia Illing[1], Pouya Faridi[1], Nicole Mifsud[1], Ziyi Huang[1], Grace Khuu[1], Marios Koutsakos[2], Luca Hensen[2], Bridie Clemens[2], Oanh Nguyen[2], Katherine Kedzierska[2], Anthony Purcell[1]

1. Infection and Immunity Program, Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia

2. Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria, Australia

Influenza viruses are a significant cause of respiratory illnesses, inflicting a high level of morbidity and mortality worldwide especially in at risk populations such as the young and elderly. Current vaccines are formulated based on the strains predicted to circulate in the upcoming season and induce predominantly antibody-based immune responses against the surface glycoproteins. Due to antigenic drift and shift, these glycoproteins can differ greatly between strains resulting in limited cross-strain protection. In contrast, $CD8^+$ T cells can recognise virus-derived peptides displayed by HLA class I molecules (HLA-I) on the surface of infected cells. These peptides are generated through the breakdown of viral proteins within the cell, hence include peptides derived from more conserved internal viral proteins, building capacity for T cell responses that target a range of strains.

To determine which viral components have the potential to stimulate $CD8^+$ T cell responses, we have developed a mass spectrometry-based epitope discovery approach, isolating peptides from the HLA-I of infected cells to determine the naturally presented virus-derived peptides available for immune surveillance. Utilising a *de novo* assisted database searching strategy (PEAKS 8.5, BSI), combined with in-house software, we have mapped the contribution of both genome templated (linear) and a newly recognised subset of non-genome templated (spliced) peptides to the array of peptides displayed by the HLA-I. Current work moves to assess the contribution of both linear and spliced peptide subsets to the anti-influenza immune response in healthy donors. The knowledge of which viral components can instigate immune responses is fundamental to the rational design of vaccines to maximise T cell responses, hence these data will help inform future strategies for the generation of a universal influenza vaccine. This work is funded by the Human Vaccines Project Michelson Prizes for Human Immunology and Vaccine Research.

589

Development of water droplet in oil reaction protocol for improving tryptic peptides recovery from small-scale sample preparation

Yuma Inamori[1], Takeshi Masuda[1, 2, 3], Shingo Ito[1, 2, 3], Sumio Ohtsuki[1, 2, 3]

1. Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, JAPAN

2. Faculty of Life Sciences, Kumamoto University, Kumamoto, JAPAN

3. AMED-CREST, Tokyo, JAPAN

Introduction

Single-cell proteomics is a key technology for the understanding of tissue and cell heterogeneities. However, it is hardly possible to perform single-cell proteomics because of low peptide recovery during sample preparation. To reduce the contact area of a sample solution with a plastic tube, we have developed a water droplet in oil (WinO) protocol. WinO is an inexpensive and easily manual protocol for small-scale preparations. In this presentation, we evaluated and optimized the WinO to improve the tryptic peptides recovery toward single-cell proteomics.

Methods

HEK293 cells were solubilized by 100 mM Tris-HCl (pH9.0) containing 12 mM sodium deoxycholate and 12 mM sodium lauroyl sarcosinate. The droplet was formed by adding 1 µL of the whole cell lysate into ethyl acetate. Proteins were reduced and alkylated, and digested with Lys-C followed with trypsin for overnight. Tryptic peptides were analyzed by a nanoLC-MS/MS.

Results and discussion

To evaluate the improvement of peptides recovery in WinO, 10 ng of proteins were digested by WinO or in-solution digestion (InSol). The identified number and total signal intensities of proteins in WinO were 1.7-fold greater than those in InSol, indicating that the peptides recovery was improved in WinO. However, the total signal intensities of 10 ng of peptides digested by WinO were only 38% compared with those of 10 ng of peptides divided from 10 µg of peptides prepared by InSol. We have assessed the sample recovery in each preparation step, and the lowest recovery step was trypsin digestion. Trypsin activity in WinO was decreased by 34% compared with that in InSol, suggested that the low recovery was due to low trypsin digestion efficiency in the droplet. In conclusion, WinO provides higher peptides recovery than InSol. WinO can be the novel technology for small-scale preparations, including single-cell analysis, by improving trypsin digestion efficiency.

59

Exploring the limits of high-resolution mass spectrometry imaging data

Jonatan J.O. Eriksson[1], Melinda M. Rezeli[1], Max M. Hefner[1], Gyorgy G. Marko-Varga[1], Frank F. Suits[2], Peter P.L. Horvatovich[3]

1. Department of Biomedical Engineering, University of Lund, Lund, Swedeb

2. IBM Research, Melbourne, Victoria, Australia

3. University of Groningen, Groningen, GRONINGEN, Netherlands

Imaging with high-resolution mass spectrometry (MSI) such as Orbitrap mass analyzers yields large datasets that contain the spatial distribution of all compounds that can be sampled from a tissue section into the mass spectrometer. In this talk we present a bioinformatics approach that allows rapid exploration of a large MSI dataset without loss of information, while generating high-quality ion images of compound isotopes in accordance with the resolution of the mass spectra.

Our approach includes peak detection in all collected spectra, construction of a consensus peak list, and generation of ion images of compound isotopologue peaks using the resolution of the mass spectra at given m/z. Scoring the content of the spatial structure by the first principal components and using spatial chaos metric allows culling of ion images with low information content.

Our tool enables spatial correlation queries using ion images, selected area or combination of the selected area and ion distribution. The performance of our approach over regular binning and other tools such as Cardinal is demonstrated by automatic generation of separate clean ion images of spiked-in drug compounds or lipids close in mass. We further demonstrate the tool's performance by its ability to detect a large number of isotopes and in-source fragments of spiked-in cancer drugs in correlation queries.

590

UniScore, a universal measure for annotated peptide product ion spectra

Tsuyoshi Tabata[1], Akiyasu C. Yoshizawa[1], Naoyuki Sugiyama[1], Yasushi Ishihama[1]

1. Kyoto University, Kyoto, KYOTO, Japan

The number of public mass spectrometry datasets has been exponentially growing in proteomics through the public domain such as jPOST, https://jpostdb.org as reusable resources. In general, public datasets consist of accumulating search results from multiple projects and institutions under the wide variety of experimental conditions, leading to linear increase of false positive hits when the registered results are just combined. Therefore, it is required to develop novel repository-scale computational workflows to control the quality of the re-analyzed datasets. Here we will propose a universal measure for annotated peptide MSMS spectra stored in public repositories named UniScore. Using the scores, we can accumulate different results from multiple search engines as well as multiple peak picking algorithms to minimize the false positive identifications.

MS raw data acquired in the global proteome analysis of Tryptic peptides from HeLa cells by Q-Exactive with 8h gradient were obtained from the jPOST repository with JPST000200. Another dataset was Phosphopeptides from mouse Hepa cells by Q-Exactive (PXD001792). The peak list was created by MaxQuant and the protein identification was performed by Mascot, X!Tandem, Comet and MaxQuant.

UniScore was calculated based on the peak annotation of MSMS spectra, independent of employed search engines and peak picking algorithms. Moreover, by optimizing the parameters for UniScore, we can reduce the false discovery rate by minimize the decoy hits. One of the major parameters of UniScore is based on the concept of peptide sequence tags (PSTs) and we accepted the number of amino acids consisting of the PSTs. So far, the UniScore-based identification generates 10-20% increase in peptide identification for relatively large-scale datasets.

We successfully developed a universal measure for peptide MSMS spectra. It is applicable for re-analysis of accumulated datasets in public repository as well as for spectral libraries.

591

Biomarkers of endometriosis

Jason Ito[1], Scott Bringans[1], Kirsten Peters[1], Tammy Casey[1], Christina Andronis[1], Richard Lipscombe[1]

1. Proteomics International, Broadway, Nedlands, WA, Australia

Endometriosis is an under-recognised and often misdiagnosed chronic disease that affects one in ten women. It occurs when cells similar to those that line a woman's uterus grow in other parts of her body, usually around the pelvis. The current gold standard for diagnosis of endometriosis is direct visualisation of the tissue with confirmation by histological analysis which can only be achieved by laparoscopy/laparotomy under a general anaesthetic. It is estimated 176 million women and girls around the world are living with endometriosis. The condition costs $7.7 billion a year in direct medical expenses and lost productivity in Australia alone.

Proteomics International has utilised its Promarker™ technology platform to search for unique biomarkers of Endometriosis. Initially, three pooled plasma samples from three distinct groups of patients were tested. The first group had symptoms of endometriosis and a surgical procedure performed to confirm this diagnosis. The second group also had symptoms but the surgical procedure did not confirm the diagnosis. The third group were healthy matched controls. The pooled plasma samples were quantitatively profiled against each other with a triplicate iTRAQ 2D-LCMS experiment and out of the 399 proteins identified (≥2 significant peptides), 30 showed significant levels of change when comparing the diseased state to the two other groups. These 30 potential biomarkers are being tested in separate groupings of individual patient plasma samples by targeted mass spectrometry to verify their potential as diagnostic biomarkers.

592

High-throughput proteome analysis using 50 cm long micro pillar array (µPAC™) columns

Jeff Op de Beeck[1], Geert Van Raemdonck[1], Paul Jacobs[1]

1. PharmaFluidics, Zwijnaarde, OOST VLAANDEREN, Belgium

In the past few years, micro pillar array column technology (µPAC™) has evolved from a tool to perform fundamental LC research to a powerful alternative for classical nano LC columns. The inherent high permeability and low 'on-column' dispersion obtained by the perfect order of the separation bed makes pillar array based chromatography unique in its kind. The peak dispersion originating from heterogeneous flow paths in the separation bed is drastically reduced and therefore components remain much more concentrated during separation (1). Apart from an improved efficiency, pillar array columns can also be designed with substantially lower flow resistances compared to packed bed columns.

Using a 50 cm long reversed phase C18 µPAC™ nano LC column in combination with a Thermo Orbitrap Elite mass spectrometer for detection, we demonstrate extremely robust and high-throughput proteome analysis at capillary flow rates up to 2 µl/min. At these flow rates, sample throughput can be increased to 20, 30, 60 and even 100 samples per day with corresponding peak capacity values (2) of respectively 300, 250, 200 and 150. When comparing the chromatographic performance that could be obtained for single protein and cell lysate tryptic digest samples to state-of-the art packed bed nano LC columns (packed with sub 2 µm silica particles), average peptide peak widths could be reduced by a factor of 1.6. For complex HeLa cell digest samples, this resulted in an increase in peptide and protein group identifications of respectively 60 and 40%.

To increase the throughput even further, a micro pillar array based trapping column was implemented into the workflow. By doing so, sample loading times for diluted protein digest samples could be reduced up to a factor of 15 without affecting the chromatographic performance.

1. (1) De Malsche, W., Gardeniers, H., and Desmet, G. (2008) Experimental Study of Porous Silicon Shell Pillars under Retentive Conditions. Anal. Chem. 80, 5391–5400
2. (2) Neue, U. D. (2008) Peak capacity in unidimensional chromatography. J. Chromatogr. A 1184, 107–130

593

Intact glycopeptide analysis by trapped ion mobility tandem mass spectrometry

Hans JCT Wessels[1], Melissa Bärenfänger[1], Merel Post[1], Kristina Marx[2], Pierre-Olivier Schmit[3], Markus Lubeck[2], Alain Van Gool[1], Dirk Lefeber[1], Jaran Jainhuknan[3]

1. Radboud UMC, Nijmegen, Netherlands

2. Bruker Daltonik GmbH, Bremen, Germany

3. Bruker France, Wissembourg, France

Introduction

Recent introduction of parallel accumulation serial fragmentation (PASEF) on a trapped ion mobility quadrupole time-of-flight mass spectrometer (timsTOF Pro) provides unique possibilities for comprehensive glycopeptide profiling in complex samples such as blood plasma. Glycosylated peptides are known for their poor ionization efficiency during electrospray ionization which we addressed by providing dopant enriched nitrogen gas into an axial desolvation vacuum assisted electrospray ionization source (CaptiveSpray nanoBooster). Here, we evaluated the effects of different dopants for glycopeptide analysis by ion mobility mass spectrometry and PASEF.

Methods

Glycopeptides from model glycoproteins and blood plasma were separated by reversed phase liquid chromatography (nanoElute, Bruker Daltonics) with online analysis in a trapped ion mobility quadrupole time-of-flight mass spectrometer (timsTOF Pro, Bruker Daltonics) using acetonitrile and primary alcohols as nitrogen gas dopants in the ESI source (CaptiveSpray nanoBooster). Data analysis was performed in DataAnalysis (Bruker Daltonics), ProteinScape/GlycoQuest (Bruker Daltonics), PEAKS studio (Bioinformatics Solutions), and by in-house developed scripts.

Results

Use of dopants in nanoflow ESI on the timsTOF Pro platform requires significant adjustment of the TIMS vacuum conditions and prohibits the use of fluorinated phosphazene calibrants due to signal loss. Alternatively, Poly-DL-Alanine can be used as calibrant which also offers additional data points for TIMS calibration. Acetonitrile and primary alcohol dopants significantly enhance the ionization efficiency of glycopeptides in ESI compared to air or nitrogen gas. Acetonitrile supercharges glycopeptides whereas primary alcohols subcharge glycopeptides. Interestingly, primary alcohols also reduce the number of charge states per ion which simplifies the ion mixture prior to spectrometric analysis. At present, we analyze for each dopant the ion mobility characteristics as well as PASEF performance for glycan- and peptide-moiety characterization.

Summary

This work present the first application and evaluation of ESI dopants for enhanced glycopeptide characterization by trapped ion mobility mass spectrometry and PASEF.

594

Comparative Analysis of Protein Expression Associated with Antitumor Activity of Methanol Extracts of *Agrimonia Pilosa Ledeb*

Hye Ock Jang[1]

1. Pusan National University, Yangsan Si, GYEONGSANGNAM-DO, South Korea

Oral cancer is a mass composed of cancer cells in the oral cavity. It accounts for about 5% of malignant tumors of the whole body. About 97% of all oral cancers are squamous cell carcinomas, followed by adenocarcinomas and rarely malignant melanomas. It occurs especially in males (about twice that of females), and it occurs in middle - aged people over 40 years of age. *Agrimonia pi!osa Ledeb*. has been traditionally known as a plant with effective antitumor activity. And it is being used for cancer therapy in China today. *Agrimonia pilosa Ledeb* has been used traditionally for treatment of abdominal pain, sore throat, headaches, bloody discharge, parasitic infections and eczema in Korea and other Asian countries since centuries. Most studies on *Agrimonia pilosa Ledeb* are related to leaves of *Agrimonia pilosa Ledeb*, and there are few studies on roots of *Agrimonia pilosa Ledeb*. However, the detailed mechanisms of *Agrimonia pilosa Ledeb* in antitumor profilees were not fully elucidated. Further, to date no report exists on the antitumor effect of *Agrimonia pilosa Ledeb* on oral squamous cell. In the present study, we investigated the antitumor activity of the methanol extract from roots of *Agrimonia pilosa Ledeb*, proteins that are changed by *Agrimonia pilosa Ledeb*, were identified using the proteomic approach, and then improving the experimental basis for the in-depth study of pharmacological effects of *Agrimonia pilosa Ledeb.*

595

High-throughput mass-spectrometry-based analysis of endogenous peptide-HLA I thermal stability improves the prediction of T cell epitopes

Emma C Jappe[1, 2], Nathan P Croft[3], Christian Garde[1], Ethan Passantino[3], Thomas Trolle[1], Jens V Kringelum[1], Anthony W Purcell[3]

1. Evaxion Biotech, Copenhagen, COPENHAGEN, Denmark

2. Department of Health Technology, Technical University of Denmark, 2800 Lyngby, Denmark

3. Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800

Publish consent withheld

1. Yewdell, J. W., Reits, E. & Neefjes, J. Making sense of mass destruction: Quantitating MHC class I antigen presentation. Nat. Rev. Immunol. 3, 952–961 (2003).
2. Gfeller, D., Bassani-Sternberg, M., Schmidt, J. & Luescher, I. F. Current tools for predicting cancer-specific T cell immunity. Oncoimmunology 5, 1–9 (2016).
3. Koşaloğlu-Yalçın, Z. et al. Predicting T cell recognition of MHC class I restricted neoepitopes. Oncoimmunology 7, 1–15 (2018).
4. Rasmussen, M. et al. Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. J. Immunol. 197, 1517–1524 (2016).
5. Blaha, D. T. et al. High-Throughput Stability Screening of Neoantigen / HLA Complexes Improves Immunogenicity Predictions. 7, 50–62 (2019).

596

Mass spectrometry-based proteomics of cellular proteins of radiation exposure time

Hwangkyo Jeong[1], Jeonghun Yeom[1], Je-Won Ryu[2], Sang-wook Lee[3], Kyunggon Kim[1, 2]

1. Convergence Medicine,Asan Institute for Life Science, Asan Medical Center, Songpa-gu, SEOUL, South Korea

2. Asan Institute for Life Science, Asan Medical Center, Seoul, South Korea

3. Radiation Oncology, Asan Institute for Life Science, Asan Medical Center, Seoul, South Korea

Radiation therapy show various cellular responses to Human lymphocytes.However, the molecular mechanisms of radiation-induced cell differentiation, malignant transformation, and apoptosis are not fully understood. In this study, changes in protein in B-lymphocyte (GM10832) in various radiation exposure environments were LFQ (Lable Free Quantification) analyzed by LC-ESI-MS / MS and Proteome Discoverer software. As a result, we were able to identify protein groups that vary with irradiation dose and protein groups that change with time after irradiation. In addition, T acute lymphoblastic leukemia cells (Molt4) were used to find protein groups induced by radiation in cancer cells. Identification of proteins for cellular responses will potentially provide deeper insights into radiation-induced cytotoxicity, particularly interesting in radiation therapy and protection programs.

597

Peroxiredoxin 6 is downregulated in platelets of first-onset psychosis

Helena P G Joaquim[2, 1], Alana C Costa[2, 1], Leda L Talib[2, 1], Wagner F Gattaz[2, 1]

1. Psychiatry, Instituto Nacional de Biomarcadores em Neuropsiquiatria (INBioN), Conselho Nacional de Desenvolvimento Cientifico e Tecnológico, Sao Paulo, Brazil., Sao Paulo, São Paulo, Brasil

2. FMUSP, São Paulo, SP, Brazil

Background: Markers of oxidative stress have already been considered as a reflection of the balance of protective and degenerative mechanisms in psychoses. The antioxidant protection system includes several enzymes, such as peroxiredoxins (PRDXs). PRDX6 is a bifunctional enzyme and both activities are associated with SCZ physiology: peroxidase and phospholipase A2 activity.

Aim: To determine the PRDX6 levels in first onset drug-naïve psychotic patients.

Method: We investigated by western blot the protein expression of peroxiredoxin 6 in platelets of 42 first onset, drug-naïve psychotic patients (28 schizophrenia and 14 bipolar disorder) as compared to 16 age and gender matched healthy controls.

Results: PRDX6 level is lower in patients with schizophrenia and bipolar disorder compared to controls (p=0.049). If we consider psychotic patients vs. healthy control the level is even more decreased (p=0.028). No associations were found between PRDX6 level and gender, age, duration of the illness and psychopathological scores.

Conclusions: The role of PRDX6 in neuropsychiatry diseases may be controversial. The binding of PRDX6 to different substrates regulates its peroxidase and $iPLA_2$ activities. Possibly, in first onset psychosis, antioxidant pathway is more prominent and informative in platelets.

598

High-throughput Targeted Lipidomics Analysis of Dihydroceramide Desaturase-1 (DES1) Knockout Mice

Arianna Jones[1], Justin Lim[2], Mackenzie Pearson[1], Santosh Kapil[1], Trevor Tippetts[3], Scott Summers[3]

1. SCIEX, Framingham, MA, United States

2. SCIEX, Singapore

3. Dept of Nutrition and Integrative Physiology, University of Utah, Salt Lake City, Utah, USA

Retention time scheduling for MRM transitions during targeted assays enables more compounds to be quantified with higher quality results. In this assay, MRMs to quantitate over 1150 lipid molecular species were combined into a single assay, covering lipids containing fatty acids with 14 to 22 carbons and 0 to 6 double bonds. The selected amide column chemistry provides reproducible isomer separation based on lipid class, to minimize isomeric overlap. This liquid chromatography separation strategy is then coupled with the time scheduled targeted MRM assay. To validate this method, we selected DES1 knockout mice to measure lipid changes in liver and adipose tissues. The method was then adjusted to optimize window width and dwell weight to enhance sensitivity and coverage. This method provides extensive, reproducible lipid coverage in complex biological samples like tissue, cells, and plasma.

Liver, epididymal white adipose tissue (eWAT), and subcutaneous white adipose tissue (sWAT) extracts were subjected to amide LC-MS/MS analysis with a QTRAP® 6500+ mass spectrometer coupled with ExionLC™ System. To streamline the development of this assay, an Excel based method optimization tool (sMRM Pro Builder) has been developed to assist with the assay refinement steps. The method is customizable, so new lipid categories, classes, and molecular species can be added to the MRMs list.

Liver, eWAT, and sWAT tissues were harvested from dihydroceramide desaturase-1 (DES1) knockout mice. DES1 is the enzyme responsible for inserting the 4,5-trans-double bond into the sphingolipid backbone causing dihyrdoceramides conversion to ceramides.

This widely targeted panel allows for quantitation of 19 lipid classes in several matrices, and yet has the sensitivity and selectivity to visualize significant changes in lipid species that are low abundant in eWAT and liver tissue.

599

Improved Metabolite Identification in a Single Injection with Data Independent Acquisition for Untargeted Metabolomics Workflow

Arianna Jones[1], Bob Proos[1]

1. SCIEX, Framingham, MA, United States

Comprehensive metabolite identification with MS/MS library spectral matching can be problematic for data dependent acquisition (DDA) workflows as it often requires multiple injections for each sample. SWATH® Acquisition, a Data Independent Acquisition (DIA) method, with optimized variable windows, provides a powerful workflow requiring only a single injection per sample for each polarity. In addition to capturing product ion spectra for all detectable analytes, SWATH Acquisition also provides the option of quantitation at either the MS or MS/MS level allowing a comprehensive qualitative and quantitative analysis of metabolites in complex biological samples like plasma. The speed of the X500R QTOF and the TripleTOF® 6600 QTOF system allows for the top 20 selection in DDA or 20 variable SWATH acquisition windows in DIA analysis with a cycle time of 651 msec for both the DDA and SWATH Acquisition methods, which provides more than 9 points across a peak of 6 seconds providing accurate and reproducible integration.

Human Plasma, NIST SRM1950, was purchased from NIST and metabolites were extracted using cold methanol. Both the DDA and SWATH Acquisition methods compared in this study were single injection workflows, using the same simple RP chromatography with a 20 minute run time per injection. Data was processed using SCIEX OS software for non-targeted peak finding with spectral library matching.

Top 20 DDA method was able to identify 476 of the features on the basis of precursor mass and MS/MS spectral matching. The 20 variable window SWATH acquisition method resulted in an additional 215 features being identified, a 45% increase, for a total of 671 compounds identified in a single injection. This study has shown that variable window SWATH acquisition can be a useful for improving compound identification in untargeted metabolomics.

600

Bead-based workflows for rapid urinary proteome profiling and simultaneous recovery of DNA, RNA and proteins for clinical proteogenomic studies

Justin Jordaan[1], Stoyan Stoychev[2], Previn Naicker[2], Ireshyn Govender[2], Isak Gerber[2]

1. Resyn Biosciences, Edenvale, GAUTENG, South Africa

2. NextGen Heath Cluster, CSIR, Pretoria, Gauteng, South Africa

The integration of cross-omics approaches, e.g. through genetic and protein isolation (proteogenomics), and analysis can provide comprehensive insights into the regulation of gene expression, cell signalling pathways, monitoring the onset of disease and subsequent progression. The collection and preservation of DNA, RNA and protein from the same patient sample is however challenging due to the different properties of these biomolecules and requirements for collection. This is often further complicated in clinical settings where minimal sample processing is possible due to a lack of suitable laboratory equipment. Urine is an attractive source for novel protein biomarker discovery due to its non-invasive collection and high abundance from patients. However, the DNA, RNA and protein content of urine is prone to degradation, and the dynamic range of e.g. proteins is greater than five orders of magnitude, with a wide molecular weight distribution due to protein cleavage that occurs in the kidneys. In this work we demonstrate workflows for the capture, processing and storage of DNA, RNA, and protein from patient samples that allow for robust, high-throughput sample processing. We compare two magnetic bead workflows for protein isolation, using hydrophilic affinity (HILIC) as well as strong anion exchange (SAX) for SPE, and benchmark against a traditional method and commercial kit. We further elaborate on the collection of proteins DNA and RNA from a range of biological samples, with downstream testing for a range of analyses.

601

Serum exosomes derived from colorectal cancer patients promote cancer cell migration

Hyeryeon Jung[1], Yu-ri Seo[1], Eugene C Yi[1]

1. Molecular medicine and biopharmaceutical sciences, Seoul National University, Seoul, South Korea

Publish consent withheld

602

Multiple Biomarker Panel to Predict Response to Tocilizumab(anti-IL6R) in Rheumatoid Arthritis Patients Using High-precision Proteomics Approach

Jinwoo Jung[1], Byoung-Kyu Cho[1], Yeong Wook Song[2], Eugene. C Yi[1]

1. Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, South Korea

2. Division of Rheumatology, Department of Internal Medicine, College of Medicine, Seoul National University, Seoul, South Korea

Rheumatoid arthritis (RA) is a common chronic and systemic autoimmune diseases that cause inflammation of the thin layer of tissue lining the joints. Interleukin-6 (IL-6), along with TNF-a and several inflammatory cytokines, acts a vital role in activation of local synovial leukocyte production and induction of chronic inflammation. A humanized anti-IL-6 receptor(IL-6R) monoclonal antibody, Tocilizumab (TCZ), has been demonstrated a significant clinical efficacy for RA patients. However, like other inflammatory cytokine blockers such as TNF-a, Interleukin-1 (IL-1), or CD20 inhibitors, some patients still show a partial respond or resistant to the treatment. This study therefore aimed at identifying protein biomarkers that could predict clinical response against TCZ in RA patients by implementing high-precision proteomics approach. We first identified 54 serum protein biomarker candidates from a large-scale serum proteome profiling of TCZ responder and non-responder groups. Selected protein biomarker candidates combined with known RA biomarkers from the literature data mining were verified by two different targeted quantification methods; multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM) with Triple-quadrupole (QqQ) and Q-Exactive (QE), respectively. Moreover, we validated the results with 47 individual serum samples using MRM and developed as a multi-biomarker panel. The constructed 4-biomarker panel showed 83% discriminate power in average between two different groups with high AUC value of 0.859. The panel also shows 82% sensitivity and 84% specificity of its innate validity. Collectively, our multi-biomarker panel implies that 4 selected proteins were able to serve as diagnostic assessments to predict the TCZ non-responders in RA patients and possible to supplement serum biomarker discovery-validation process in the clinical field based on integrative proteomic approach.

603

Analysis of electrostatic interaction between ganglioside GM3 and transmembrane peptide

Kazuya Kabayama[1, 2, 3], Yuka Nimura[1], Yuya Asahina[4], Shinya Hanashima[1], Hironobu Hojo[4], Michio Murata[1], Koichi Fukase[1, 2, 3]

1. Dept. of Chem., Grad. Sch. of Sci., Osaka Univ., Toyonaka, OSAKA, Japan

2. MS-CORE, PRC, Grad. Sch. of Sci., Osaka Univ., Toyonaka, OSAKA, Japan

3. Inst. for Radiation Sciences, Osaka Univ., Suita, OSAKA, Japan

4. Inst. for Protein Research, Osaka Univ., Suita, OSAKA, Japan

Theinsulin receptor is a single transmembrane protein that regulates blood glucose level. The reduction of sensitivity to insulin leads to type 2 diabetes. It has been suggested that the accumulation of ganglioside GM3 on cell membrane reduce insulin signal in type 2 diabetes. The interaction of GM3 with the insulin receptor is considered to be due to an electrostatic interaction between a sialic acid of GM3 and the lysine residue of the insulin receptor[1)](Figure 1). However, the detailed analysis has not been elucidated.

In order to observe this interaction, we constructed a model system by incorporating fluorescently labeled insulin receptor transmembrane peptides into liposomes containing GM3.

First, in order to clarify interaction of these molecules in liposomes, we decided the optimal conditions of phase-separating liposomes containing GM3. We found that liposomes consisting of DOPC, GM3 and cholesterol did not undergo phase separation, whereas sphingomyelin (SM) containing liposomes showed phase separation. This result suggests that the charge repulsion between GM3 may affect the liposome phase separation.

Next, we synthesized the transmembrane peptide of insulin receptor. At the isolation of this peptide by HPLC system, the purity was not enough because of its hydrophobicity. We improved this problem by using the isopeptide method[2)]. The synthesized peptides were incorporated into liposomes and observed by 3D imaging. As a result, it became possible to observe the phase state of the whole liposome which was difficult to discriminate by usual 2D imaging. We successfully constructed an evaluation system for detailed analysis.

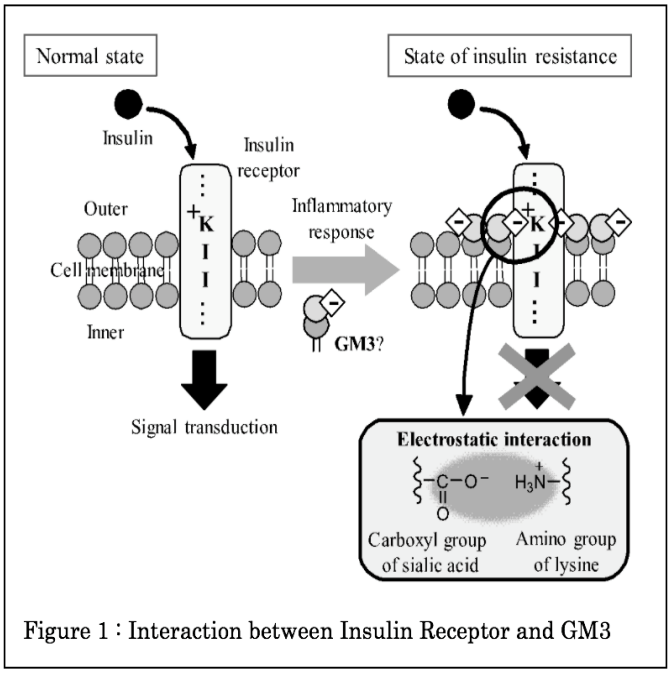


Figure 1 : Interaction between Insulin Receptor and GM3

References

- Kabayama, *et al., Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 13678-13683.
- Sohma, *et al., J. Peptide. Sci.*, **2007**, *88*(2), 253-262.

604

Semisynthesis of α2,6-sialyltransferase bearing a high mannose type oligosaccharide and evaluation of glycosylation functions

Yasuhiro Kajihara[1]

1. Department of Chemistry, Graduate School of Science, Osaka University, Osaka, Japan

Sialyltransferases are essential enzymes for the biosynthesis of sialyloligosaccharides. However, it is known that sialyltransferases expressed in E. coli show very poor enzymatic activities. We consider that the biosynthesis of these series of sialyltransferases requires glycosylation and the glycoprotein quality control (GQC) system, which promotes the native folding of glycoproteins in the endoplasmic reticulum (ER). In order to study the necessity of glycan and GQC in folding of sialyltransferases, we synthesized human β-galactoside α2,6-sialyltransferase (ST6GAL1) with and without a homogeneous high-mannose type oligosaccharide which is a tag of GQC.

The membrane extracellular domain of ST6GAL1 consists of 318 amino acids. We planned to synthesize ST6GAL1 from three peptide segments: sialyltransferase (1-53)-thioester, (54-95)-glycopeptide-hydrazide, (96-318)-peptide, and these segments were prepared by Boc Solid Phase Peptide Synthesis and E. coli expression. High-mannose type oligosaccharide isolated from egg yolk was used for the preparation of the glycopeptide segment. These peptides were ligated from N-terminus to C-terminus by native chemical ligation. Finally, enzymatic active sialyloligosaccharide could be obtained by folding experiments of chemically synthesized ST6GAL1-polypeptide. Kinetic parameters (Km, Kcat) for the chemically synthesized sialyltransferase are evaluated.

ST6GAL1 could be folded as an enzymatic active form, regardless of whether it had a glycan or not, if protein aggregation could be suppressed. Because an expression system in mammalian cell robustly yields an active ST6GAL1, the glycan and GQC seem to suppress aggregation efficiently and to promote the folding of ST6GAL1.

605

Quantitation pipeline evaluation in the context of a multi-run TMT spike-in experiment

Karthik Kamath[1], **Yunqi Wu**[1], **Jemma Wu**[1], **Dana Pascovici**[1], **Matthew J McKay**[1], **Ardeshir Amirkhani**[1], **Mehdi Mirzaei**[1]

1. Australian Proteome Analysis Facility, Sydney, NSW, Australia

The popularity of TMT proteomic analysis has been increasing with several computational pipelines now being available. Novel methods of data analysis including data imputation and normalisation have been proposed, with the more challenging task typically being that of experiments requiring several runs. We also have a previously published statistical analysis pipeline for multi-run TMT analysis termed *TMTPrePro* which uses the quantitation provided by ProteomeDiscoverer.

Here we present an internal spike-in evaluation, comprising a carefully designed spike-in experiment containing three 10-plex TMT replicate runs with 2%, 5% and 10% of know ratio of yeast peptides each spiked into a mice brain tissue lysate. The spiked-

in TMT samples were fractionated by offline HpH chromatography and analysed using a Q-Exactive-HFX mass spectrometer followed by quantification at the MS2 level. The goal was to comprehensively evaluate all sources of variation introduced from the LC-MS/MS analysis through to the statistical data analysis pipeline, and to provide a useful benchmark dataset.

We validated our quantitative pipeline in two scenarios, one including sample ratios to a common reference in each run, and an alternate scenario including cross-batch normalisation for peak areas. We observed internal reference scaling (IRS) normalisation enabled reduction of batch effect enabling good run-to-run concordance with coefficient of variation below 10%. The results confirmed well known aspects such as ratio compression of up to 28-60%. Furthermore, the benchmarking data enabled us to choose an optimal cut-offs for differential expression, including assessing the need for multiple testing corrections. In our hands, 1.2-fold change and uncorrected *p*-values provided the best balance of true/false positives.

606

Dried blood spot samples as a diagnostic and prognostic tool for neonatal sepsis in preterm infants

Azra Karamehmedovic[1, 2], Julie Hibbert[3], Duc Ninh Nguyen[4], Andrew Currie[5, 6], Per T. Sangild[4], Allan Stensballe[1]

1. Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

2. Sino-Danish Center for Education and Research, University of Chinese Academy Sciences, Beijing, China

3. Centre for Neonatal Research & Education and Division of Paediatrics, University of Western Australia, Perth, Australia

4. Department of Veterinary and Animal Sciences, University of Copenhagen, Copenhagen, Denmark

5. Centre for Neonatal Research & Education, University of Western Australia, Perth, Australia

6. Medical and Molecular Sciences, School of Veterinary and Life Sciences, Murdoch University, Perth, Australia

Background & Significance

Neonatal sepsis is associated with short and long-term complications and high mortality if untreated (1). The incidence of neonatal sepsis has been decreasing in extremely low birth weight and preterm infants, yet the incidence of neonatal sepsis associated mortality and length of hospital stay has failed to decrease (2). Rapid and accurate diagnosis based on laboratory and clinical findings remains a challenge despite extensive investigation of the molecular mechanisms behind neonatal sepsis (1). Our novel study investigates the proteome, metabolome and liquid biomarkers such as PTMs and EVs for prognostic markers of late onset sepsis using dried blood spot (DBS), comparing whole blood with plasma in a cross-species comparison.

Methods

DBS and plasma samples from preterm piglet experimental sepsis models and human preterm infants were used for identification of potential diagnostic and prognostic markers. Proteomics and metabolomics investigations were performed on 3 mm punched DBS samples lysed in SDS and analyzed using high-end LC-MS/MS using DIA and PASEF followed by datamining.

Results and Conclusion

Recently, it has been shown that stability of proteins and metabolites in DBS samples persists and may serve as tool for diagnosis and clinical prognostics (3). Our optimized multiple-extraction strategy allowed insight into sepsis markers (>1300 proteins) and differences in whole blood vs. plasma based diagnostics.

1. Ng S, Strunk T, Jiang P, Muk T, Sangild PT, Currie A. Precision Medicine for Neonatal Sepsis. Front Mol Biosci. 2018;5:70.
2. Berkhout DJC, van Keulen BJ, Niemarkt HJ, Bessem JR, de Boode WP, Cossey V, et al. Late-onset Sepsis in Preterm Infants Can Be Detected Preclinically by Fecal Volatile Organic Compound Analysis: A Prospective, Multicenter Cohort Study. Clin Infect Dis [Internet]. 2018 Jun 21 [cited 2019 Apr 25];68(1):70–7. Available from: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciy383/5040327
3. Björkesten J, Enroth S, Shen Q, Wik L, Hougaard DM, Cohen AS, et al. Stability of Proteins in Dried Blood Spot Biobanks. Mol Cell Proteomics [Internet]. 2017 [cited 2019 Jun 14];16(7):1286–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28501802

607

Proteomic responses reveal signal transduction and transcription factor rearrangement to NaCl-induced salinity stress of *Pistachio* Rootstocks

Mohammad Abari[1, 2], Ramesh Katam[1], Rigo X Espina Santiago[1, 3], Ghazala MUSTAFA[4], Nasser Mahna[2]

1. Florida A&M University, Tallaahssee, FLORIDA, United States

2. Department of Horticulture, University of Tabriz, Tabriz, Iran

3. Agro-Environmental Science Department, Universidad de Puerto Rico, Mayaguez, Puerto Rico, United States

4. Department of Plant Sciences, Quaid-i-Azam University, Islamabad , Pakistan

Pistachio (*Pistacia vera* L.), cultivated in arid and semi-arid regions, is one of the most important nuts worldwide. However, the mechanisms underlying salinity tolerance of this plant is not well understood. Studies were carried out both at physiological and molecular level to unravel the metabolic pathways associated with the salt tolerance mechanisms in various cultivars. Five one-year-old pistachio rootstocks were treated with four saline water regimes for 100 days. Salinity decreased total chlorophyll content and carotenoids in the leaves, and ascorbic acid and total soluble proteins in both leaves and roots. Total phenolic compounds, proline, glycine betaine, and H_2O_2 content increased in all rootstocks. Three different ion exclusion strategies were observed in rootstocks, (i) Na+ exclusion in UCB-1, (ii) Cl- exclusion in Badami, (iii) and similar concentrations of Na+ and Cl- in the leaves and roots of Ghazvini, Akbari and Kale-Ghouchi. Rootstock UCB-1 exhibited better intonate to the salinity stress followed by

Badami, Ghazvini, KG, and Akbari. Two contrasting rootstocks to salinity stress UCB-1 and Akbari were used for further comparative proteome studies. Total proteins were isolated from the roots treated with different NaCl concentrations. The proteins were characterized using high throughput LC MS/MS spectrometry searched against *Citrus* database. Over 1600 protein IDs were detected among which, the comparative analysis revealed 153 more abundant and 69 low abundant proteins in UCB-1 and 340 more abundant and 18 low abundant proteins in Akbari. The majority of identified proteins have the functions related to stress responsive proteins, signal transduction, cell wall and cytoskeleton metabolism, and transcription factor. The data suggests a strong linkage of molecular mechanism with the physiological traits in the cultivars with various salt tolerances, and confirmed with enzyme assays and transcript analysis. The results lead to further functional elucidation and genetic engineering.

608

Protein glycosylation in a 2D vs 3D cell culture of ovarian cancer cells

Shauna Winchester[1], Liisa Kautto[1, 2], Nicolle H Packer[1, 2]

1. Macquarie University, NSW, Australia

2. ARC Centre of Excellence for Nanoscale Biophotonics (CNBP), Macquarie University, Sydney , NSW, Australia

Ovarian cancer is the leading gynaecological cause of death in women. Altered glycosylation is a feature of ovarian cancer tumour progression and metastasis with increased levels of high mannose structures and bisecting *N*-glycans found in cancerous cells when compared to non-cancerous cells. The enzyme that is responsible for forming bisecting *N*-glycans is β 1,4-*N*-acetylglucosaminyltransferase III (MGAT3) which, due to the nature of the glycosylation pathway, prevents further branching of glycan structures.

To date the majority of cell-based assays use cells grown in the traditional two-dimensional (2D) monolayer cultured on a flat and rigid surface. The move to three-dimensional (3D) cell culture has become increasingly popular in the research community as it allows studies of cellular responses in settings that more closely resemble the *in vivo* environment. Growing cells on biological scaffolds, eg. decellularised whole organs, in which the extracellular matrix (ECM) has been left intact, provides a 3D cell growth. ECM also contains hormones and growth factors which are important to a cell's ongoing growth and survival.

In this project we investigated the changes in cell physiology and glycan expression on cell membrane proteins of cells grown in three different methods: traditional 2D flat cell culture, 3D spheroid cell culture and 3D tissue scaffold culture. We analysed differences in expressed glycosylation on proteins from three ovarian cancer cell lines and their MGAT3 deleted counterparts. The differences in specific glycan expression between the different cell culture systems were imaged by fluorescent lectins (eg. E-PHA lectin specific for bisecting GlcNAc).

609

Elucidation of the biosynthetic mechanism of fucosylated haptoglobin

Hiromu Kawaguchi[1], Shinji Takamatsu[2], Eiji Miyoshi[2], Miyako Nakano[1]

1. Unit of Biotechnology, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, HIROSHIMA, Japan

2. Department of Molecular Biochemistry & Clinical Investigation, Graduate School of Medicine, Osaka University, Suita, OSAKA, Japan

[Background] Haptoglobin (Hpt) is an acute inflammatory glycoprotein which mainly produced in the liver. We have previously reported that fucosylated Hpt (Fuc-Hpt) is observed in the sera of patients with cancers of various organs, whereas there is a trace of that in the sera of healthy individuals on ground that fucosyltransferases hardly express in the liver. Therefore, Fuc-Hpt is currently under development as a novel cancer biomarker. However, it has not been clarified where and how Fuc-Hpt is produced in cancer patients. We elucidated the production mechanism using cancer cell lines.

[Methods] Human liver cancer cell line and two human colon cancer cell lines were used. Hpt was purified from the cell-culture supernatant using the antibody column of Hpt. Amount of Hpt from each cell line was measured by WB. The cell-culture supernatant of the colon cancer cell lines were added to the liver cancer cell line and were cultured. *N*-Glycans of Hpt were released with PNGase F, and the released *N*-glycans were analyzed by LC-ESI MS.

[Results] Hpt was produced by the liver cancer cell line, but not by the colon cancer cell lines. Hpt was also produced in the liver cancer cell line cultured with the cell-culture supernatant of the colon cancer cell lines. Interestingly, there was difference with fucosylation between these Hpts.

[Conclusion] We concluded that Fuc-Hpt would be not produced by cancer cells of patients with colon cancer because colon cancer cell lines did not produce Hpt. Fucosylation of Hpt produced by the liver cancer cell line co-cultured with conditioned medium of the colon cancer cell line was changed, compared to Hpt produced by the liver cancer cell line. Therefore, it is suggested that the inducing substance which changes the fucosylation of Hpt would be produced by colon cancer cells.

[Keywords] haptoglobin, fucosylation, cancer-biomarker

610

Proteogenomics-based identifying neoantigens in refractory cancers using xenograft mice

Yuko Nakamura[1], Susumu Iizumi[2], Hiroki Ueda[3], Shogo Yamamoto[3], Kenji Tatsuno[3], Aya Nakayama[1], Kazuki Yamamoto[1], Taku Koro[4], Kyoko Hasegawa[2], Norihiro Nakamura[2], Hiroyuki Aburatani[3], Tetsuro Sasada[4], Takeshi Kawamura[1, 3]

1. Isotope Science Center, The University of Tokyo, Bunkyo-Ku, Tokyo, Japan

2. BrightPath Biotherapeutics Co., Ltd., Kawasaki, Kanagawa, Japan

3. Research Center for Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo, Japan

4. Kanagawa Cancer Center Research Institute, Yokohama, Kanagawa, Japan

Background

Recently, neoantigens have been paid great attention among cancer-specific antigens. Neoantigens are self-peptides that are displayed on major histocompatibility complexes (MHC) with cancer specific mutations that are never seen in normal tissues. We performed combined searching by the next generation sequencing (NGS) and mass spectrometry for MHC-associated neoantigens in the refractory cancers.

Methodologies

DNAs and RNAs are extracted and sequenced from cultured colon cancer cell lines and patient tissues (total 9 samples) from pancreatic and biliary cancers. Among the samples, 4 tissues are transplanted to mice to prepare xenograft mice. Proteins are extracted from the xenograft models, immunoprecipitated with anti-MHC antibody, and measured by a mass spectrometer. Peptide sequences are searched against database obtained from the exome and transcriptome from the identical tissue.

Findings

After a systematic examination of experimental conditions using the cultured cells, we have established a workflow that can detect above 1,000 MHC-associated peptides from 3 mg of protein lysates from xenograft mice. We have also identified several neoantigen candidates by DNA and RNA sequencing followed by peptide sequencing with mass spectrometric analysis.

Concluding

In identifying neoantigens, we have found that peptidomics have been added the detection confidence of the identified neoantigens. We are further improving sensitivity to the level to detect directly from patient tissues without xenograft formation. The identified neoantigen candidates are tested for immunogenicity *in vitro*.

611

Promotion of proteomic data sharing through a specialized data journal

Shin Kawano[1, 2], Shujiro Okuda[3], Masaki Matsumoto[4], Yasushi Ishihama[5]

1. Research Organization of Information and Systems, Kashiwa, Japan

2. Toyama University of International Studies, Toyama, Japan

3. Niigata University, Niigata, Japan

4. Kyushu University, Fukuoka, Japan

5. Kyoto University, Kyoto, Japan

Background

Proteomics is a rapidly growing research area that produces large amounts of data, which has led to complex challenges around data management. Proteomics data and datasets need to become more Findable, Accessible, Interoperable, and Re-usable (FAIR[1]) to promote data sharing. To address these challenges and facilitate better proteome data management, the field needs to develop formal structures and procedures. To this end, the Japanese Proteomics Society decided to launch a new data journal – the Journal of Proteome Data and Methods (JPDM).

Preparations

In order to launch JPDM, we prepared the following documents: Publishing plan, Instruction to authors, Guide to reviewers, Ethics policies, License to publish form, and Frequently asked questions. All documents are available at https://www.jhupo.org/jpdm/. We also prepared a submission system (https://mc.manuscriptcentral.com/jpdm/) using the ScholarOne platform.

Features of JPDM

JPDM is a peer-reviewed and fully open access journal provided via the J-stage platform. JPDM publishes four article types: Data descriptors, Protocols, Data processing notes, and Reviews. The main content of the journal is Data descriptors, which is described detailed metadata of mass-based proteomics datasets. JPDM collaborates with ProteomeXchange[2] (PX), which is a consortium of proteome data repositories including jPOST[3], PRIDE[4], MassIVE, etc. When data producers register their datasets in a PX repository then post detailed metadata in JPDM, the journal sends a feedback to the PX repository. JPDM contributes to add a value for proteome datasets by providing detailed metadata.

Conclusion

We launched the new data journal for proteomics community named JPDM. The journal is peer-reviewed and fully open access, and the website is https://www.jhupo.org/jpdm/. JPDM realizes the FAIR principal in the proteomics field by providing rich information. JPDM welcomes lots of submission from community members.

1. Wilkinson MD et al., The FAIR Guiding Principles for scientific data management and stewardship, Sci Data. 3:160018 (2016).
2. Deutsch EW et al., The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res. 45:D1100-D1106 (2017).
3. Moriya Y et al., The jPOST environment: an integrated proteomics data repository and database. Nucleic Acids Res. 47:D1218-D1224 (2019).

4. Perez-Riverol Y et al., The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 47:D442-D450 (2019).

612

Reliable and Deep Proteome Coverage by Gas-Phase Fractionation of Peptides with a FAIMS Pro Interface on a Novel Quadrupole Orbitrap

Julia Kraegenbring[1], Tabiwang N Arrey[1], Michael W Belford[2], Satendra Prasad[2], Markus Kellmann[1], Martin Zeller[1], Alexander Harder[1]

1. Thermo Fisher Scientific, Bremen, BREMEN, Germany

2. Thermo Fisher Scientific, San Jose, CA 95134, USA

Introduction

While great advances have been achieved in the performance of mass spectrometers, deep proteome coverage is still impaired for highly complex samples with high dynamic concentration ranges. To address this challenge a multitude of offline fractionation techniques are employed. However, these are time-consuming and mostly use higher sample amounts. Employing gas-phase fractionation with a FAIMS Pro interface on the new Orbitrap Exploris 480 mass spectrometer mitigates these challenges.

Data

The data demonstrate that the singly charged ion population can be filtered out effectively by applying compensation voltages larger than -40 V. As a consequence, the chemical background can be significantly reduced even in single CV runs. This aspect is especially beneficial for the peptide species eluting at higher organic concentrations because of co-eluting contaminants.

Internal switching between CVs enhances the range of identifiable peptide ion species. With CV switching between -45 V, -55 V, and -65 V, we were able to boost protein identification by as much as 15 % in 90 min LC-gradients and 1 µg HeLa digest accompanied by 30 % more unique peptide IDs. With lower sample loads, we still gain 10% protein IDs even though peptide IDs are slightly decreased which points to the enhanced sampling with FAIMS. In fact, by internal switching of compensation voltages, the number of protein identifications can be increased in a single shot proteomic experiment comparable to a traditional 2D-LC setup.

613

Mutations in HIV GAG Peptides and Their Effect on Cellular Immune Response in Pakistani Patients

Saeed SK Khan[1]

1. Dow University of Health Sciences, Karachi, SINDH, Pakistan

Background

The infected cells present viral antigens to the CTLs through MHC class I, leading to the generation of immune response and destruction of infected cells. Antigen presentation to the CTL involve, proteasomal cleavage, transport to the ER, HLA binding, and presentation to TCR. Response of the CTLs with particular HLA.

Objectives: To study Mutations in *gag* gene of HIV-1 and their effect on proteosomal degradation of peptides.Association of HLA types with selected viral mutations and its effect on the immune response.

Methodology: Samples were collected from patients followed by DNA extraction which was used in HLA typing and in Nested PCR to amplify the HIV *gag* gene followed by sequencing . The sequences were analyzed using bioinformatics tools to identify mutations and for potential proteosomal degradation sites using Netchop. The selected peptides containing the proteosomal degradation sites were synthesized and used in protoeosomal degradation assay. Digested peptides were run on SDS gel to observe the degradation pattern . The digest products were send for MALDI-MS to confirm the exact cutting sites of the peptides.

Results

We identified T303V mutation which is a stronger proteosomal degradation site as compared to T303T in majority of the sequences that we retrieved from database . *The in-vitro* proteosomal degradation assay and MALDI-MS also support these results. The HLA types of the patients were relatively diverse but HLA A*68 was found more prevalent than other HLA types.

Conclusion

The HIV infection in the region is relatively new and the virus is evolving according to the local environment and immune pressure by the patients. We have identified T303V mutation in locally spread virus. We intend to use the peptide containing this particular residue in ELISpot experiments to further elucidate the interaction between HIV and the host immune system.

614

Deciphering host-pathogen-drug interactions: optimisation of antibiotic therapy

Zhi Ying Kho[1], Meiling Han[1], Yan Zhu[1], Cheng Huang[2], Ralf Schittenhelm[2], Joel Selkrig[3], Mohammad Azad[1], Jian Li[1]

1. Biomedicine Discovery Institute, Infection and Immunity Program, Department of Microbiology, Monash University, Clayton, Victoria, Australia

2. Monash Proteomics & Metabolomics Facility and Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

3. European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg, Baden-Württemberg, Germany

Background: Opportunistic pathogen *Acinetobacter baumannii* is a leading cause of nosocomial infections worldwide. Polymyxins are the last-resort antibiotics to treat multidrug resistant (MDR) *A. baumannii* infections. However, the elusive complex interplay among human innate immunity, *A. baumannii* and polymyxins has limited the optimisation of current antibiotic therapy.

Aim: This study aimed to elucidate this tripartite crosstalk using a label-free liquid chromatography-mass spectrometry-based proteomics approach.

Methods: THP-1 human macrophages were infected by a highly virulent MDR *A. baumannii* isolate AB5075 at multiplicity of infection (MOI) 1000 for 3 h, followed by another 1 h polymyxin B (30 mg/L) treatment. Protein expression profiling of both macrophages and interacting bacteria were conducted.

Results: Our results identified 202, 549 and 607 differentially expressed bacterial proteins post macrophage monotherapy, polymyxin monotherapy, and macrophage-polymyxin combination treatment, respectively. Notably, macrophage-polymyxin combination augmented the number of down-regulated bacterial proteins involved in the following fitness-related pathways compared to each individual treatment: (i) cell wall/membrane/envelope biogenesis (KdsD, KdsB biosynthetic enzymes of lipopolysaccharide); (ii) translation, ribosomal structure and biogenesis (aminoacyl-tRNA synthetases); (iii) central metabolism (transport and metabolism of amino acids, lipids, carbohydrates, nucleotides); and (iv) energy production (NuoA, PPA, CyoB enzymes in oxidative phosphorylation). Besides, macrophage-polymyxin combination enhanced down-regulation of regulatory and transport proteins in bacterial two-component systems (BaeR, PstS), and siderophore biosynthetic enzymes (EntABCE) associated with iron uptake. Our results indicate reduced bacterial adaptive capability particularly in macrophage-polymyxin combination, and dysregulated iron homeostasis necessary for bacterial survival and virulence. The ongoing macrophage proteomics analysis will facilitate identification of key candidate host factors and provide a better picture in depicting such tripartite crosstalk.

Conclusion: This is the first proteomics study revealing the adaptive response of *A. baumannii* in macrophage and/or polymyxin treatment. Our mechanistic findings will contribute to optimising polymyxin therapy in patients.

615

FusionPro, a versatile proteogenomic tool for identification of novel fusion transcripts and their potential translation products in cells

Chae-Yeon Kim[1], Keun Na[1], Saeram Park[1], Seul-Ki Jeong[1], Jin-Young Cho[1], Heon Shin[1], Min Jung Lee[1], Gyoonhee Han[2], Young-Ki Paik[1]

1. Yonsei Proteome Research Center, Yonsei University, Seoul, Seodaemun-gu, Republic of Korea

2. Department of Pharmacy, College of Pharmacy, Yonsei University, Seoul, Seodaemun-gu, Republic of Korea

Fusion proteoforms are translation products derived from gene fusion. Although very rare, the fusion proteoforms play important roles in biomedical science. For example, fusion proteoforms influence the development of tumors by serving as cancer markers or cell cycle regulators. Although numerous studies have reported bioinformatics tools that can predict fusion transcripts, few proteogenomic tools are available that can predict and identify proteoforms. In this study, we develop a versatile proteogenomic tool “FusionPro”, which facilitates the identification of fusion transcripts and their potential translatable peptides. FusionPro provides an independent gene fusion prediction module and can build sequence databases for annotated fusion proteoforms. FusionPro shows greater sensitivity than the available fusion finders when analyzing simulated or real RNA sequencing datasets. We use FusionPro to identify 18 fusion junction peptides and three potential fusion-derived peptides by MS/MS-based analysis of leukemia cell lines (Jurkat and K562) and ovarian cancer tissues from the Clinical Proteomic Tumor Analysis Consortium. Among the identified fusion proteins, we molecularly validate two fusion junction isoforms and a translation product of *FAM133B:CDK6.* Moreover, sequence analysis suggests that the fusion protein participates in the cell cycle progression. In addition, our prediction results indicate that fusion transcripts often have multiple fusion junctions and that these fusion junctions tend to be distributed in a non-random pattern at both the chromosome and gene levels. Thus, FusionPro allows users to detect various types of fusion translation products using a transcriptome-informed approach and to gain a comprehensive understanding of the formation and biological roles of fusion proteoforms.

616

Identification of cancer stem cell specific transcription factors of B7H3 using DP-MS

HanByeol Kim[1], Yu-ri Seo[1], Hyeryeon Jung[1], Eugene C Yi[1]

1. Global Translational Research on Molecular Medicine and Biopharmaceutical Sciences, Seoul National University, Korea, Jongno-gu, SEOUL, South Korea

B7H3 is a member of the B7 family and is known as an immune checkpoint molecule. B7H3 is known to inhibit T cell function by being expressed in immune cells such as APCs and macrophages. Recent studies have shown that B7H3 is expressed in various cancer cells which affects tumor growth and EMT (epithelial–mesenchymal transition) as well as immune evasion of cancer cells.

In addition, our previous study showed that the expression of B7H3 was elevated in cancer stem cell (CSC) derived from MDA-MB453 breast cancer cells compared with non-cancer stem cell (NCSC) through in-situ cell surface biotinylation. This suggests that B7H3 expressed in CSC plays a major role in cancer stem cell immunity, progression and therapeutic resistance of CSC. We have investigated the mechanism by which the expression of B7H3 is increased in CSC and found that transcription of B7H3 was elevated in CSC. Here we performed a DNA pull-down mass spectrometry (DP-MS) with a biotinylated promoter region of B7H3 with CSC and NCSC cell lysates and identified specific transcription factors regulating B7H3 expression in CSC. In this study, we present novel transcription factors that bind to the promoter of B7H3 in CSC using DP-MS method.

617

Comparative proteome analysis for three-dimensional (3D) spheroids and conventional two-dimensional (2D) cultured ovarian cancer cell revealed the differential pathway alteration

Hee-Sung Ahn[1], Jeonghun Yeom[1], Jiyoung Yu[1], Yumi Oh[1], Hwangkyo Jeong[1], Yujin Jo[1], Sangyong Jon[2], Kyunggon Kim[1]

1. Asan Medical Center/University of Ulsan, Songpa-gu, SEOUL, South Korea

2. Dept. of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea

In previous study, tumor cell under three-dimensional (3D) spheroids culture condition makes different phenotype compared to two-dimensional (2D) culture condition. To make a reliable cell line experiment, spheroid culture condition have been thought to be reasonable but there have been no clear molecular characterization about this issue. In this study, comparative proteome analysis between spheroid cell cultured using home-developed membrane (poly(2,4,6,8-tetravinyl-2,4,6,8- tetramethyl cyclotetrasiloxane) (pV4D4), and commercial hydrophilic membrane ultra-low-attachment (ULA). Conventional 2D monolayer-cultured control cell was also analysed using high resolution LC-MS platform coupled by multiplexing peptide labeling technology. And quantitative proteome of conditioned media from each culture condition were analyzed to investigate the difference at secretory part. As a result, EGFR pathway and mitochondrial dysfunction pathway is involved in pV4D4-based spheroid culture condition compare to other culture condition. From this study, we could conclude that pV4D4-based spheroid cultured cell shows stem-like characteristics and that this 3D spheroid culture can be a useful tool to investigate cancer stem cell and its resistance to chemotherapy.

618

A mitochondrial proteome profile indicative of type 2 diabetes mellitus in skeletal muscles

Su-Jin kim[1], Sehyun Chae[2], Young Do Koo[3], Hokeun Kim[1], Byung Yong Ahn[3], Daehee Hwang[2], Kyong Soo Park[3], Sang-Won Lee[1]

1. Department of Chemistry, Center for Proteogenome Research, Korea University, Seoul, Republic of Korea

2. Department of New Biology, Daegu Gyeongbuk Institute of Science and Technology, Daegu, Republic of Korea

3. Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

The pathogenesis of type 2 diabetes mellitus (T2DM) is closely associated with mitochondrial functions in insulin responsive tissues. The mitochondrial proteome, compared with the mitochondrial genome, which only contains 37 genes in humans, can provide more comprehensive information for thousands of mitochondrial proteins regarding T2DM-associated mitochondrial functions. However, T2DM-associated protein signatures in insulin-responsive tissues are still unclear. Here, we performed extensive proteome profiling of mitochondria from skeletal muscles in nine T2DM patients and nine nondiabetic controls. A comparison of the mitochondrial proteomes identified 335 differentially expressed proteins (DEPs) between T2DM and nondiabetic samples. Functional and network analyses of the DEPs showed that mitochondrial metabolic processes were downregulated and mitochondria-associated ER membrane (MAM) processes were upregulated. Of the DEPs, we selected two (NDUFS3 and COX2) for downregulated oxidative phosphorylation and three (CALR, SORT, and RAB1A) for upregulated calcium and protein transport as representative mitochondrial and MAM processes, respectively, and then confirmed their differential expression in independent mouse and human samples. Therefore, we propose that these five proteins be used as a potential protein profile that is indicative of the dysregulation of mitochondrial functions in T2DM, representing downregulated oxidative phosphorylation and upregulated MAM functions.

Acknowledgements
This study was supported by the Collaborative Genome Program for Fostering New Post-Genome Industry of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (MSIT) (NRF-2017M3C9A5031597)

619

Combining efforts of the Russian proteomic consortium in missing proteins' mining

Olga I. Kiseleva[1], Ekaterina V. Poverennaya[1], Ekaterina V. Ilgisonis[1], Svetlana E. Novikova[1], Arthur T. Kopylov[1], Alexander I. Archakov[1], Elena A. Ponomarenko[1]

1. Orekhovich Institute Of Biomedical Chemistry, Moscow, MOSCOW, Russia

Background

Despite direct or indirect efforts of proteomic community, the fraction of blind spots on the protein map is still significant. Almost 14% of human master proteins () have no experimental validation up to now. Apparently, proteomics has reached the stage where all easy scores are achieved, and every next protein identification requires more intension and curiosity in expansion of unusual types of biomaterial and/or conditions.

Methodologies

We encompassed omnigeneous data obtained mass spectrometrically by members of Russian Proteomic community on more than 25 types of non-trivial biological samples and cell lines. These data were processed in a uniform manner by three search engines (X!Tandem, MS-GF+, OMSSA) being a part of SearchGUI package. We accompanied MS-data with the results of RNASeq, neXtProt and GPMdb analyses to estimate probability of unique peptide detection and hence the possibility to identify protein.

Findings

The study resulted in detection of 7 missing proteins with two peptides. Moreover, 149 missing proteins were detected with single proteotypic peptide. We analyzed the gene expression levels to suggest feasible targets for further validation of missing and uncertain protein observations, which will fully meet the requirements of the international consortium.

Conclusion

All proteins are on unequal terms from the very beginning: some of them have competitive advantage to be recaptured. Each case of protein identification is unique (and missing protein – especially) and requires custom-tailored approach and criteria of reliabilty. We believe that non-standard methodological and bioinformatical solutions applied to unusual biomaterials will be fruitful sources of unknown fragments of the proteomic puzzle. Taking into account this creative task, we invite the international community to reconsider specific quality requirements for the proteomic data to “missing” and uncertain proteins.

62

Synthesis of glycoconjugate vaccines and antibodies for new cancer immunotherapies

Koichi Fukase[1], Yoshiyuki Manabe[1], Kazuya Kabayama[1], Julinton Sianturi[1], Li-Ting Chiu[2], Tsung-Che Chang[1], Kento Tokunaga[1], Masahiro Tanemura[3], Eiji Miyoshi[4], Shang-Cheng Hung[2], Yuka Nimura[1], Yukari Fujimoto[5], Shino Ohshima[6], Yoshie Kametani[6], Chun-Cheng Lin[7]

1. Department of Chemistry, Graduate School of Science, Osaka University, Toyonaka, OSAKA, Japan

2. Genomics Research Center, Academia Sinica, Taipei, Taiwan

3. Department of Gastroenterological Surgery, Osaka Police Hospital, Osaka, Japan

4. Department of Molecular Biochemistry and Clinical Investigation, Graduate school of Medicine,, Osaka University, Suita, Osaka, Japan

5. Department of Chemistry, Faculty of Science and Technology, Keio University, Yokohama, Kanagawa, Japan

6. Faculty of Medicine, School of Medicine, Tokai University, Isehara, Kanagawa, Japan

7. Department of Chemistry, National Tsing Hua University, Hsinchu , Taiwan

We have investigated the synthesis of various microbial components that activate or modulate immune responses via innate immune receptors. Here, we report on the application of their immunomodulatory effects to the developments of immune adjuvants and cancer vaccines.Self-adjuvanting vaccines consisting in antigens and adjuvants have been investigated to improve vaccine efficacy and safety. We synthesized self-adjuvanting cancer vaccines composed of a trimeric tumor-associated carbohydrate antigen SialyI-Tn (STn), a TLR2 agonist Pam3CSK4 as an adjuvant, and a T-helper epitope, since clustered STn antigens are highly expressed on many cancer cells. Immunization of vaccines in mice induced the anti-triSTn IgG antibodies, which recognized triSTn-expressing cell lines PANC-1 and HepG2[1].

Humans do not have α-gal epitope but have natural antibodies (Ab) against α-gal. The reaction of anti-Gal Abs with α-gal causes hyperacute rejection in xenogeneic organ transplantation. In this study, we applied the ADC (antibody-drug conjugate) concept to develop a novel α-gal-based tumor immunotherapy.The α-gal trisaccharide was efficiently synthesized by one-pot or one-flow glycosylation procedures. After chemical synthesis of α-gal, we achieved concise chemical conjugation of antibodies with α-gal. The α-gal-Ab conjugates showed increased complement dependent cytotoxicity (CDC) by recruiting natural Abs through the interaction between α-gal and anti-Gal Abs. Utilization of α-gal dendrimer allowed introduction of large amounts of α-gal epitope to the antibody. The conjugated antibodies with α-gal dendrimer showed potent CDC without loss of affinity for the target cell. The method described here will enable re-development of antibodies to improve their potency[2].

1. [1] Chang, T-C.; Manabe, Y.; Fujimoto, Y.; Ohshima, S.; Kametani, Y.; Kabayama, K.; Nimura, Y.; Lin, C-C.; Fukase, K. Angew. Chem. Int. Ed. Engl. 2018, 57, 8219-8224.
2. [2] Sianturi, J.; Manabe, Y.; Li, H-S.; Chiu, L-T.; Chang, T-C.; Tokunaga, K.; Kabayama, K.; Tanemura, M.; Takamatsu, S.; Miyoshi, E.; Hung, S-C.; Fukase, K. Angew. Chem. Int. Ed. Engl. 2019, 58, 4526-4530.

620

The use of unique peptide-tags for sensitive detection of yeast proteome

Keiji Kito[1]

1. Meiji University, Kawasaki, KANAGAWA, Japan

Technical advances in mass spectrometry (MS)-based proteomic analysis have enabled proteome-wide analysis with high coverage and sensitivity. Detection of individual proteins generally relies on peptide identification via tandem mass analysis. Therefore, successful detection of each protein depends on the presence of proteolytic peptide sequences that are suited for detection and identification by MS analysis. However, a subset of proteins, especially of small molecular weight or consisting of biased sequence, contain few peptide sequences that can be detected with high sensitivity and generate good quality of tandem mass spectra. In addition, extreme complexity in protein sample makes it difficult to detect low-abundance proteins with high dynamic range.

To overcome these problems, we have developed a strategy using a genetically modified yeast strain in which each protein is tagged with unique peptide-tag suited for MS detection and identification. Only peptide-tags are purified and subjected to MS analysis, leading to a reduction in sample complexity and allowing for sensitive detection even in the case of low-abundance proteins. A single yeast strain in which unique peptide-tags, which were selected from peptide library generated from DNA template produced by degenerate PCR and chemically synthesized, were fused to multiple proteins was successfully constructed with high efficiency by the use of CRISPR/Cas9 system. We present a proof of concept of this strategy using such a yeast strain best-optimized for proteomic analysis.

621

Suppression of terminal modifications of N-glycans by bisecting GlcNAc

Yasuhiko Kizuka[1], Miyako Nakano[2], Naoyuki Taniguchi[3]

1. Gifu University, Gifu, GIFU, Japan

2. Hiroshima University, Hiroshima, Japan

3. Osaka International Cancer Institute, Osaka

Bisecting GlcNAc, a branched GlcNAc in N-glycan synthesized by GnT-III (encoded by MGAT3 gene), is most highly expressed in brain. Our previous studies using Mgat3-deficient mice showed that bisecting GlcNAc promotes Alzheimer's pathology by positively regulating the functions of amyloid beta-producing enzyme BACE1. However, physiological functions of bisecting GlcNAc remain largely unclear.

Several in vitro biochemical studies already showed that introduction of bisecting GlcNAc suppresses further actions of other GlcNAc-branching enzymes, such as GnT-IV and -V, suggesting that bisecting GlcNAc regulates overall N-glycan profiles in vivo. To clarify this issue, we analyzed N-glycan structures of Mgat3-deficient brain and found that various terminal modifications were upregulated in Mgat3-knockout, such as Lewis-fucose, sialic acid and HNK-1 epitopes. The levels of glycosyltransferase mRNAs and their donor substrates were unchanged between WT and Mgat3-knockout. In contrast, our enzyme assays using bisected and non-bisected acceptor substrates revealed that most enzymes working on N-glycan terminals prefer the non-bisected glycan as a substrate, indicating that the upregulation of terminal N-glycan epitopes were attributed tosubstrate specificity of glycosyltransferases. Molecular dynamics simulation of glycosyltransferase-glycan complexes showed that the presence of bisecting GlcNAc changed overall N-glycan conformation from extended type to back-fold type where alpha1,6-arm loses interaction with the enzymes, suggesting that the bisected glycan is not preferred by various glycosyltransferases for N-glycan terminals due to alteration of N-glycan conformation. These results highlight the roles of bisecting GlcNAc as a general suppressor of terminal modifications of N-glycans, providing new insight into how protein N-glycosylation is regulated in cells.

622

Proteomics for vaccine discovery: development of two potential whole protein vaccines with efficacy against *Pseudomonas aeruginosa*

William Klare[1], Paula Niewold[2], Claudio Counoupas[2], Jamie Triccas[2], Stuart Cordwell[1, 3, 4]

1. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

2. Department of Infectious Diseases and Immunology, University of Sydney, Sydney, NSW, Australia

3. Discipline of Pathology, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

4. Sydney Mass Spectrometry, University of Sydney, Sydney, NSW, Australia

One of the biggest challenges to quality of life for an individual with Cystic Fibrosis (CF) is the high rate of incidence of infection with the opportunistic pathogen *Pseudomonas aeruginosa*. Infections are typically lifelong, causing significant morbidity and mortality, and lead to high rates of divergent within-host evolution, often resulting in the presence of multiple infection phenotypes that are multidrug resistant. This renders traditional therapies and interventions ineffective, with no efficacious vaccine available to date. We profiled within-host adaptation by investigating a pair of isogenic clonal epidemic isolates (AES-1R and AES-1M), isolated from the same patient 11 years apart. Using an integrated multi-omic strategy, we investigated cellular differences by parallel proteomics, metabolomics and lipidomics when grown in an artificial sputum-like medium that reflects the physiology of the CF lung. This near-complete cellular characterisation was followed by time-resolved proteomics within the same media to identify proteins crucial to initiation of infection as possible vaccine candidates. Candidate selection was further refined with the aid of the comparative integrated multi-omics data, known recognition by human sera, *in vivo* expression and comprehensive bioinformatic characterisation. Three proteins were identified, synthesised as peptides, and subsequently screened for protection against infection, as well as mechanisms of immunogenicity via ELISA and flow cytometry. We show high levels of efficacy (99.9% reduction) for two out of three protein vaccines against *P. aeruginosa* in a mouse infection model. Presented is a workflow for identification and optimisation of peptide vaccine candidates that is scalable, effective and widely applicable to other culturable bacterial species.

623

Complete multi-omic physiological characterisation of clonal *Pseudomonas aeruginosa* adaptation to the CF lung

William Klare[1], Paula Niewold[2], Brian Forde[3], Scott Beatson[3], Jamie Triccas[2], Stuart Cordwell[1, 4, 5]

1. School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

2. Department of Infectious Diseases and Immunology, University of Sydney, Sydney, NSW, Australia

3. School of Chemistry and Biological Sciences, University of Queensland, Brisbane, QLD, Australia

4. Discipline of Pathology, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

5. Sydney Mass Spectrometry, University of Sydney, Sydney, NSW, Australia

One of the biggest challenges to quality of life for an individual with Cystic Fibrosis (CF) is the high rate of incidence of infection with the frequent isolate *Pseudomonas aeruginosa*. Infections are typically lifelong, resulting in significant morbidity and mortality, and lead to high rates of divergent within-host evolution, often resulting in the presence of multiple infection phenotypes. This renders traditional therapies / interventions ineffective. We profiled within-host adaptation by investigating a pair of isogenic clonal epidemic isolates (AES-1R and AES-1M), isolated from the same patient 11 years apart. Using a combined -omic strategy, isolates were grown in an artificial sputum-like medium that reflects the physiology of the CF lung. We undertook intra- and extra-cellular comparisons at the proteomics level using offline HILIC peptide fractionation coupled to reversed phase LC-MS/MS, which enabled the identification of functional clusters associated with virulence that were highly expressed only in the initial colonizing isolate. Proteome data was integrated with metabolome and lipidome data extracted from the same cells, using a targeted (MRM) and untargeted mass spectrometric method, respectively. Integration of multiple -omics revealed distinct metabolic preferences between the two strains that reflects niche adaptation. Differences in virulence capabilities were assessed through confocal scanning laser microscopy to investigate changes in biofilm structure, cell membrane glycolipid analysis by MALDI-TOF MS, as well as virulence in the classical *C. elegans* slow-killing and murine lung-infection models. Host-pathogen interactions were investigated using 16-colour flow cytometry within the murine lung, in tandem with qPCR to quantify *in vivo* virulence network expression. The results obtained provide one of the most comprehensive assessments to date of the consequences of over a decade of within-host evolution on the cellular physiology of *P. aeruginosa* in adaptation and persistence within the context of CF.

624

Matrix assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) for monitoring of drug response in primary cancer spheroids

Parul Mittal[1], Zoe K Price[2], Noor Alia Lokman[2], Carmela Ricciardelli[2], Martin Oehler[3], Manuela Klingler-Hoffmann[4], Peter Hoffmann[4]

1. Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adealaide, SA, Australia

2. Discipline of Obstetrics and Gynaecology, Adelaide Medical School, Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia

3. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

4. University of South Australia, Mawson Lakes, SA, Australia

Malignant ascites is a fluid, which builds up in the abdomen and contains cancer cells in the form of single cells or multicellular clusters called spheroids. Malignant ascites has been observed in patients suffering from ovarian, cervical, gastric, colorectal, pancreatic, endometrial or primary liver cancer. The spheroids are believed to play a major role in chemo resistance and metastasis of the cancer. To ease the discomfort of patients, malignant ascites (MA) is often drained from the abdomen using a procedure called paracentesis. MA retrieved via this minimal invasive procedure, is a great source for cancer spheroids, which can be used for testing chemotherapeutic drugs and drug combinations. We have optimised a workflow for concurrent monitoring of drug accumulation, drug response and drug metabolites in cancer spheroids. This new approach might be used to choose the best targeted therapy for each patient and thereby facilitate personalised medicine.

625

Analysis of the interaction between NF1-associated factor TCTP and translation elongation factors by cross-linking mass spectrometry coupled with affinity purification

Daiki Kobayashi[1], Kota Mukugi[1], Takaho Tokuda[1], Norie Araki[1]

1. Kumamoto University, Kumamoto, KUMAMOTO, Japan

Neurofibromatosis type 1 (NF1) is an autosomal dominant disease that predisposes individuals to developing benign neurofibromas and malignant peripheral nerve sheath tumors (MPNST). The mechanism of NF1-tumoigenesis or the curatives have not been established. Using iPEACH (Kobayashi *et al*, MCP 2013), we previously found that translationally controlled tumor protein (TCTP) is a novel biological target for NF1-associated tumors (Kobayashi *et al*. JBC 2014), and the complex formation of TCTP and protein elongation factors (EF1A2 and GEF proteins for EF1A2) was identified as a translational regulator specific in NF1 associated tumors (Kobayashi *et al*, MCP 2019). In this study, cross-linking mass spectrometry combined with affinity purification (AP) using FLAG-antibody was utilized to analyze the structural insight in the complex of TCTP and elongation factors.

For AP analysis, FLAG-tagged TCTP and EF1A2 were overexpressed into Hela cells, and the Flag-TCTP- and Flag-EF1A2-binding proteins were co-immunoprecipitated using magnetic beads coupled with FLAG-antibody. These proteins were treated with cross-linker disuccinimidyl sulfoxide (DSSO) and digested with trypsin on beads. The tryptic cross-linked peptides were analyzed by XlinkX workflow (Lui *et al.* Nat Methods 2015) using nanoLC-Orbitrap Fusion ETD system.

Cross-linking mass spectrometry analyses specifically identified the 76 and 140 DSSO-crosslinked peptides in TCTP- and EF1A2-FLAG-IP fractions (FDR < 1% and XlinkX score > 20). Totally, eleven peptides crosslinked in the complex of TCTP-EF1A2-elongation factors were identified with high confidence. Among them, the crosslinking between K19 in TCTP and K41/K44 in EF1A2 was observed. These data suggest that TCTP inhibits the EF1A2 homodimer formation and mediates the GDP/GTP exchange reaction on EF1A2 via the direct binding to EF1A2 near the GTP-binding site.

These results could provide a novel treatment strategy for the NF1-associated tumors targeting the translational machinery driven by the specific TCTP-EF1A2 interaction.

1. 1) M. Hirayama et al., Mol. Cell. Proteomics, 12, 1377-94 (2013). 2) D. Kobayashi et al., J. Biol. Chem., 289, 26314-26 (2014). 3) D. Kobayashi et al., Mol. Cell. Proteomics, 18, 245-262 (2019). 4) F. Lui et al., Nat Methods., 12, 1179-84 (2015)

626

Adipose Stem Cell Characterization

Naomi Koh Belic[1], Jerran Santos[1], Bruce Milthorpe[1], Matthew Padula[1]

1. University of Technology Sydney, Broadway, NSW, Australia

Human adipose stem cells are widely used as treatments for a plethora of disorders, despite the minimal scientific evidence of safety let alone efficacy. These unproven stem cell treatments being offered by predatory clinics around the globe, more than 60 of which are practicing in Australia. Our understanding of human adipose stem cells is limited, in particular very little is known about their phenotype, as their proteome has not previously been characterised. In order to identify this valuable proteome, this project isolated and characterised human adipose stem cells from 10 healthy patient's abdominal lipoaspirates. These cells were passaged to achieve a homogenous cell culture and their proteome was characterised through a comprehensive analysis of cellular proteins, extracellular vesicles and secreted cytokines. The samples were run on our Q Exactive™ Plus Orbitrap Mass Spectrometer resulting in a quantitative proteomic profile of human adipose stem cells. Identified cellular proteins provide vital insight into cellular function, while analysis of membrane bound proteins provided an extensive catalogue of cell surface markers that are useful for antibody-based assay development. The stem cell derived extracellular vesicle proteome was also examined because stem cells secrete extracellular vesicles in substantial quantities and they are known to play a significant role in cancer, injury healing and immune suppression. Secreted cellular proteins such as cytokines also facilitate cellular communication of immune signals and warrant investigation. 27 cytokines were investigated through the utilisation of a Multiplex Immunoassay. This study produced a comprehensive data set of human adipose stem cell proteins, which is a unique resource that ultimately investigates the biological phenotype of human adipose stem cells. This is an invaluable tool for researchers and clinicians as it will assist in developing this much needed understanding of human adipose stem cells that are already being used for clinical applications.

627

Method development for clean-up of OCT-embedded tissue samples using mixed mode SPE cartridges

Jennifer M. S. Koh[1], Natasha Lucas[1], Dylan Xavier[1], Andrew Robinson[1], Steve G. Williams[1], Peter G. Hains[1], Phil J. Robinson[1]

1. ProCan, Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

Solid phase extraction (SPE) is important in proteomics workflows to remove impurities and concentrate the peptides before they are analysed by liquid chromatography-mass spectrometry (LC-MS). Optimal cutting temperature (OCT) is a specimen matrix of liquid polymers used for cryostat sectioning of fresh frozen tissues. However, OCT must be removed prior to LC-MS since the polymers compromise the equipment and analyses. OCT is typically removed by tedious washing steps over 1-2 hours, with some degree of sample loss with each wash step. Tissue lysis and digestion follows OCT removal then the digests are passed through reversed-phase C18 cartridges. C18 resin has a strong affinity for OCT (i.e. polymers); hence if any OCT remains after washing, it will outcompete the peptides, leaving a poorly recovered sample which is highly concentrated with OCT and not suitable for LC-MS.

We present a new SPE method using mixed mode strong cation exchange (MCX) cartridges for effective OCT removal. We optimised the method to maximise the removal of OCT and minimise the loss of peptides during SPE. OCT-embedded 30 µm rat kidney sections and human prostate tumour sections were lysed, digested and cleaned up using MCX cartridges. The digests were analysed on Triple TOF 6600 mass spectrometers equipped with micro-LC systems (SCIEX). The MCX cartridges effectively removed OCT from all digests. The presence of up to 85% OCT in the initial tissue samples did not affect lysis or digestion efficiency performed in a Barocycler. The identified proteins and peptides were comparable to those obtained from samples cleaned with C18 (after standard OCT removal).

MCX SPE allows tissue samples to be processed and analysed in a much shorter time frame by completely eliminating manual washing of OCT-embedded sections. The new method is reliable and reproducible, and also minimises the operator's risk of exposure to biological samples.

628

Identification of glioblastoma-selective secreted proteins by secretome analysis of cancer cell lines

Tomohiro Kohata[1], Shingo Ito[1, 3, 2], Takeshi Masuda[1, 3, 2], Mitsutoshi Nakada[4], Sumio Ohtsuki[1, 3, 2]

1. Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, JAPAN

2. AMED-CREST, Tokyo, JAPAN

3. Faculty of Life Sciences, Kumamoto University, Kumamoto, JAPAN

4. Department of Neurosurgery, Kanazawa University, Kanazawa, JAPAN

Background

Glioblastoma (GBM) is known as malignant brain tumor, and GBM biomarkers are required to detect cancer at earlier stage for improving therapeutic outcomes. However, GBM biomarkers possessing the potential for clinical applications have not been identified yet. In the previous studies, GBM biomarkers were screened from body-fluid. However, abundant proteins in the body-fluid interfere identification of proteins secreted from tumor tissues. Therefore, the purpose of the present study was to identify GBM-selective secreted proteins from the conditioned-mediums (CMs) of GBM cells and validated their expression in the cerebrospinal fluid (CSF) of GBM or non-brain tumor patients.

Methods

GBM (U87, U251, T98G) and non-GBM (MDA-MB-231, MCF-7, Caco-2) cells were cultured with FBS-free medium. After incubation, CMs were collected and the proteins were concentrated by acetone precipitation. The peptides digested by Lys-C and trypsin were identified by data-dependent acquisition and quantitated by data-independent acquisition (SWATH-MS) using TripleTOF5600 (SCIEX, USA). The identified proteins were quantified by targeted proteomics with spiking internal-standard peptides by MRM of QTRAP6500 (SCIEX).

Preliminary data

We identified 2,371 proteins from the CMs, and 1,338 proteins were identified by more than 3 peptides. Among 1,338 proteins, 19 proteins were detected only in the CMs of GBM cells. To validate the GBM-selective secretion, the expression of 19 proteins were quantified in the CMs and CSF by targeted proteomics. As a result, 19 proteins exhibited higher expression levels in CMs of GBM cells than those of non-GBM cells. The expression of laminin subunit alpha-4 (LAMA4) and osteopontin (OPN) were significantly greater in CSF of 22 GBM patients than in 11 non-brain tumor patients and showed AUC greater than 0.86. Therefore, LAMA4 and OPN were suggested to be secreted from GBM tissues. In conclusion, LAMA4 and OPN were identified GBM-selective secreted proteins, and are candidates as biomarkers and therapeutic targets for GBM.

629

Search for novel diagnostic markers of pancreatic ductal adenocarcinoma using secretome and proteome analysis

kaori konno[1], Shigetsugu Takano[2], Yoshio Kodera[3], Hideyuki Yoshitomi[2], Masayuki Ohtsuka[2], Fumio Nomura[4], Kazuyuki Sogawa[1]

1. Department of Biochemistry, Azabu univercity, Sagamihara, Kanagawa, Japan

2. Department of General Surgery, Chiba University, Chiba-city, Chiba, Japan

3. Center for Disease Proteomics, Kitasato University, Sagamihara-city, Kanagawa, Japan

4. Divisions of Clinical Mass Spectrometry and Clinical Genetics, Chiba University Hospital, Chiba-city, Chiba, Japan

Pancreatic ductal adenocarcinoma (PDAC) accounts for 85–90% of all pancreatic tumours. The median survival of all PDAC patients is less than 6 months, and the recent 5-year-survival rate is approximately 8%. One of the most crucial reasons for the poor prognosis is the lack of early diagnostic markers for PDAC. To overcome this and improve the outcomes of patients with PDAC, there is an urgent need to identify highly sensitive and specific markers for early detection. The widely used serum-circulating markers for PDAC, carbohydrate tumour–associated antigen 19-9 (CA19-9) and carcinoembryonic antigen (CEA), are not sufficiently accurate to be used as early diagnostic markers. In the present study, we employed the TMT method to generate comparative protein profiles of PDAC cell lines. Furthermore, we compared the serum levels of candidate proteins to evaluate the ability to discriminate between PDAC and healthy controls. Using tandem mass tag labelling and LC-MS/MS, we performed comparative analyses of secreted proteins in culturing media obtained from human PDAC cell lines (MIA PaCa-2, PANC-1, and BxPC-3) to identify serum biomarkers for PDAC. In validation studies, we evaluated the discriminatory power of candidate proteins. Of 27 proteins, 13 proteins were identified with the increase of protein levels among these all three cell lines, and Protein X was selected for further analysis. The sera levels of Protein X were significantly higher in the preoperative PDAC patients than those in the postoperative ones ($P<0.001$). In future study, we will verify this result with more number of samples, and pursue the usefulness of Protein X as a novel diagnostic marker in PDAC.

63

Stereo-selective synthesis of sialic acid containing glycoconjugates

Hiromune Ando[1]

1. Gifu University, Gifu-shi, GIFU PREFECT., Japan

The synthesis of sialic acid containing glycoconjugates remains a challenge in carbohydrate chemistry, impeding the cultivation of biological and therapeutic potentials of sialo-glycans. We have long been seeking for a robust method for the synthesis of sialo-glycolipids such as gangliosides. In the earlier part of this paper, I will describe the synthesis of highly complex gangliosides mainly found in echinoderms, which possess neuritegenic activity. The syntheses were successfully achieved based on the

highly reactive synthetic units that were developed in our group; *N*-Troc-sialyl donor, 1,5-lactamized sialyl acceptor and glucosyl ceramide [1-3]. The synthetic units also allowed for the synthesis of fluorescently labeled gangliosides useful for single molecular imaging of lipid raft domain in the cell membrane [4-7]. In the later part, I will report a robust and comprehensive method for selective α-glycosidation of sialic acid, in which bicyclic sialic acid was utilized as the synthetic equivalent of the bridgehead oxocarbenium cation to direct α-glycoside formation [8]. This method ensured fully α-selective sialidation with a broad spectrum of hydroxyl compounds in high yields. Furthermore, the C-glycoside formation of sialic acid and the synthesis of oligomeric sialic acid have been achieved.

1. H. Tamai et al. Angew. Chem. Int. Ed. 2011, 50, 2330-2333.
2. K. Goto et al. Chem. Eur. J. 2016, 22, 8323-8331.
3. K. Goto et al. Org. Lett. 2019 in press. https://doi.org/10.1021/acs.orglett.9b01229.
4. N. Komura et al. Nat. Chem. Biol. 2016, 12, 402-410.
5. N. Komura et al. Trends Carbohydr. Res. 2017, 9, 1-26.
6. N. Komura et al. Methods Enzymol. 2017, 597, 239-263.
7. M. Koikeda et al. J. Carbohydr. Chem. 2019 in press. DOI: 10.1080/07328303.2019.1609019
8. N. Komura et al. Science, 2019, 364, 677-680.

630

Proteomics mass spectrometry analysis reveals subclass-specific molecular aberrations within triple-negative breast cancer

Max KOSOK[1, 2], Asfa ALLI SHAIK[1], Jayantha Gunaratne[1, 2], Boon Huat BAY[2]

*1. Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR), Singapore*

2. Department of Anatomy, Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore

Triple-negative breast cancer (TNBC) is considered as the most aggressive cancer subtype of the breast, and its intrinsic heterogeneity impedes elucidation of dysregulated molecular pathways and subsequently obstructs the development of targeted therapies. In this study we performed proteomics mass spectrometry of human breast cancer cell lines, resulting in the identification of two TNBC subclasses, basal A and basal B, with distinct proteomic profiles and associated molecular mechanisms. Within this subclasses, we found especially kinases and proteases to display unique expression pattern, indicating prominent functional roles in orchestrating discrete molecular functions. Focusing on those two protein classes, we used a comprehensive and advanced computational workflow to further interrogate subclass specific protein interactions and their associated perturbations, revealing deregulated pathways and possible targets for each TNBC subclass. While we found cell cycle, apoptosis and keratinization to be responsible for tumorigenesis in basal A subclass via NF-kB and WNT signaling. We identified kinase AXL and TGFBR1/2, and proteases FAP and MMP2/14, which play a central role in invasion and metastasis, to regulate EMT and TGF signaling in the generally more aggressive basal B subclass.

These observations highlight a distinct involvement of proteases and kinases in a subclass-specific regulation of tumorigenesis within TNBC. Thus, it has to be emphasized, that such proteins have to be exploited for targeted therapy in a subclass-specific manner, rather than for TNBC in general.

631

Multiplexed quantification of plasma proteins using isotopically labelled protein standards

David Kotol[1]

1. KTH - Royal Institute of Technology, Solna, STOCKHOLM, Sweden

The human blood plasma proteome contains an immense amount of information for determining a patient's phenotype on the molecular level. The changes in protein levels as an evidence for a trait have to meet strict requirements regarding accuracy and reproducibility to become a clinical assay fulfilling personalized and precision medicine endeavors. Here, we present an assay repertoire of 474 isotopically labelled protein standards mapping towards 351 proteins that can be used for multiplexed absolute quantification of proteins by mass spectrometry (MS) operated in data independent acquisition (DIA) mode. Each of these standards is overlapping each target protein sequence by shorter 50-150 amino acid stretches and they are added directly to the plasma sample prior to trypsin digestion. This enables robust quantification while each protein standard spike-in is individually tuned to its endogenous level, thereby allowing for precise one-point calibration.

632

Evaluation of a tims-Q-TOF instrument for targeted proteomics

Antoine Lesur[1], Pierre-Olivier Schmit[2], Joseph Longworth[1], Francois Bernardin[1], Gary Kruppa[3], Gunnar Dittmar[1]

1. Luxembourg Institute of Health, Strassen, Luxembourg

2. Bruker France, Wissembourg, France

3. Bruker Daltonics Inc, Billerica, MA, USA

Background

Targeted proteomics approaches are now commonly used, either to perform targeted biomarker candidate discovery or to validate candidate biomarkers that had been pointed out by untargeted discovery proteomics approaches. In this study, we are performing a first evaluation of the potential of the recently introduced nano-ESI tims Q-TOF architecture for targeted proteomics approaches. More specifically, we want to check if the time and space concentration of ions that results of the Trapped Ion Mobility

Spectrometry (tims) separation process can translate in to increased sensitivity and specificity performances for targeted proteomics approaches.

Methods

An equimolar mixture of 259 quantified synthetic peptides labelled with stable isotopes (AQUA) was diluted in a 100ng/µl human cell line digest. The dilution series covered 6 concentration levels ranging from 31.25 amol/µl to 25 fmol/µl. All samples were separated by nano-HPLC with a 60 min gradient and analysed on a high-resolution timsTOF Pro instrument (Bruker Daltonics) operated in data dependent PASEF mode or an exploratory targeted TIMS-PRM acquisition modes. The global sensitivity, selectivity and detectability of the different acquisition modes was evaluated with the latest version of the Skyline Ô software.

Findings

The PASEF acquisition of cell lysate digest spiked with 25fmol, 6,25fmol, 1562.5amol, 500amol, 125amol and 31.25 amol of the AQUA mixture allowed to identify 253/235/205/104/48 and 4 of the original 259 AQUA peptides, respectively. Using an exploratory tims-PRM approach with a 100 ms tims trapping time, and prior to any collision energy optimization, 110/168/205/213 and all of the AQUA peptides could be quantified at a the 31/125/500/1562/6250 amol level, respectively. Increasing the tims trapping time allowed to increase the detected S/N ratio. The results obtained after a more complete optimization will also be presented.

Conclusions

We have demonstrated a real (yet) unexploited potential of the tims-Q-TOF architecture for targeted proteomics approaches.

633

Physiology driven diversity in the peptidome of cow urine: A mass spectrometric view

ROHIT KUMAR[1], SUMIT KUMAR SINGH[1], Syed Azmal Ali[1], Sudarshan Kumar[1], J.K. Kaushik[1], A.K. Mohanty[1]

1. NATIONAL DAIRY RESEARCH INSTITUTE, KARNAL, HARYANA, India

Cow urine is an excretory fluid of biological origin that reflects summary of events occurring in body as a result of changing physiology or pathological conditions. Historically bovine urine is in use for various applications. However, very limited work is available on scientific evaluation of its effectiveness. To explore the active principles behind the activity of cow urine we focused our investigation on endogenous peptides derived from cow urine from three important physiological states of Sahiwal cows *viz.* heifer, lactation and pregnancy. These peptides possess several bioactive properties such as antimicrobial, antiangiogenic, antioxidant antithrombotic etc. Endogenous peptides are small size peptides (<10 kDa) which were obtained by using ultra-filtration assembly with a 10 kDa molecular weight cut-off filter. The filtrate containing peptides were used in solid phase extraction (SPE) for separation of peptides. Our preliminary investigation showed that the extracted peptides are endowed with antimicrobial activity as confirmed by disk diffusion assay (6 mm discs coated with urinary peptides) on *S. aureus*. MS/MS analysis of the crude peptides extracted from SPE further revealed 6238, 5465 and 3773 peptides from heifer, lactation and pregnant respectively with 10 individual samples for each group. Amino acids composition profile of the peptide sequences from all the groups follows the same pattern with the alanine, glycine, leucine, proline and serine being the most abundant amino acids. Intrestingly, most of the sequences were derived from collagen protein. Nevertheless, the types of sequences are mostly unique to individual physiological states. Using several databases (CAMPR3, Antiinflam TumorHPD etc.) we predicted potential functions of peptides such as anti-hypertension, anticancer and antimicrobial activity. Our study suggests that the diverse roles of cow urine in various applications are possibly due to the wide diversity of peptides present in cow urine.

634

Centrosome guides spatial activation of Rac to control cell polarization and directed cell migration

Hung-Wei Cheng[1], Cheng-Te Hsiao[2], Yin-Quan Chen[3], Chi-Ming Huang[1], Seng-I Chan[1], Arthur Chiou[4], Jean-Cheng Kuo[3, 1]

1. Institute of Biochemistry and Molecular Biology, National Yang-Ming University, Taipei, Taiwan

2. Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

3. Cancer Progression Research Center, National Yang-Ming University, Taipei, Taiwan

4. Institute of Biophotonics, National Yang-Ming University, Taipei, Taiwan

Directed cell migration requires centrosome-mediated cell polarization and dynamical control of focal adhesions (FAs). To examine how FAs cooperate with centrosomes for directed cell migration, we used centrosome-deficient cells and found that loss of centrosomes enhanced the formation of acentrosomal microtubules, which failed to form polarized structures in wound-edge cells. In acentrosomal cells, we detected higher levels of Rac1-GEF TRIO on microtubules and FAs. Acentrosomal microtubules deliver TRIO to FAs for Rac1 regulation. Indeed, centrosome disruption induced excessive Rac1 activation around the cell periphery via TRIO, causing rapid FA turnover, a disorganized actin meshwork, randomly protruding lamellipodia as well as loss of cell polarity. This study reveals the importance of centrosomes to balance the assembly of centrosomal and acentrosomal microtubules and to deliver microtubule-associated TRIO proteins to FAs at the cell front for proper spatial activation of Rac1, FA turnover, lamillipodial protrusion and cell polarization, thereby allowing for directed cell migration.

635

Novel functional proteins coded by the human genome discovered in metastases of melanoma patients

Magdalena Kuras[1], Aniel Sanchez[1], Jimmy Rodriguez[2], Indira Pla[1], Krzysztof Pawlowski[1], Johan Malm[1], Melinda Rezeli[2], Roger Appelqvist[2], Lazaro Betancourt[2], György Marko-Varga[2]

1. Department of Translational Medicine, Division of Clinical Protein Science and Imaging, Lund, Sweden

2. Department of Biomedical Engineering, Division of Clinical Protein Science and Imaging, Lund, Sweden

Background: Due to the high complexity and wide dynamic range of the human proteome, several predicted protein products have not yet been identified in proteomic experiments. Proteins that have been predicted but are currently unverified are classified as 'missing proteins' (MP). These proteins could provide essential clues to the interpretation of biological processes, and offer new research and therapeutic strategies to address the remaining clinical challenges.

Method: Proteins were extracted and digested from fresh frozen melanoma metastases samples [1]. Peptides were TMT-labelled, fractionated and analyzed by nLC-MS/MS on a Q Exactive HF-X mass spectrometer. Raw data were processed with Proteome Discoverer v2.3. Protein evidence (PE) was determined using the criteria from neXtProt and the C_HPP [2]. The bioinformatics analysis of relational networks between proteins that correlated with novel PE5 proteins was performed by ingenuity pathway analysis. Protein family annotation (PFAM) of the PE2 proteins was detected using DAVID.

Results: In this study, nine MPs were confidently identified by mass spectrometry in melanoma metastases. Some MPs significantly correlated with proteins that possess identical PFAM structural domains and other MPs were significantly-associated with cancer-related proteins. All proteins were classified based on the PE categories reported by neXtProt and annotated according to the HUPO guidelines [3]. Four proteins were uniquely identified within the context of metastatic cancer progression, although previously supported only by transcript presence, and five proteins previously marked as proteins of uncertain evidence and suspected to be pseudogenes were explicitly linked to mechanisms of melanoma metastasis. In addition, 24 other missing proteins were identified in up to 140 melanoma metastases.

Conclusion: Deep proteomic analyses were performed on more than 100 tissue samples from malignant melanoma patients. To our knowledge, this is the first study where unknown and novel proteins have been annotated in metastatic melanoma tumor tissue.

1. 1. Kuras, M.; Betancourt, L. H.; Rezeli, M.; Rodriguez, J.; Szasz, M.; Zhou, Q.; Andersson, R.; Marko-Varga, G., Assessing Automated Sample Preparation Technologies for High-throughput Proteomics of Frozen Well Characterized Tissues from Swedish Biobanks. J Proteome Res 2018.
2. 2. Omenn, G. S.; Lane, L.; Overall, C. M.; Corrales, F. J.; Liu, S.; Snyder, M.; Baker, M. S.; Deutsch, E. W., Progress on Identifying and Characterizing the Human Proteome: 2018 Metrics from the HUPO Human Proteome Project. J Proteome Res 2018, 17 (12), 4031-4041.
3. 3. Omenn, G. S.; Lane, L.; Lundberg, E. K.; Overall, C. M.; Deutsch, E. W., Progress on the HUPO Draft Human Proteome: 2017 Metrics of the Human Proteome Project. J Proteome Res 2017, 16 (12), 4281-4287.

636

Glycoproteomic changes in response to prion infection

Dominic KURIAN[1]

1. Roslin Institute, University of Edinburgh, Edinburgh, MIDLOTHIAN, United Kingdom

Chronic neurodegenerative diseases, such as Alzheimer's, Parkinson's and Prion diseases are all characterized by the protein conformational changes that result in misfolding and aggregation of one of the host proteins. Among these, Prion diseases are distinct from other neurodegenerative diseases, in that the misfolded protein (PrP^{Sc}) is infectious. The PrP^{Sc} is a misfolded and aggregated beta-sheet-rich isoform of the normal cellular prion protein (PrP). PrP is a glycoprotein with two N-linked glycans and this post-translational modification has critical roles in prion protein expression, distribution on neuronal cells and disease pathogenesis. Recent evidence in prion pathogenesis suggests that apart from the glycosylation of prion protein, the entire cellular glycosylation machinery is modified during pathogenesis. Herein I report the comparative analysis of prion-infected and un-infected mouse brain using integrated global proteomics and glycoproteomics approaches. Enrichment of glycoproteins using hydrazide beads, both at the protein level and tryptic peptide levels was performed to maximize glycoprotein/peptide coverage. N-glycopeptides were released from hydrazide support by PNGase-F removal of glycans and the resulting peptides were then labeled with stable isotopic dimethyl labeling and are analyzed by quantitative LC-MS analysis. In order to improve glycoprotein coverage, I have also applied alternate methods, such as HILIC and ZIC-HILIC for enrichment. Changes at global proteome and N-glycosite levels will be studied in prion-infected and normal brains, so that the prion disease associated alterations in glycosylation site occupancy can be measured.

637

Identification of isotope clusters from mass spectra using neural network model

Daewook Kwon[1], Seunghyuk Choi[1], Eunok Paek[1]

1. Hanyang University, Seoul, Seoul, SEOUL, South Korea

Mass spectrometry-based proteomics plays an important role in identifying peptides. Peptide identification strongly depends on a precursor mass estimated from the preceding precursor scan. However, we often observe that estimated precursor masses include isotopic errors. Determining isotope clusters is the first step in determining correct precursor masses. Existing tools such as RAPID (1) and MS-Deconv (2) adopted heuristic functions to recognize correct isotope clusters. Such heuristic functions have

been developed based on the characteristics of theoretical isotope clusters, but in isotope clusters of an experimental scan include noise and may overlap with different isotope clusters. Here, we propose a machine learning approach to identify correct isotope clusters, and it has a benefit of accommodating characteristics of experimental isotope clusters.

We designed an artificial neural network model to train characteristics of isotope clusters. The model takes a monoisotopic mass and intensities of the first to the twelfth peaks in a cluster as input, and predicts whether the given cluster is an isotope cluster or not.

To train the model, we collected ~4.2M peptide spectrum matches (PSMs) from a previous study (3). Detected isotope clusters (DICs) corresponding to the precursor of each PSM were extracted using both RAPID and MS-Deconv, and we filtered out ~2.95M DICs, whose spectral contrast angle (4) against theoretical isotope clusters (5) is below 0.80. We generated ~1.25M negative isotope clusters (NICs) consisting of partial peaks of selected ~1.25M DICs.

We applied 5-fold cross validation to prevent overfitting. The accuracy was 96.68% on average. We used DICs and NICs derived from different experimental methods (6,7) to test the model. The sensitivity and specificity were 97.35% and 85.85%, respectively.

1. K. S. Park, J. Y. Yoon, S. H. Lee, E. Paek, H. J. Park, H. J. Jung, S. W. Lee, “Isotopic peak intensity ratio based algorithm for determination of isotopic clusters and monoisotopic masses of polypeptides from high-resolution mass spectrometric data” Anal. Chem., vol.80, 7294-7303, 2008.
2. Liu X, Inbar Y, Dorrestein PC, et al. “Deconvolution and database search of complex tandem mass spectra of intact proteins: a combinatorial approach”, Mol. Cell. Proteomics., vol.9(12),27782-2782, 2010.
3. Zolg DP, Whilhelm M, Schnatbaum K, Zerweck J, Knaute T, Delanghe B, Kuster B, “Building ProteomeTools based on a complete synthetic human proteome”, Nat. Methods., vol.14(3), 259-262, 2017.
4. Umut H. Toprak, Ludovic C. Gillet, Alessio Maiolica, Pedro Navarro, Alexander Leitner, Ruedi Aebersold, “Conserved Peptide Fragmentation as a Benchmarking Tool for Mass Spectrometers and a Discriminating Feature for Targeted Proteomics”, Mol Cell Proteomics., vol.13(8), 2056-2071, 2014.
5. Martin Loos, Christian Gerber, Francesco Corona, Juliane Hollender, and Heinz Singer, “Accelerated Isotope Fine Structure Calculation Using Pruned Transition Trees”, Anal. Chem., vol.87(11), 5738-5744, 2015.
6. Philip Thomas, Trevor G. Smart, “HEK293 cell line: A vehicle for the expression of recombinant proteins”, Journal of Pharmacological and Toxicological Methods, vol.51(3), 187-200, 2005.
7. Geiger T, Wehner A, Schaab C, Cox J, Mann M, “Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins”, Mol Cell Proteomics., 11:M111-014050, 2012.

638

The breast cancer extracellular vesicle proteome recapitulates molecular features of the cell of origin

Rebecca E Lane[1], Darren Korbie[1], Ahmed Mohamed[2], Kum Kum Khanna[2], Michelle Hill[3, 2], Matt Trau[1, 4]

1. Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, St Lucia, QLD, Australia

2. QIMR Berghofer Medical Research Institute, Herston, QLD, Australia

3. The University of Queensland Diamantina Institute, Translational Research Institute, Woolloongabba, QLD, Australia

4. School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD, Australia

Introduction: Extracellular vesicles (EVs) have garnered much recent attention as circulating biomarkers for non-invasive tumour monitoring [1]. EVs are membranous nanoparticles which are secreted from a variety of cell types, including tumour cells. The clinical utility of EVs as circulating biomarkers is contingent on a thorough understanding of the correlation between tumour-EV cargo and the molecular features of the originating tumour cell. As such, the current study aimed to characterise the breast cancer EV proteome and evaluate if cellular features including ER and HER2 expression were reflected in the EV content.

Methods: EVs were isolated from the conditioned media of nine breast cancer cell lines using an optimised size exclusion chromatography protocol[2]. The isolated EVs were enumerated and physically sized using nanoparticle tracking analysis (NTA). Following this, label-free shotgun proteomics was performed using a nanoLC-MS/MS workflow, and proteins that were differentially abundant based on the ER and HER2 status of the originating cell were identified.

Results: NTA confirmed that nano-sized vesicles were recovered from all cell lines. The proteomics data suggested breast cancer cell-derived EVs shared core protein cargo, including established EV markers such as clathrin heavy chain 1, programmed cell death 6-interacting protein (ALIX) and CD9 antigen. Upregulated proteins in ER+ EVs included E-cadherin (CDH1), prominin-2 (PROM2) and mucin-5B (MUC5B). Unsurprisingly, HER2 protein was upregulated in HER2+ EVs, along with sushi domain-containing protein 2 (SUSD2) and mucin-6 (MUC6). A signature of 12 EV proteins was derived that appeared to recapitulate the ER and HER2 expression status of the originating cell.

Conclusion: This study represents an important foundational step in understanding the correlation between breast cancer cell and EV content. With further validation, the identified protein candidates may have utility as part of a liquid biopsy for characterising key molecular features of breast cancer.

1. 1. Lane, R. E., Korbie, D., Hill, M. M. & Trau, M. Extracellular vesicles as circulating cancer biomarkers: opportunities and challenges. Clinical and translational medicine 7, 14, doi:10.1186/s40169-018-0192-7 (2018).
2. 2. Lane, R. E., Korbie, D., Trau, M. & Hill, M. M. Optimizing Size Exclusion Chromatography for Extracellular Vesicle Enrichment and Proteomic Analysis from Clinically Relevant Samples. Proteomics, e1800156, doi:10.1002/pmic.201800156 (2019).

639

RIAna.py facilitates analysis of stable isotope labeling mass spectrometry experiments for protein turnover quantification

Edward Lau[1], Dean E Hammond[2], Robert J Beynon[3]

1. Stanford Cardiovascular Institute, Stanford University, Palo Alto, CA, USA

2. Cellular and Molecular Physiology, University of Liverpool, Liverpool, United Kingdom

3. Biochemistry, University of Liverpool, Liverpool, United Kingdom

The turnover rates of proteins are an important proteome parameter that is specific to organisms, tissues, and cellular states, and show promise as a new source of disease markers. Measurement of protein turnover rates may be achieved in intact animal models by stable isotope labeling using various isotope labels including heavy lysine, valine, or water, followed by mass spectrometry analysis of multiple time points. To enable large-scale data analysis, there is an unmet need for software tools which can automate the extraction of relative isotope abundance information from mass spectra and which is broadly applicable to multiple labeling strategies.

We describe RIAna.py (Relative Isotope Abundance Analyzer), an open-source software tool in Python designed to facilitate analysis for multiple isotope labels under a unified workflow. RIAna accepts as inputs mzid or tab-delimited files from database search results, and the corresponding mzML raw data through the Pymzml package. RIAna extracts the intensities of centroided MS1 peaks for each isotopomer for each qualifying peptide within user-defined retention time and mass precision ranges. It then returns the integrated areas-under-curve of the isotopomer chromatograms using the trapezoid method in SciPy. We applied RIAna toward a large dataset on inbred C57BL/6 mice labeled with [2]H2O or [13]C6-lysine (see companion poster by RB and DH). High-resolution mass spectrometry data were acquired on an Orbitrap instrument then searched against UniProt. Peptide mass isotopomers were then integrated over 0.5 min retention time windows at isotopomers 0,1,2,3,4,5 for [2]H2O labeling, and 0,6,12 for [13]C6-lysine labeling. Kinetic curve-fitting was performed using a custom R script with quasi-Newtonian optimization. Our analysis shows that RIAna provides isotopomer intensities and kinetic rate constants comparable to manual analysis but in a fraction of time required, and is applicable to different labeling methods. RIAna is freely available under MIT license on our GitHub repository.

64

Site-specific conjugation of dendrimer probes to the Fc glycans of monoclonal antibodies

Nima Sayyadi[1, 2], Edward Moh[1, 3], Douglas Campbell [4], Brad Walsh[4], Nicolle Packer[1, 2]

1. Macquarie University, Macquarie Park, NSW, Australia

2. ARC Centre of Excellence for Nanoscale Biophotonics , Macquarie University, Sydney, NSW, Australia

3. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

4. Minomic International Ltd, Sydney, NSW, Australia

Publish consent withheld

640

Involvement of CXCL5 in thermogenic function of inguinal white adipose tissue

Dabin Lee[1], Dong Wook Kim[1], Sanghyuk Yoon[1], Je-Yoel Cho[1]

1. Department of Biochemistry, BK21 PLUS Program for Creative Veterinary Science Research and Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, South Korea

Publish consent withheld

641

Automation of in-solution digestion for large-scale clinical proteomics through Selected Reaction Monitoring Mass Spectrometry (SRM-MS) in a cost-effective way

Jihyeon Lee[2, 1], Hyunsoo Kim[2, 3], Areum Sohn[2, 1], Injoon Yeo[2, 4], Youngsoo Kim[2, 1, 3, 4]

1. Department of Biomedical Sciences, Seoul National University , Seoul

2. Seoul National University, Jongro-gu, SEOUL, South Korea

3. Institute of Medical and Biological Engineering, MRC, Seoul National University, Seoul

4. Interdisciplinary Program of Bioengineering, Seoul National University , Seoul

Reproducibly quantifying biomarkers in large-scale cohorts remains a challenge in clinical proteomics due to multi-step process for sample preparation which consumes a great deal of time and human labor. Furthermore, many technical variations introduced from the multi-step process should be minimized for reproducible results of clinical assays. The rising demand for making reproducible results of research has led researchers to introduce the robotic liquid handling platform to process of preparation for the protein quantification. However, high cost of consumables for the automatic platform is the biggest obstacle for automation, although it considerably relieves the burden of large scale study. In this study we assessed the reproducibility of automated in-solution digestion while reducing the usage of consumables for automation. The quantification of 26 multiplex assays by Selected reaction monitoring-mass spectrometry (SRM-MS) was conducted in four sets of 24 pooled human serum aliquots. The fixed

number of pipette tips (same or fewer than the number of samples) are used to dispense each reagent in each set during digestion process. Reduction of related consumables such as wasted reagents, and reagent stock plates were accompanied by the reduced number of tips for each step. The reproducibility was evaluated by comparing the Coefficient of variations (CVs) values of MS-based quantification data from each set. As a results, the comparable reproducibility could be maintained, while the cost of consumables reduced up to one-sixth of the standard experiment (24 tips used for 24 samples). The reduced cost for automation will enable researchers to facilitate automated workflow for their large-scale clinical research with less cost burden.

642

Discrimination of cancer cell types using ganglioside fingerprinting based on LC-MS/MS

Jua Lee[1], Heeyoun Hwang[2], Jaekyung Yun[1], Sun-il Choi[3], Yun-Hee Kim[3], Yong Sam Kim[4], Hyun Joo An[1]

1. Chungnam National University, Yuseong-gu, Daejeon, South Korea

2. Biomedical Omics Research, Korea Basic Science Institute, Cheongju, South korea

3. Research Institute and hospital, National Cancer Center, Goyang, South korea

4. Cancer Biomarkers Development Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South korea

Gangliosides, a subtype of glycosphingolipids containing sialic acid, play important roles in various biological processes such as cell-cell recognition and interaction, and modulation of membrane protein function. In particular, it has attracted considerable attention for a long time due to their close association with cancer pathology such as modified expression patterns in cancer cells related to apoptosis resistance. Although the important biological functions of the gangliosides continue to be revealed, there is only few study on comprehensive profiling of gangliosides due to the lack of effective analytical method for proving their structural complexity (hydrophobic ceramide moiety and hydrophilic glycan chain). In this study, we performed profiling and structural elucidation of gangliosides in the different cancer cell lines such as CFPAC1, A549, NCI-H358, MCF7, and Caski using negative ion detection mode nano LC/MS and MS/MS. Extracted gangliosides were assigned based on accurate mass, known glycobiology, and established mathematical model, and then the structures were further elucidated using tandem MS, retention time. In total, over 70 ganglioside compositions were identified from 5 cancer cell lines. Notably, the different cell lines were obviously distinguished through the difference of relative abundance of various gangliosides. In particular, the distribution of glycan moiety was more distinctly differed in each cell line than the ceramide moiety between different cell lines. These results demonstrated that cancer origins and molecular subtypes could be definitely differentiated based on distinctive ganglioside profiles. This result shows the feasibility of cell surface ganglioside as effective targets for diagnosis of cancer cell types.

643

Proteogenomic Characterization of Human Early-Onset Gastric Cancer

Sang-Won Lee[1], Nam Dowoon[1], Dong-Gi Mun[1], Ji-Won Hong[1]

1. Korea University, Seoul, NA, South Korea

We report an integrated proteogenomic analysis of early onset gastric cancers (EOGCs) using exome and mRNA sequencing and global proteome, phosphoproteome, and glycoproteome profiling. Single amino acid variants (SAAVs) identified by exome sequencing and supported by proteomic and mRNA transcriptomic data enabled effective prioritization of candidate driver genes that are functional at the protein level. Protein abundance changes between paired tumor and adjacent normal tissues were poorly predicted by mRNA transcript abundances, indicating that protein measurements increased the reliable evaluation of cancer-related alterations of cellular processes. Integrated proteogenomic analysis identified four subtypes of EOGCs. Four subtypes were associated with proliferation, immune response, metabolism, and invasion, respectively; and associations of the subtypes with immune- and invasion-related pathways were identified mainly by phosphorylation and N-glycosylation data. Therefore, integrated proteogenomic analysis provided a more complete and better refined molecular characterization of EOGCs than that by genomic analysis alone, affording a paradigm for enhanced understanding of cancer biology and a roadmap for patient stratification as it relates to this disease.

644

Harnessing machine-learning techniques to accurately identify protein complexes and their changes based on SEC-SWATH/DIA data

Chen Li[2, 1], Andrea Fossati[1, 3], Moritz Heusel[4], Fabian Frommelt[1], Federico Uliana[1], Isabell Bludau[1], Varun Sharma[1], Matteo Manica[5], Peng Xue[1, 6], María Rodríguez Martínez[5], Patrick G.A. Pedrioli[1], Anthony W. Purcell[2], Matthias Gstaiger[1], Rudolf Aebersold[7, 1]

1. Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, Zürich, Switzerland

2. Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria, Australia

3. Department of Biology, Institute of Molecular Health Sciences, ETH Zürich, Zürich, Switzerland

4. Division of Infection Medicine (BMC), Department of Clinical Sciences, Lund University, Lund, Sweden

5. IBM Research Zürich, Zürich, Switzerland

6. Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

7. Faculty of Science, University of Zürich, Zürich, Switzerland

Biological functions are usually performed and regulated by complexes of interacting proteins. In most cases, transcriptomic and proteomic measurements have focused on individual molecules and their functions. Despite the availability of a range of methods for the analysis of specific protein complexes, systematic analysis of the ensemble of protein complexes in a sample has remained

challenging, highlighting the urgent need to identify protein complexes and their functions using high-throughput proteomic techniques. Approaches based on biochemical fractionation of intact complexes and correlation of protein profiles have shown promising performance when combining size exclusion chromatography (SEC) with highly accurate protein quantification by SWATH-MS. However, interpretation of co-fractionation datasets to protein complex composition, abundance and dynamic rearrangements remains challenging due to the limited experimentally-verified protein complexes in public knowledgebases[1], because of which novel protein complexes and their changes are not identified without prior evidence. Here, we used advanced machine-learning techniques to construct a systematic framework based on Random Forest (RF) to identify novel protein complexes from SEC-SWATH-MS data and to characterize their changes across different experimental/medical conditions. With raw protein matrices of different conditions, our framework can accurately identify novel protein complexes and their changes across conditions, using hypothesis generation, data rescaling and differential analysis modules in the software package. Experiments demonstrated that the RF model achieved outstanding performance with an average AUC of 0.948, accuracy of 88.3% and MCC of 0.762 via five-fold cross-validation on three in-house/published SEC-SWATH-MS datasets (mapped to the CORUM database). Independent tests across different data acquisition methods (DDA *vs.* DIA) and species (human *vs.* mouse) also highlighted the strong generalizability of the RF model and the robustness of the prediction performance. The software is being implemented and will be freely available for prediction and analysis of novel protein complexes and their changes across conditions from co-fractionation-MS datasets.

1. Moritz Heusel, Isabell Bludau, George Rosenberger, Robin Hafen, Max Frank, Amir Banaei-Esfahani, Audrey van Drogen, Ben C. Collins, Matthias Gstaiger, and Ruedi Aebersold. Complex-centric proteome profiling by SEC-SWATH-MS. Molecular Systems Biology (2019) 15, e8438

645

Temporal profiling of the diabetic and non-diabetic heart during myocardial reperfusion injury

Desmond K Li[1, 2], Alexander W Rookyard[3, 2], Stuart J Cordwell[3, 1, 2, 4], Melanie Y White[3, 1, 2]

1. Discipline of Pathology, School of Medical Sciences, The University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

4. Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia

Cardiovascular disease is the leading causes of death worldwide with ischemic heart disease (IHD) the largest contributor. As an imbalance in oxygen supply and demand, ischemic damage can be minimised by timely reperfusion, however reperfusion itself can also result in further damage. Individuals with diabetes are at a greater risk of developing IHD with higher mortality rate and poorer outcomes following an ischemic event. The current study utilises a rat model of type 2 diabetes (T2D) to investigate how the heart recovers after an ischemic event by profiling the changes of protein phosphorylation and metabolite abundance in physiological and T2D conditions. Rat hearts were subjected to Langendorff perfusion to produce increasing periods of reperfusion, up to 60 min (60R) after 15 min of global ischemia (15I). Phosphoproteomic analysis was conducted using isobaric tags across the reperfusion time course, prior to mass spectrometry (MS) while metabolomics was performed using a targeted MS approach. Significant changes in abundance were identified in >8000 phosphopeptides and 112 metabolites across the time course, with a majority significantly altered within the first 5 minutes reperfusion (>75% for phosphopeptide, >65% for metabolites). Differential analysis between the diabetic and non-diabetic hearts revealed altered regulation of a number of kinases including CaMK2, FAK, PI3K, MAP2K2 and associated pathways, which all have a role in healthy heart function. This was confirmed using metabolomics which showed that high energy phosphates were being differentially utilised. By elucidating the differences in signalling pathways and kinase modulators between diabetic and non-diabetic hearts during reperfusion, novel therapeutic targets are possible to improve cardiovascular outcomes for diabetics after an ischemic event.

646

The effect of post-translational modifications on the molecular phenotype of the liver in type 2 diabetes

Desmond K Li[1, 2], Alexander W Rookyard[3, 2], Lauren E Smith[1, 2], Zeynep Sumer-Bayraktar[3, 2], Yen Chin Koay[2, 4], Holly McEwan[2, 5], Anthony Don[2, 5], John O'Sullivan[2, 4], Stuart J Cordwell[3, 1, 2, 6], Melanie Y White[3, 1, 2]

1. Discipline of Pathology, School of Medical Sciences, The University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

4. Heart Research Institute, Sydney, NSW, Australia

5. ACRF Centenary Cancer Research Centre, Sydney, NSW, Australia

6. Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia

Altered glucose metabolism in response to peripheral insulin resistance is a hallmark of type 2 diabetes (T2D), clinically observed as the inability to maintain postprandial blood glucose levels (BGL). Associated with energetic excess arising from caloric overload, T2D is linked to excess non-esterified fatty acid production and rising nutrient levels, which influence metabolic processes. The liver plays a pivotal role in the pathogenesis of T2D, as both an insulin sensitive organ and by gluconeogenesis, whereby glycogen stores are liberated, further elevating BGL. It is important to understand the molecular adaptations of the liver to the metabolic flux and insulin resistance arising from T2D. We developed a T2D rat model which complements energetic excess from elevated dietary fat intake and insulin insufficiency from pancreatic beta cell damage. To investigate cellular changes to T2D, we performed a proteomic analysis which included targeting post-translational modifications by phosphoproteomics, cysteine redox proteomics and glycomics. To quantify alterations in protein and PTM abundance, samples were isobarically

tagged prior to mass spectrometry (MS). Discovery glycomics was achieved by untargeted MS methods. We used untargeted lipidomic and targeted metabolomic analysis for functional validation. We quantified close to 20,000 phosphopeptides, 7,500 cysteine redox peptides, 7,000 proteins, and 18 N-glycan structures. PTM analysis showed that amino acid metabolic pathways were greatly altered by both phosphorylation and redox modifications. Using metabolomics, we observed a concurrent change in branched chain amino acids (BCAA) and metabolites facilitating amino acid metabolism. Given that our model alters only dietary fat intake, and hepatic availability to glucose via altered insulin levels, altered level of BCAA are of particular interest.

647

FORECAST: a novel computational strategy Fusing Omics datasets with liteRaturE information to aCcelerAte the progress of gene annotation

Dong Li[1]

1. Beijing Institute of Life Omics, Beijing, BEIJING, China

Manual annotation for the significant genes from the routine bioinformatics analyses of large-scale omics datasets has become the bottleneck for further biomedical knowledge discovery. To address this challenge, a novel computational strategy called FORECAST was developed to realize seamless fusion between omics datasets and literature information for automatic gene-centric knowledge mining. In the case of cancer gene discovery, FORECAST exhibits superior efficacy over routine strategies of either omics data analysis or literature mining. Identified novel cancer genes also proved the high capability of FORECAST to mine novel knowledge. To meet the requirements of large-scale omics datasets analysis, FORECAST shows great robustness and efficiency when implemented in high-volume, noisy and incomplete biological omics and literature datasets. This strategy is promising to accelerate the knowledge mining in the era of Big Data.

648

Suborgan proteome of honeybee workers reveals induced signal transduction in antennal lobe to drive distinct reproductive investment by alloparental care

Jianke Li[1], **Xufeng Zhang**[1]

1. Institute of Apicultural research, Chinese Academy of Agricultural Sciences, Haidian District, BEIJING, China

Raising female-destined larvae into new queens by alloparental care of nurse bees in provision of royal jelly (RJ) is critical for reproductive investment of honeybees (*Apis mellifera*) by alloparental care, which requires more than one brain region involving in neural responses. However, knowledge of how mushroom bodies (MBs) and antennal lobes (ALs) of brain influence the alloparental care of honeybees is largely unknown. We compared the proteome of MBs and ALs in the brain of two strains of honeybees with different reproductive investment in terms of high (royal jelly bees selected from Italian bees > 40 years, RJbs) and low (Italian bees, ITbs) RJ production. In newly emerged bees, the induced pathways associated with energetic metabolism in MBs and ALs of RJbs comparing to ITbs, indicate that the developmental trajectory of brain has reshaped since eclosion. In nurse ALs of RJbs comparing to ITbs, the pathways and kinases related to modulate vesicular transport, signal transduction, synaptic plasticity, and long-term memory consolidation are functionally enhanced to boost the activity of neuronal cell to sense the brood odor. The difference in ALs is also found in forager bees between both bee lines, of which pathway related to protein metabolism are functionally activated, however there was no difference in MBs between them. Among of all the differences, of which the enhanced activity of nurse ALs plays the major role in driving the stronger reproductive investment in RJbs than in ITbs. Our findings indicate that RJbs has adapted a strong olfactory sense to larvae by alloparental care due to the genetic selection of increasing royal jelly outputs. This gains a novel insight into a neural biology of honeybees, and is potentially useful for neurophysiology for honeybees and other social insects.

649

Proteomics and Metabolomics to Predict Outcome for Pre-Disease Individuals Using Edge Biomarkers

Qing-Run Li[1], **Yi-Di Sun**[1], **Wan-Wei Zhang**[1], **Jia-Rui Wu**[1], **Rong Zeng**[1]

1. CAS Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, CAS, SHANGHAI, CHINA

Early prediction of the long-term outcome at the pre-disease stage is often a big challenge since the divergence between pre-disease and healthy is trivial and fluctuant. We developed a novel method using correlation information as features to predict 6-year glycemic status in pre-diabetic individuals, based on data of proteome and metabolomics. We first collected plasma samples of the population that the pre-diabetic individuals at the baseline level. After 6-year follow-up, 45 of these subjects returned to normal fasting blood-glucose level, another 45 remained pre-diabetes and the other 45 developed into type 2 diabetes. Then, we profiled the proteins and carnitines of the blood samples which were collected 6 years ago by mass spectrometry.

Next, to integrate these data, the edge biomarker method which used the correlation of molecules as biomarker was employed. In detail, this method transforms the molecular expression data into the correlation components of each molecular pair, which involving feature selection, classifier training and phenotype prediction on the edge-level data, and finally to construct molecular pairs from proteomics and metabolomics data. High cross-validation accuracy and functional analysis of the selected edge biomarkers suggested its clinical potentials. Evidently, most of the molecules were associated with diabetes reported by previous work. We demonstrated that edge biomarkers of proteomics and metabolomics data could effectively predict long-term outcome in a pre-disease population.

65

Development of gold nanoparticle-based multivalent photoaffinity probes toward exploration of carbohydrate-protein interaction

Kaori Sakurai[1]

1. Tokyo University of Agriculture and Technology, Koganei-shi, TOKYO, Japan

Carbohydrate-protein interactions mediate cellular signals, which are crucial in a diverse array of biological and pathological processes. Despite their importance, many of carbohydrate-protein interactions remain unknown due to inherent difficulties in studying them. While chemical methods involving affinity probes and affinity matrixes offer most direct and rapid ways to explore cellular targets, their utility is often limited particularly in the case with carbohydrate binding proteins with low affinity. With the aim to provide the first step in elucidating the biological roles of carbohydrates, photoaffinity labeling has been employed as a promising chemical strategy for the detection and the identification of carbohydrate-binding proteins in their native environment. In this presentation, I will introduce a new approach, which employs multivalent photoaffinity probes based on gold-nanoparticles for streamlined analysis of carbohydrate-protein interactions. Gold nanoparticles are attractive scaffolds for chemical probes based on several advantages. The probe design can be optimized rapidly owing to facile and modular functionalization of gold nanoparticles with various composition of a ligand and a photoreactive group. Multivalent display of ligands and photoreactive groups promotes increased protein binding affinity as well as enhanced protein reactivity. Proteins can be easily enriched on gold nanoparticle probes by centrifugal separation from the protein mixture. Therefore, gold nanoparticles can serve both as a scaffold for affinity probes in solution and as an affinity matrix separable from solution. Since preparation of chemical probes and their application to target identification studies are often laborious, multistep processes, our approach should be useful for expediting the identification and analysis of target proteins.

650

Development of a sensitive and specific targeted mass spectrometry assay for proNGF

Xiang Li[1], **Sam Faulkner**[1], **Matthew D. Dun**[1], **Hubert Hondermarck**[1]

1. School of Biomedical Sciences and Pharmacy, Hunter Medical Research Institute, The University of Newcastle, Callaghan, NSW 2308, Australia

Background: The precursor for nerve growth factor (proNGF) is emerging as a biomarker and therapeutic target in cancer and neurodegenerative diseases. However, current quantification of proNGF in clinical samples, such as serum and tissues, is based on ELISA and immunohistochemistry and suffers from the limitations of these methods in terms of specificity and sensitivity.

Aims: To design a sensitive and specific parallel reaction monitoring mass spectrometry (PRM-MS) assay to detect and quantify proNGF in human cell extracts and serum samples.

Methods: PRM-MS was performed using a high-resolution Orbitrap mass spectrometer coupled with nanoflow-liquid chromatography (LC-MS/MS). Initially, tryptic peptide transitions were optimized using stable isotope labelled proNGF peptides in human sera. ProNGF peptides were detected and quantified compared to the abundance of the stable isotope control. Quantification obtained in PRM-MS vs ELISA were also compared.

Results: A proNGF specific peptide at 577.3198++ m/z presenting no similarities with other human proteins was selected. Quantification was based on the corresponding stable isotope control (observed in positive ion mode). Using this assay, proNGF was detected and quantified in cell extracts and human sera. The detection of proNGF was also validated by immunoprecipitation. Although both PRM-MS and ELISA could detect proNGF in the human serum, there was a limited concordance between quantification obtained in PRM-MS vs ELISA.

Conclusion: This study demonstrates that PRM-MS is an efficient technique for the detection and quantification of proNGF in human samples. These results have potential clinical significance in cancer and neurological diseases where proNGF is involved.

651

Global lipidomics reveals broad dysregulation of lipid metabolism in triple negative breast cancer development

Yan Li[1], **Fangmeng Fu**[1], **Ling Lin**[2], **Songping Lin**[1], **Jinxing Lv**[1], **Chuan Wang**[1]

1. Department of Breast Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian, China

2. Institutes of Biomedical Sciences of Shanghai Medical School, Fudan University, Shanghai, China

Triple Negative Breast Cancer (TNBC) is an aggressive disease, accounting for 15% -23.8% of all breast cancers. Due to the complexity and diversity of lipid molecules, along with the challenges in analytical techniques development and comprehensive lipid database construction, altered lipid composition and reprogrammed lipid metabolism have not been fully elucidated during TNBC progression. Here, we performed a highly sensitive shotgun lipidomic approach to explore the global lipidome in tumor tissues and corresponding para-tumor tissues from Grade I-II and Grade III TNBC patients. Totally, 1556 intact lipids were identified in different stage of TNBC tissues and paired para-cancerous tissues. Palmitic acyl (C16:0)-containing glycerophospholipids (GPs) were significantly reduced in tumor tissues compared with adjacent nontumor tissues, while C18:0 and C20:5-containing GPs were remarkably upregulated. Systematic lipidomic changes occurring in different groups were then assessed by two widely used multivariate methods - principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA). The first principal component (PC1) accounted for 69% of the total variance and separated TNBC tumor tissues from corresponding para-tumor tissues. The second principal component (PC2) accounts for 12.1% of the variance in the data. PLS-DA was applied to enhance the separation between the groups. PE(18:0/18:2) with the highest VIP value was the most powerful group discriminator in differentiating tumor tissues from para-tumor tissues. It is interesting to note that the decreased saturated fatty acyl containing GPs (such as C16:0 and C14:0) and the increased unsaturated fatty acyl comprising GPs (such

as C18:0, C20:4 and C20:5) in tumor tissues compared with para-tumor ones. Thus we further examined the effects of palmitic acid treatment on cell proliferation and invasion ability *in vitro* via MDA-MB-231 cell line. Our data collectively herald the biomedical potential of using altered lipid metabolism for treating triple negative breast cancer.

652

Precision *de novo* peptide sequencing using mirror proteases of Ac-LysargiNase and trypsin

Yanchang Li[1], Hao Yang[2], Mingzhi Zhao[1], Xi Wang[2], Feilin Wu[1], Junling Zhang[1], Simin He[2], Hao Chi[2], Ping Xu[1]

1. Beijing Proteome Research Center, Beijing, China

2. Institute of Computing Technology, CAS, Beijing, China

De novo sequencing is an alternative approach to identify unknown proteins, post-translational modifications (PTM) and amino acid mutations, which deduces peptide sequences directly from tandem mass spectra (MS2) instead of searching reference databases. *De novo* peptide sequencing for large-scale proteomics remains challenging because of the lack of full coverage of ion series in tandem mass spectra. We previously developed a mirror protease of trypsin, acetylated LysargiNase (Ac-LysargiNase), with superior activity and stability. The mirror spectrum pairs derived from the Ac-LysargiNase and trypsin treated samples can generate full *b* and *y* ion series. The *b* and *y* ion series provide mutual complementarity and examination of each other, which allowed us to develop a novel algorithm, pNovoM, for *de novo* sequencing. Using pNovoM to sequence peptides of purified proteins, the accuracy of the sequence was close to 100%. More importantly, from a large-scale yeast proteome samples digested with trypsin and Ac-LysargiNase individually, 48% of all tandem mass spectra formed mirror spectrum pairs, 97% of which contained full coverage of ion series, resulting in precise *de novo* sequencing of full-length peptides by pNovoM. This enabled pNovoM to successfully sequence 21,249 peptides from 3,753 proteins and interpreted 44–152% more spectra than pNovo+ and PEAKS at a 5% FDR at the spectrum level. Moreover, the mirror protease strategy had an obvious advantage in sequencing long peptides. We believe that the combination of mirror protease strategy and pNovoM will be an effective approach for precision *de novo* sequencing on both single proteins and proteome samples. Futhermore, we would like to additionally combine NeuCode labeling with mirror strategy to distinguish ion types for improvement of *de novo* sequencing in the furture.

1. Hao Yang, Yanchang Li, Mingzhi Zhao, et al. (2019) Precision De Novo Peptide Sequencing Using Mirror Proteases of Ac-LysargiNase and Trypsin for Large-scale Proteomics. Mol Cell Proteomics 18:773-785.

653

Proteomics links ubiquitin chain topology change to transcription factor activation

Yanchang Li[1], Mark A. Villamil[2], Qiuyan Lan[3], Eric B. Dammer[1], Yuan Gao[1], Junzhu Wu[3], Fuchu He[1], Peter Kaiser[2], Ping Xu[1]

1. Beijing Proteome Research Center, Beijing, China

2. University of California Irvine, Irvine, CA, USA

3. Wuhan University, Wuhan, Hubei, China

A surprising complexity of ubiquitin signaling has emerged with identification of different ubiquitin chain topologies. However, mechanisms of how the diverse ubiquitin codes control biological processes remain poorly understood. Here, we use a SILAC label-swap approach that combines high-resolution protein fractionation and an LC–MS/MS platform to examine the proteomes of two yeast strains, a wild-type (WT) strain and a ubiquitin mutant (K11R) strain, which could be used to identify yeast proteins that are regulated by lysine 11 (K11)-linked ubiquitin chains. Our results reveal a profound downregulation of methionine biosynthesis enzymes in the K11R mutant strain, indicating a so far unappreciated role of K11-linked ubiquitin chains in regulating the SCF^{Met30}-Met4 network. The entire Met4 pathway, which links cell proliferation with sulfur amino acid metabolism, was selected for mechanistic studies. Previously we demonstrated that a K48-linked ubiquitin chain represses the transcription factor Met4. Here we show that efficient Met4 activation requires a K11-linked topology. Mechanistically our results propose that the K48 chain binds to a topology-selective tandem ubiquitin binding region in Met4 and competes with binding of the basal transcription machinery to the same region. The change to K11 enriched chain architecture releases this competition and permits binding of the basal transcription complex to activate transcription.

1. Yanchang Li, Eric B. Dammer, Yuan Gao, et al. Proteomics links ubiquitin chain topology change to transcription factor activation. Mol Cell, 2019 (accepted).

654

UbiBrowser 2.0: expanded prediction for human proteome-wide deubiquitinase-substrate interactions

Yang Li[4, 2, 3, 1], Xun Wang[4, 2, 3, 1], Fuchu He[4, 2, 3, 1], Dong Li[4, 2, 3, 1]

1. State Key Laboratory of Proteomics, Beijing

2. Beijing Proteome Research Center, Beijing

3. Beijing Institute of Lifeomics, Beijing

4. National Center For Protein Science-Beijing, Changping, BEIJING, China

Publish consent withheld

655

The quest for the perfect tumor specific antigen - characterizing the oncogenic MHC ligandomes for novel therapeutic targets

Jennie R Lill[1]

1. Genentech, SOUTH SAN FRANCISCO, CALIFORNIA, United States

Background

Cancer Immunology is a rapidly evolving and exciting field with many new strategies for harnessing the immune system to eliminate tumors on the horizon. We have previously reported proteomic studies on the characterization and targeting of neo-antigens, leading to the development of a therapeutic approach for generating cancer vaccines. These vaccines are already showing promise in the clinic. Our current strategy expands on this, and explores novel mechanisms to find additional classes of tumor specific antigens for targeting using a combination of transcriptomics, proteomics and proteo-informatic tools.

Methodologies

We have employed a combination of ribosomal sequencing, MHC I proteomics and in depth proteo-informatics on a variety of primary cell lines, monoallelic engineered cells and primary tumors. We compared the effects of various cellular perturbations and small molecule inhibitors to decipher how to reveal new tumor specific antigens for potential application as biotherapeutics. Using high resolution mass spectrometry in combination with exome sequencing, ribosomal sequencing and computational strategies we reveal repertoires of peptides as targets for oncological intervention.

Findings

By performing various cellular perturbations and characterizing the MHC I ligandome using proteomics, we have identified several new tumor specific antigens which are currently being tested for their ability to elicit a potent immune response. If successful, these antigens will be added to our arsenal of cancer immunotherapeutic tools.

656

Increasing the ease of use of nanoflow with plug and play low flow source

Justin Lim[1], **Christie Hunter**[2], **Leroi De Souza**[3], **Carmai Seto**[3], **Yang Kang**[3], **Leigh Bedford**[3], **Bradley Schneider**[3]

1. SCIEX, Singapore

2. SCIEX, USA

3. SCIEX, Canada

When performing electrospray ionization mass spectrometry, reducing chromatographic flow rates can improve sampling efficiency and therefore increase sensitivity. When operating in the nanoflow regime (50-500 nL/min) very high sampling efficiencies are possible. However, nanoflow ionization can require significant user expertise and advanced tuning to get best performance. Significant research was done to determine whether most source parameters could be optimized and locked in, to take out most of the user interactions. Using Design of Experiments, x,y,z positioning, tip protrusion, ionization voltage, nebulization gas and other parameters were examined. Spray electrodes were developed for maximal robustness. This led to the development of the OptiFlow® Source, a single source that would cover the full spectrum of low flow rates for high sensitivity LC-MS analysis.

Performance was evaluated relative to a highly tuned NanoSpray® Source III on both a TripleTOF® 6600+ system and a QTRAP® 6500+ system. Equivalent chromatographic performance was observed comparing peak width and area. 30 day testing was performed to check spray electrode robustness and minimal change in backpressure was observed. The SWATH® Acquisition Performance kit was run with both sources and very equivalent performance was again observed, within 5% in IDA and SWATH acquisition modes. Next, plug and play cartridge based nanoflow columns will be tested with different phases.

657

Discovery and verification of phosphoprotein signature for predicting prognosis of patients with hepatocellular carcinoma

Yu-Tsun Lin[1], **Kun-Yi Chien**[1, 2, 3], **Jau-Song Yu**[1, 2, 3]

1. Department of Biochemistry & Molecular Biology, Chang Gung University, Taoyuan, Taiwan

2. Graduate Institute of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan

3. Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

Hepatocellular carcinoma (HCC) is one of the most common cancers worldwide. Along with a rising number of diagnostic biomarkers were discovered, more and more patients can be diagnosed early. Additionally, patients with early-stage HCC are able to be treated curatively. Unfortunately, 40-70% HCC patients still suffer from recurrence within five years. Accordingly, various reports tried to develop prognosis signature in HCC to find potential prognostic predictors. However, most of the previous studies focused on mRNA, miRNA, and proteins. Phosphorylation catalyzed different vital signaling cascades which may contribute to poor prognosis of HCC patients. Therefore, phosphorylation-related events have generally been considered biomarkers and/or viable therapeutic targets for HCC. In the present study, we aim to discover potential phosphorylation-related biomarkers for predicting prognosis and/or potential therapeutic targets for developing new treatment option for HCC patients.

First, we established a pre-fractionation workflow using strong anion-exchange chromatography under continuous pH gradient, which can efficiently increase the number of identified proteins and phosphopeptides. Next, we applied this pre-fraction workflow and iTRAQ-based labeling strategy to quantitatively analyze proteome and phosphoproteome of paired tissue samples (cancerous and adjacent non-cancerous tissues) from HCC patients with good (no recurrence in 3 years after operation) or poor prognosis (recurrence in one year after operation). In the four pooled tissue samples, we obtained relatively quantitative results of 20,693 phosphopeptides and 9,311 proteins. We found that phosphoprotein signature outperformed protein profile to distinguish poor prognosis from good prognosis group. Moreover, we introduced super-SILAC mix as internal standard into the pooled tissue samples from discovery phase to preliminarily verify dozens of prognosis-related proteins/phosphoproteins by PRM assay. We'll further validate the significant candidates in hundreds of HCC tissue specimens to determine practical prognostic predictors and investigate the feasibility of treatment targets.

658

Quantitative proteomics reveals stage-specific protein regulation of triple negative breast cancer

Yuxiang Lin[1], Ling Lin[2], Fangmeng Fu[1], Songping Lin[1], Juanjuan Xie[2], Pengyuan Yang[2], Chuan Wang[1], Huali Shen[2]

1. Department of Breast Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian, China

2. Institutes of Biomedical Sciences of Shanghai Medical School, Fudan University, Shanghai, China

Triple negative breast cancer (TNBC) is a heterogeneous disease with more aggressive clinical courses than other subtypes of breast cancer. To data, functional proteomic analysis provides complementary information that can be integrated with the genomic and transcriptomic data to explore the clinic-pathological characteristic differences as well as the accurate classification for breast cancer subtypes. We performed a highly sensitive proteomic approach of iTRAQ-labeling coupled LC-MS/MS to obtain the global proteome and unravel protein signatures in tumor tissues and corresponding para-tumor tissues from 24 patients with Grade I-II and Grade III primary TNBC. Totally, 5,401 unique proteins were identified and quantified in different stage of TNBC. 865 proteins were changed in patients with Grade I or II TNBC, among which 309 were up-regulated and 556 were down-regulated. Meanwhile, for patients with Grade III TNBC, 359 proteins were increased and 672 proteins were decreased. Differentially expressed proteins were further analyzed by bioinformatic analyses, including GO function classification annotation, ingenuity pathway analysis and KEGG enrichment analysis. Comparing to para-cancerous tissues, various signaling pathways and metabolic processes, including PPAR pathways, PI3K-Akt pathway, one-carbon metabolism, amino acid synthesis, and lipid metabolism were activated in TNBC cancer tissues. Interestingly, death receptor signaling was significantly activated in Grade I-II TNBC, however, remarkably inhibited in Grade III TNBC. Our proteomic data presented precise quantification of potential signatures, signaling pathways, regulatory networks and characteristic differences in each clinic-pathological subgroup. The proteome provides complementary information for TNBC accurate subtype classification and therapeutic targets research.

659

Analysis of the renal redox proteome in type 2 diabetes mellitus

Shivanjali J. Lingam[1], Desmond K. Li[1], Stuart J. Cordwell[2], Melanie Y. White[2]

1. Discipline of Pathology, School of Medical Sciences, Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

2. Discipline of Pathology, School of Medical Sciences, School of Life and Environmental Sciences, Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

Type 2 diabetes mellitus (T2DM) mediated renal dysfunction, or diabetic nephropathy (DN) is a progressive chronic complication which further increases the risk of T2DM mortality. Several studies have found reactive oxygen and nitrogen species (ROS/RNS) as key mediators of signal transduction during T2DM and DN leading to increased oxidative stress. Concurrent decreases in the capability of the cellular antioxidant defences to compensate for this T2DM-induced oxidative stress, results in elevated redox-mediated post-translational modifications (PTMs) of reactive cysteine residues. Here, we examined the cysteine redox-modified renal proteome in T2DM using rats that were fed either a standard CHOW (C) (12% fat) or high fat (HF) (42% fat) diet for 8 weeks with T2DM induced in 50% of the animals after 4 weeks utilising a low dose of streptozotocin (STZ; 35mg/kg); a pancreatic β-cell toxin. The remaining 50% were injected with a buffer vehicle (Cit). At the cessation of the feeding protocol, 9 animals per treatment group (C Cit; C STZ; HF Cit; HF STZ) were euthanised and kidneys excised. Tissue lysates were subjected to western blotting using anti-glutathione, anti-sulfenic acid, and anti-S-nitrosylation antibodies. Western blots revealed gross changes in glutathionylation, sulfenylation and nitrosylation PTMs associated with diet and pathology. To identify discrete sites targeted by these PTMs we performed isobaric tagging (TMT) and enriched for modified peptides using thiol-affinity chromatography, with subsequent analysis by LC-MS/MS on a Thermo Q-Exactive HF-X. Significantly modified proteins, changing in abundance between one or more treatment group in comparison to C Cit, were mapped to enriched regulatory pathways such as tricarboxylic acid cycle, phosphoinositol 3-kinase/protein kinase B pathway and glycolysis/gluconeogenesis pathway. Our data suggests that the rat renal redoxome is sensitive to diet and T2DM, indicating a possible physiological remodelling role for cysteine oxidation in DN.

660

PromarkerD: A novel test for predicting rapid decline in renal function in type 2 diabetes

Scott Bringans[1], Richard Lipscombe[1], Kirsten Peters[1, 2], Tammy Casey[1], Jason Ito[1], Wendy Davis[2], Timothy Davis[2]

1. Proteomics International, Broadway, Nedlands, WA, Australia

2. Medical School, University of Western Australia, Perth, WA, Australia

Chronic kidney disease (CKD) affects one in three adults with diabetes, accounting for 40,000 deaths and $100 billion (USD) in healthcare spending annually. The current tests for CKD, urinary albumin:creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR), have limited accuracy to predict CKD progression.

PromarkerD is a blood test that measures three plasma protein biomarkers (CD5L, ApoA4, and IBP3) combined with three clinical factors (age, HDL-cholesterol and eGFR) to predict risk of renal decline in patients with type 2 diabetes (T2D) over the ensuing 4 years. PromarkerD was developed using a proteomics workflow in patients with T2D drawn from the longitudinal observational Fremantle Diabetes Study Phase II (FDS2). Two versions of the test were developed, a targeted mass spectrometry based assay and an ELISA assay, both of which were used to measure the PromarkerD plasma biomarkers. Risk predictions were compared between the two platforms using Bland and Altman plot analysis.

The plasma biomarkers add significant incremental benefit to conventional clinical risk factors for predicting rapid decline in renal function in T2D. The biomarkers outperform both eGFR and ACR for predicting future renal decline and provide physicians with a more informed approach to managing diabetic kidney disease and patient care. PromarkerD may also be useful for risk stratification in future clinical trials.

661

Screening of DNA aptamers toward oral cancer biomarkers using plate-based SELEX and next-generation sequencing

CHIEN-CHUN LIU[1], Yung-Chin Hsiao[1], Jau-Song Yu[1]

1. Molecular Medicine Research Center, CHANG GUNG UNIVERSITY, TAOYUAN, Taiwan

Identification of proteins as biomarkers for early detection of cancer and therapeutic targets for cancer treatment are important issues in precision medicine. Oral cavity squamous cell carcinoma (OSCC) is a common cancer worldwide and represents a serious and growing problem in many parts of the globe, including Taiwan. To help early detection of OSCC, our group have put efforts in the past decades to discover/verify potential protein biomarkers using saliva, a kind of protein-rich body fluid that directly contacts the oral lesions, and we developed a four-protein marker panel, which may be a clinically effective tool for detecting OSCC and monitoring high-risk oral potentially malignant disorders (Yu et al., PNAS USA, 2016 Oct 11;113(41):11549-11554). Here, we attempted to screen DNA aptamers toward these four biomarkers for further diagnosis assay development. A plated-based platform of systematic evolution of ligands by exponential enrichment (SELEX) was developed for screening DNA aptamers against one of OSCC biomarkers, kininogen-1 (KNG1). After ten rounds of SELEX, the reserved DNA pools were submitted to next-generation sequencing (NGS), and six anti-KNG1 aptamers were discovered. The K_D value of these aptamers toward KNG1 can achieve 10^{-9} M level. Aptamers have many advantage comparing with antibodies, such as small size, stability, easy modification, low cost, and batch-to-batch consistency. Using aptamer to develop detection kits is an alternative strategy for biomarker measurement. These newly generated anti-KNG1 aptamers represent valuable tools for future development of the sandwich-ELISA assay and lateral flow strip assay.

662

Construction of miRNA-RNA networks underlying Taiwanese colorectal cancer for novel targets of therapy and precision medicine

Hsuan Liu[2, 3, 1, 4, 5], Pei-Ting Hsu[5], Yi-Shiuan Lai[3]

1. Molecular Medicine Research Center, Chang Gung University, Taoyuan, Guishan, Taiwan

2. Chang Gung University, Tao-Yuan, Taiwan

3. Department of Cell and Molecular Biology, Chang Gung University, Taoyuan, Guishan, Taiwan

4. Division of Colon and Rectal Surgery, Chang Gung Memorial Hospital, Linkou, Taipei new city, Taiwan

5. Graduate Institute of Biomedical Sciences, Chang Gung University, Taoyuan, Guishan, Taiwan

With growing cancer omics data, the complete functional genomic regulatory network of cancer has been progressively constructed, leading to many new breakthroughs in cancer treatment. Although numerous studies have identified microRNAs as tumor markers, the complete regulatory network and molecular mechanism of their actions is still fragmented, with many key clinical questions concerning treatment and patient outcome still unresolved. This project therefore aims to employ a comprehensive and systematics approach to demarcate and identify novel miRNA-mRNA gene regulatory network and further provide an in-depth understanding of their molecular actions and pathological mechanisms. In the first part, relying on our in-house colorectal cancer mRNA and small RNA sequencing data of 104 pairs of Taiwanese colorectal cancer (CRC) samples, we performed a miRNA-mRNA regulation network analysis to comparatively profile distinctions between tumors and matched normal tissues. After integration with the public TCGA CRC database, we identified two novel miRNAs, which were significantly up-regulated in Taiwanese CRC specimens but conversely down-regulated in the U.S.-based TCGA dataset. We also validated their tumor-associated expression profiles in a second independent clinical cohort. Functionally, cell-based studies showed that these miRNAs exert significant effect on the growth, migration, and invasiveness of CRC tumor cells. In addition, through NGS-based RNA-sequencing and bioinformatics analysis, several potential target genes of these miRNAs and the associated regulatory networks were delineated. Using gene expression and proteomics-based analyses, we further showed that these miRNAs would

negatively regulate the expression of several cell cycle- and migration-related genes and proteins. Viewed together, our study provides support to an ethnic basis of CRC tumorigenesis and further uncovers two novel miRNA biomarkers with functional and translational implications.

68

Proteomic profiling of plasma lipoprotein particles as a tool to identify novel subspecies.

Sonja Hartwig[1, 2], Joerg Kotzka[1, 2], Stefan Lehr[1, 2]

1. German Diabetes Center (DDZ), German Diabetes Center, Duesseldorf, NRW, Germany

2. German Center for Diabetes Research (DZD), Duesseldorf, NRW, Germany

Introduction:

Plasma lipoprotein particles are complex microemulsions of proteins and lipids, which can be divided into seven classes based on size, lipid composition, and apolipoproteins (chylomicrons, chylomicron remnants, VLDL, IDL, LDL, HDL and Lp (a)). Recent studies reveal that the protein composition of each lipoprotein particle class is highly diverse and modulations of these may have significant impact on the development of metabolic diseases. The aim of this study is to comprehensively characterize the protein signatures of the known lipoprotein particles derived from human plasma samples to identify novel subspecies.

Methods:

Human plasma samples (0.5ml) were separated by size exclusion FPLC on a Superpose 6 column. To assign the fractions to dedicated lipoprotein classes apolipoprotein marker proteins were monitored by western blot analyses and matched to cholesterol and triglyceride content. In order to avoid contamination from co-eluting plasma proteins lipoprotein particles fractions were cleaned up with Calcium Silicate Hydrate (CSH). Subsequently, isolated lipoprotein particles were measured by high resolution mass spectrometry (Orbitrap Lumos).

Results:

Continuous FPLC-separation revealed 20 fractions containing lipoprotein particles. By means of the marker protein and cholesterol/triglyceride distribution, fractions were assigned to distinct lipoprotein particles, i.e. VLDL, IDL, LDL and HDL. Proteomic profiling by mass spectrometry assigned more than 400 different proteins to the lipoprotein fractions.

Conclusion:

Our proteomic profiling approach provide a basis to identify novel compositions of known lipoprotein particles which might represent as predictive markers for multifactorial metabolic diseases including type 2 diabetes.

69

Relating high density lipoprotein (HDL) particle composition to clinical signaling capacity

Sandra Goetze[1, 2], Mathias Cardner[3], Mustafa Yalcinkaya[4], Lucia Rohrer[4], Edlira Luca[5], Miroslav Balasz[6], Gerhard Liebisch[7], Manfred Claassen[8], Jan Krützfeldt[5], Christian Wolfrum[6], Niko Beerenwinkel[3], Arnold von Eckardstein[4], Bernd Wollscheid[1, 2]

1. Department of Health Sciences and Technology (D-HEST)/Institute of Molecular Systems Biology (IMSB), ETHZ, Zurich, Switzerland

2. Swiss Institute of Bioinformatics (SIB), Zurich, Switzerland

3. BSSE, ETHZ, Zurich, Switzerland

4. Institute for Clinical Chemistry, University Hospital Zurich, Zurich, Switzerland

5. Department of Diabetology and Endocrinology, University Hospital Zurich, Zurich, Switzerland

6. Dep. of Health Sciences and Technology, ETHZ, Zurich, Switzerland

7. Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Regensburg, Regensburg, Germany

8. Institute of Molecular Systems Biology, ETHZ, Zurich, Switzerland

HDL is a complex molecular particle mediating reverse cholesterol transport in the human body. High plasma levels of HDL cholesterol are clinically associated with a lower risk of coronary heart disease (CHD) and diabetes mellitus type 2 (T2DM). However, HDL could not yet be successfully exploited for the prevention or treatment of disease. This is mainly because the structure-function relationship of this complex particle is still unresolved. A prerequisite for establishing such a structure-function relationship would be the detailed molecular knowledge about the components of the HDL particle in health and disease.

Here, we set out to characterize the protein and lipid composition of HDL particles from normal individuals and patients with CHD and/or T2DM. To link particle composition with phenotypic functionality, we implemented a set of eleven cellular *in vitro* assays as functional readouts for vasoprotective and anti-diabetic effects mediated by HDL. To quantitatively analyze the HDL particle proteotype of a larger patient cohort we made use of data-independent acquisition mass spectrometry (DIA/SWATH-MS). We first established a DIA/SWATH library from pooled HDL samples, which resulted in a DIA spectral library representing 356 protein groups. We used the HDL library in order to digitize a clinical HDL sample cohort consisting of 166 patients including healthy controls, resulting in the quantitation of 182 protein groups across the cohort. Simultaneously, the phenotypic impact of these HDL particles was tested in our cellular model systems as measures for clinical/biological functionality. Bioinformatics analysis using elastic net regularization revealed novel structure-function determinants which were tested *in vitro* using reconstituted minimal HDL particles of defined protein/lipid composition. Here, GPLD1 was found to be one of the major determinants of HDL-mediated anti-apoptotic function.

Together, we established an integrative approach and bioinformatical framework supporting the discovery of novel HDL-determinants of clinical relevance.

1. Gordon, S. M., Deng, J., Lu, L. J. & Davidson, W. S. Proteomic characterization of human plasma high density lipoprotein fractionated by gel filtration chromatography. J. Proteome Res. 9, 5239–5249 (2010).

2. Kratzer, A., Giral, H. & Landmesser, U. High-density lipoproteins as modulators of endothelial cell functions: alterations in patients with coronary artery disease. Cardiovasc. Res. 103, 350–361 (2014).
3. von Eckardstein, A. & Widmann, C. High-density lipoprotein, beta cells, and diabetes. Cardiovasc. Res. 103, 384–394 (2014).
4. von Eckardstein, A. A Novel Biomarker Approach to Exploit HDL for Risk Assessment. J. Am. Coll. Cardiol. 73, 2146–2149 (2019).

7

Spontaneous chemical modifications in long-lived proteins prevent lysosomal degradation: implications for age-related diseases

Ryan R Julian[1], Tyler Lambeth[1], Dylan Riggs[1]

1. UC Riverside, Riverside, CA, United States

Lysosomal malfunction is associated with many age-related disorders including Alzheimer's disease and macular degeneration. The lysosome plays a crucial role in autophagy and is responsible for breaking down proteins into amino acids, but certain spontaneous chemical modifications associated with long-lived proteins can interfere with this process. In particular, isomerization/epimerization (which can accumulate rapidly in peptides or disordered protein domains containing aspartic acid or by deamidation of asparagine) are known to interfere with proteolysis. Unfortunately, isomerization/epimerization do not lead to a change in mass, making them invisible to traditional proteomics. However, it is possible to distinguish peptide isomers in MS/MS spectra if the dissociation method is sensitive to structural differences. Typically collision-induced dissociation struggles to identify isomers, but we have demonstrated that enhanced structural sensitivity of radical-directed dissociation is ideal for isomer/epimer identification. We have identified many sites of isomerization in long-lived proteins derived from the eye lens, but have also demonstrated that such modifications can occur in a few weeks and are relevant to other tissues.

Here, we explore the effect of isomerization/epimerization on lysosomal degradation by examining proteolysis of a variety of model peptides with a suite of cathepsins, the major proteases in the lysosome. It is demonstrated that proteolysis is prevented to varying degrees by all isomer/epimers. In some instances, protection of 13 residues is afforded by a single, centrally located isomerized residue. This behavior can easily be rationalized by examining the structural details of protease active sites, which are defined by a binding groove where the peptide backbone must be precisely localized for hydrolysis. Isomerization/epimerization both represent structural changes that cannot be accommodated by these binding grooves.

These results foreshadow consequences for autophagy and proteostasis. The inability of lysosomes to break down isomerized peptides should lead to eventual lysosomal storage, which is observed in Alzheimer's disease.

70

Cross-species tissue proteomics analyses in cardiovascular disease: unraveling Ariadne's thread

Marika Mokou[1], Julie Klein[2], Manousos Makridakis[1], Vasiliki Bitsika[1], Jean-Loup Bascands[3], Jean Sebastien Saulnier-Blache[2], William Mullen[4], Michael Sacherer[5], Jerome Zoidakis[1], Burkert Pieske[6], Harald Mischak[4], Maria Roubelakis[7], Joost P. Schanstra[2], Antonia Vlahou[1]

1. Biotechnology Laboratory, Biomedical Research Foundation of the Academy of Athens, Athens, Greece

2. U1048, Institut of Cardiovascular and Metabolic Disease, Institut National de la Santé et de la Recherche Médicale (INSERM), Toulouse, France

3. U1188, Sainte Clotilde, Institut National de la Santé et de la Recherche Médicale (INSERM), La Réunion, France

4. BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, United Kingdom

5. Department of Cardiology, Medical University of Graz, Graz, Austria

6. Department of Internal Medicine and Cardiology, Charité University Medicine, Berlin, Germany

7. Laboratory of Biology, University of Athens, School of Medicine, Athens, Greece

Background: Tissue proteomics holds great promise to unravel the pathophysiological mechanisms of cardiovascular disease (CVD). In this study, through a cross-species proteomics-based integrated approach we intend to shed light on the molecular mechanisms of CVD.

Methods: LC-MS/MS analysis was performed on thoracic aortas from the Ldlr-/- and ApoE-/- mouse models in the absence or presence of STZ-induced diabetes and their wild type (WT) littermates as well as in vessels from patients with CVD and healthy individuals. Key findings were evaluated by western blot and *in vitro* using the MTS, transwell migration, and tube formation assay.

Results: The proteomic landscape was initially characterized in diabetic and non-diabetic Ldlr-/- atherosclerotic mouse models that led to identification of 284 differentially expressed proteins compared to WT mice. To exclude protein changes specific to the disease background, high-throughput proteomic analysis was also performed in diabetic ApoE-/- mice and 321 proteins were identified differentially expressed in comparison to WT mice. Among them, 177 proteins were common and showed similar expression trend throughout all atherosclerotic models. The high relevance of these proteins with the disease was further supported by *in silico* analysis. To translate these findings to human disease, LC-MS/MS was subsequently performed in human vessels from patients with CVD and healthy individuals. Through a cross-species comparison overlapping pathways and proteins between mouse and humans were highlighted. The most pronounced overexpression in disease was observed for the KDM5D histone demethylase which was accompanied by a reduction of its substrate -the trimethylated form of H3K4. An increase of this substrate through KDM5 inhibition on human endothelial cells decreased cell proliferation, migration and tube-forming ability *in vitro*.

Conclusions: This cross-species proteomics approach supports the involvement of the KDM5 epigenetic modulators in CVD progression. Furthermore, the high resolution proteomic datasets hold the promise of unraveling the complexity of CVD mechanisms.

701

Profiling of intact N-linked glycopeptides in the rat brain

Jialin Liu[1], Yi Wan[2], Pengyuan Yang[1], Xing Chen[2]

1. IBS, Fudan university, Shanghai, China

2. College of Chemistry and Molecular Engineering, Peking university, Beijing, China

Protein N-linked glycosylation plays a key role in various biological and pathological processes in the brain. Interestingly, different brain cells such as neurons and astrocytes can have distinct N-glycosylation patterns. However, it remains challenging to analyze brain N-glycosylation in a cell-type specific manner. Here we report the glycoproteomic studies of N-linked glycosylation in primary rat neurons and astrocytes. By combining primary brain cell culture and mass spectrometry-based analysis of intact glycopepties, we identified 6461 glycopeptides, 1684 glycosides and 770 glycoproteins, including 144 sialic acid glycoproteins. These N-linked glycosylated proteins are highly enriched in cell communication and adhesion. Profiling of intact glycopeptides provides a powerful tool for investing the biological function of brain glycosylation.

1. P. Fang, X. Wang, Y. Xue, M. Liu, et al., In-depth mapping of the mouse brain Nglycoproteome reveals widespread N-glycosylation of diverse brain proteins, Oncotarget 7 (2016) 38796–38809

702

Lysine Succinylome Analysis Demonstrates the Involvement of the Manganese-stabilizing Protein PsbO Succinylation in Cyanobacterial Photosynthesis Regulation

Xin Liu[1, 2], Mingkun Yang[2], Feng Ge[2]

1. University of Chinese Academy of Sciences, Beijing, China

2. Institute of Hydrobiology,Chinese Academy of Sciences, Wuhan, Hubei Province, China

Lysine succinylation is a newly-identified protein post-translational modification pathway, which is present in both eukaryotic and prokaryotic cells, but its extent and function in photosynthetic organisms remain unexplored. We performed a global succinylome analysis of a model cyanobacterium *Synechococcus* sp. PCC 7002 using high-accuracy tandem mass spectrometry (MS/MS) analysis in combination with the enrichment of succinylated peptides from digested cell lysates and subsequent peptide identification. In total, we identified 265 lysine succinylation sites on 129 proteins, which were involved in various biological processes; however, a large proportion of the succinylation sites were present on proteins in the photosynthetic pathway. Functional studies showed that succinylation of the manganese-stabilizing protein PsbO can decrease the oxygen evolution rates of the photosystem (PS) II center and the efficiency of energy transfer during the photosynthetic reaction. Our results demonstrate that succinylation may be a mechanism involved in acclimation to high light in *Synechococcus* as well as play a regulatory role in photosynthesis. Molecular dynamics simulations elucidated a mechanism that may allow succinylation to influence the efficiency of photosynthesis by altering the conformation of PsbO, thereby hindering the interaction between PsbO and the PSII core. Our findings provide novel insights into the functions of succinylation in the regulation of photosynthesis, thereby suggesting that reversible succinylation may be an important regulatory mechanism during photosynthesis in *Synechococcus*, as well as in other photosynthetic organisms.

703

Enrichment of Lowly-Hydrophilic *N*- and *O*-glycopeptides using Ion-Pairing ZIC-HILIC-SPE

Ian Loke[1], Harry Tjondro[2], Rebeca Sakuma[2], Katalin Medzihradszky[3], Morten Thaysen-Andersen[2]

1. Department of Biological Sciences, National University of Singapore, Singapore

2. Department of Molecular Sciences, Maccquarie University, Sydney, New South Wales, Australia

3. UCSF Mass Spectrometry Facility, University of California, San Francisco, California, USA

Trifluoroacetic acid-based Ion-pairing zwitterionic hydrophilic interaction liquid chromatography solid phase extraction (TFA-IP-ZIC-HILIC-SPE) is a key sample processing step in quantitative glycoproteomics that provides unbiased enrichment of common *N*-glycopeptides from complex peptide mixtures[1]. With the discovery of truncated and lowly-hydrophilic *N*-glycans in the mammalian glycoproteome spanning the paucimannosidic ($Man_{1-3}GlcNAc_2Fuc_{0-1}$)[2] and chitobiose core ($GlcNAc_{1-2}Fuc_{0-1}$)[3,4] type structures, it becomes important to investigate if such *N*-glycopeptides and *O*-glycopeptides are efficiently enriched using TFA-IP-ZIC-HILIC-SPE. We here investigate this by performing quantitative LC-MS/MS profiling of various tryptic and non-tryptic peptide mixtures containing human *N*- and *O*-glycopeptides carrying paucimannosidic, chitobiose and *O*-GlcNAc moieties using ZIC-HILIC-SPE enrichment with formic acid (FA) and TFA as mobile phase additives. The paucimannosidic peptides including the lowly hydrophilic $Man_1GlcNAc_2$-peptides were quantitatively retained using TFA, but the even less hydrophilic chitobiose core peptides, in particular the $GlcNAc_1$-peptides, were often under-represented. Interestingly, using the less hydrophobic ion-pairing reagent FA, both paucimannosidic and chitobiose core peptides were more often quantitatively retained. The *O*-GlcNAc peptide profiling recapitulated these trends by demonstrating dramatically better retention using FA-IP-ZIC-HILIC-SPE than under TFA condition. The comparatively better retention of lowly hydrophilic glycopeptides with FA-IP-ZIC-HILIC-SPE was rationalised using *in silico* calculations of the peptide hydropathy ($\Delta G_{oct\text{-}water}$) identified relative to published and putative glycopeptides in the human and mouse glycoproteomes, which suggest a significant peptide carrier influence on their retention propensity. In conclusion, we here show that paucimannosidic peptides, but not necessarily chitobiose core peptides, generally are well-retained

when using the conventional TFA-IP-ZIC-HILIC-SPE enrichment whereas the less used FA-IP-ZIC-HILIC-SPE appears to be a universally better strategy for quantitative glycoproteomics of mixtures containing lowly-hydrophilic glycopeptides.

1. Mysling et al. Anal Chem. 2010. 82(13):5598.
2. Loke et al. Mol Cell Proteomics. 2017. 16(8):1507.
3. Trinidad et al. Mol Cell Proteomics. 2013. 12(12):3474.
4. Loke et al. Biomolecules. 2015. 5(3):1832.

704

propro: a web-based and cloud computing platform for Data-Independent Acquisition Mass Spectrometry Data Interpretation and Visualization

Miaoshan Lu[1], Ruimin Wang[1], Shaowei An[1], Tiannan Guo[1], Changbin Yu[1]

1. Westlake University, Hangzhou, NOT US OR CANADA, China

Data-independent acquisition (DIA) mass spectrometry is gaining popularity in proteomics research however the data interpretation and visualization are limiting its application and further development due to the huge data amount and the requirement of heavy computing.

Here, we developed a novel one-stop cloud platform for convenient DIA data interpretation and visualization which can be deployed on both private and public cloud servers. The advancement of propro over existing software tools include: 1) a novel data format "Aird" with up to 89%-95% compression ratio; 2) experiment analysis with a 420000 target peptides library only cost ~15 minutes by using a personal computer (CPU: i7 7700 4.2GHz, 16G memory, HDD Disk) but with up to ~20% more peptide identifications; 3) Graphic interfaces for every intermediate steps which enable users to perform the entire workflow transparently; 4) availability of OpenAPI for community-based further development of the software; 5) also allows effective analysis and visualization of parallel reaction monitoring (PRM) data sets.

In conclusion, propro is a novel and comprehensive cloud platform for Interpretation and visualizing DIA data sets. It is available at Github: https://github.com/Propro-Studio/propro-server. Everyone can visit propro on http://www.propro.club

705

Development of lectin-based glycopeptide-enrichment approach coupled with MRM assay for verifying oral cancer candidate glycobiomarkers

Yu-Ting Lu[1], Chia-Chun Wu[1], Wei-Fan Chiang[2], Jau-Song Yu[1]

1. Chang Gung University, Taoyuan, KWEI-SHAN, Taiwan

2. Department of Oral and Maxillofacial Surgery, Chi-Mei Medical Center, Tainan, Liouying, Taiwan

Oral squamous cell carcinoma (OSCC) is the fifth most common tumor causing high mortality in Taiwan, but currently there are still no available biomarkers for detecting early-stage OSCC. Abnormal glycosylation is known as a pivotal regulatory mechanism in tumor malignancy, and saliva is a non-invasive body fluid derived from oral cavity with abundant glycoproteins. Recently, we have applied iTRAQ-based quantitative glycoproteomics coupled with pisum sativum agglutinin (PSA) enrichment of glycopeptides to analyze the differential glycoproteome profiles in saliva samples from healthy controls, subjects with oral potentially malignant disorders and OSCC, from which several glycopeptides showing significantly increased levels in OSCC saliva were discovered. To facilitate the future verification of these target glycopeptides in saliva samples, the present study aims to develop an automatic KingFisher-assisted glycopeptides enrichment procedure. We chose toluenesulfonyl (tosyl)-activated magnetic beads to couple PSA and optimized the conditions for coupling maximal amounts of PSA to the beads and for blocking and pre-cleaning beads to reduce nonspecific binding of protein contaminants prior to LC-MRM-MS analysis. We also evaluated the appropriate amount of saliva sample for LC-MRM-MS assay using the afore-mentioned optimized procedure for glycopeptie enrichment. Afterward, we incorporated the optimized condition for the magnetic bead-based glycopeptides enrichment procedure into the KingFisher magnetic bead processor and successfully created a workflow for the entire process. We found that the KF-Flex-96-DW (deep well)-head magnetic base performed better than the KF-Flex-96-PCR-head magnetic base in fully suspending the magnetic beads (5 mm) and moving beads to other plates. This established workflow for automatic KingFisher-assisted glycopeptides enrichment procedure will be applied to quantify target glycopeptides in individual saliva samples from OSCC patients and non-OSCC subjects.

706

Avoiding abundance bias in the functional annotation of PTM proteins

David B Lyon[1], Lars J Jensen[2], Christian von Mering[1]

1. Systems Biology / Metagenomics, University of Zurich / Swiss Institute for Bioinformatics, Zurich, Switzerland

2. Cellular Network Biology, University of Copenhagen, Novo Nordisk Center for Protein Research, Copenhagen, Denmark

Mass Spectrometry based shotgun-proteomics is inherently biased towards abundant proteins. This bias can influence GO-term enrichment analysis by showing enriched terms for abundant rather than e.g. post-translationally-modified (PTM) proteins. We have developed a method to correct for this bias and a freely accessible web-tool (https://agotool.sund.ku.dk/) to facilitate the use for the scientific community. We're expanding our protein-centric enrichment tool to additionally perform analysis for the following categories Reactome, Interpro and Pfam domains, as well as Disease Ontology, Brenda Tissue Ontology, and PubMed identifiers. Apart from regular automated updates, we provide protein-group support, redundancy reduction, as well as an API for programmatic access.

707

Quantitative phosphoproteomics in analysis of HipA-mediated persistence

Boris Macek[1]

1. University of Tuebingen, Tuebingen, Germany

Background: Bacterial persistence, the ability to survive antibiotic treatment by entering a physiologically dormant state, is a serious biomedical problem. Protein Ser/Thr kinase HipA, the first toxin connected to bacterial multidrug tolerance (persistence), exerts its function by phosphorylating glutamate--tRNA ligase GltX, leading to a halt in translation, accumulation of ppGpp and induction of persistence. Intriguingly, its variant HipA7 is able to induce significantly higher levels of persistence despite being less toxic for the cell.

Objectives: We postulated that the phenotypic difference between HipA and HipA7 may be driven by diverse substrate pools of the two kinase variants. We aimed to analyze regulatory networks involved in bacterial persistence using MS-based proteomics.

Methods: We ectopically expressed *hipA* and *hipA7* in *E. coli* and monitored their in vivo substrates during growth inhibition and resuscitation using SILAC-based quantitative phosphoproteomics. We applied dynamic SILAC to study protein synthesis and turnover during HipA-mediated persistence and resuscitation.

Results: Our assays confirmed that both forms of the HipA kinase phosphorylate GltX as the main substrate. Importantly, HipA phosphorylated several additional substrates involved in translation, transcription and replication, such as ribosomal protein L11 and SeqA, which were further validated in vitro. Conversely, HipA7 had a lower kinase activity, no additional substrates under tested conditions and showed a similar substrate pool only when expressed at significantly higher levels. The two forms of the kinase also differed in autophosphorylation level, which was significantly lower in HipA7. When expressed from the chromosome, HipA7 phosphorylated GltX and another substrate, a putative Ser/Thr kinase. Our preliminary data show that this novel kinase is also involved in persistence, revealing the first protein phosphorylation-based network involved in regulation of persistence. Our results contribute to understanding of HipA action and present a resource for future studies of bacterial persistence.

708

A comparison of cartridge and bead-based sample preparation strategies for bottom-up proteome analysis of detergent-containing samples

Jessica L. Nickerson[1], **Kent MacLean**[2], **Alan A. Doucette**[1]

1. Chemistry, Dalhousie University, Halifax, NS, Canada

2. Proteoform Scientific, Halifax, NS, Canada

Introduction

Detergents including sodium dodecyl sulfate (SDS) are becoming increasingly accepted in proteomics workflows. This is owing to a growing list of semi-automated approaches, designed to effectively purify and digest protein mixtures ahead of bottom up LC-MS/MS analysis. Since the advent of filter-aided sample preparation (FASP) as a cartridge-based format which simplifies SDS removal and protein digestion, numerous other strategies have been reported enabling the MS acquisition of the resulting peptide digests.[1] Among the cartridge and bead-based technologies are Suspension Trapping S-Trap, in-StageTip (iST), Single-Pot, Solid-Phase-Enhanced Sample Preparation (SP3), and the ProTrap XG, all of whom report successful acquisition of peptide lists from SDS-containing samples.[1,2,3,4,5]Here, we directly evaluate the performance of these sample preparation strategies looking at four specific figures of merit: the protein/ peptide recovery, sample purity as judged by residual level of SDS, protein digestion efficiency, and sample throughput/ process time.

Methods

Sample preparation will be conducted within FASP, S-Trap, iST, SP3 and ProTrap XG, based on the respective user manuals. Test proteome mixtures will be prepared in SDS at varying concentrations of protein to assess the effectiveness of the sample preparation at different sample loads. Protein recovery will be quantified by the bicinchoninic acid (BCA) assay, as well as SDS PAGE, and LC-UV analysis. Purity will be assessed by quantifying residual SDS using the methylene blue active substances assay. The resulting peptide mixtures will be subject to bottom up LC-MS/MS analysis, assessing both the number of peptides, degree of missed cleavage, and coverage of the proteome with respect to intrinsic protein properties (molecular weight, hydrophobicity).

Results

Methods are in place to evaluate performance (purity/ yield/ digestion efficiency) of each sample preparation strategy.

Conclusions

Multiple strategies exist for bottom-up MS analysis of detergent-containing samples. This study provides unbiased assessment of the performance characteristics of several popular strategies.

1. 1. Mann et al. (2007). Nature Protocols, 2, 1896-1906.
2. 2. Mann et al. (2009). Nature Methods, 6, 359-62.
3. 3. Krijgsveld et al. (2014). Molecular Systems Biology, 10.
4. 4. Yu et al. (2018). Journal of Proteome Research, 17, 2917-2924.
5. 5. Doucette et al. (2015). Journal of Proteomics, 118, 140-150.

709

Proteome biomarkers for reflux aspiration in cystic fibrosis

Rosemary E Maher[1], Robert W Lord[2], Victoria M Harman[1], Jaclyn A Smith[2], Andrew M Jones[2], Paul McNamara[3], Robert J Beynon[1]

1. Centre for Proteome Research, University of Liverpool, Liverpool, Merseyside, United Kingdom

2. Manchester Adult Cystic Fibrosis Centre, University Hospital of South Manchester, Manchester, United Kingdom

3. Respiratory Unit, Alder Hey Children's Hospital NHS Foundation Trust, Liverpool, United Kingdom

The introduction of gastric contents into the respiratory tract, a process known as reflux aspiration, potentially contributes to lung damage in cystic fibrosis (CF). At present, there is no method to identify patients with reflux aspiration, presenting a major challenge for the management of patients presumed to be at risk. Here we describe a quantitative proteomics approach to assess differences in the sputum proteome between CF patients based on gastro-oesophageal reflux (GOR) measures to identify protein biomarkers for reflux aspiration.

Spontaneous sputum samples were collected from CF subjects (n=37), and sputum was induced from healthy volunteers (HV, n=33) for comparison. The sputum proteome was analysed by bottom-up shotgun proteomic analysis using liquid chromatography/mass spectrometry. All CF patients had GOR measured using 24-hour pH-impedance; sputum samples were compared from those with the highest (n=5) and lowest (n=5) risk for reflux aspiration, based on total, proximal and/or supine reflux episodes.

There were significant differences between the CF and HV sputum proteome with large increases in inflammatory proteins, predominantly neutrophil granulocyte proteins. There were also significant differences in the proteome between CF patients with highest and lowest reflux measures. These protein differences were not indicative of inflammation but other underlying cellular processes. No proteins of gastric origin were identified by shotgun proteomic analysis.

These data suggest the sputum proteome of CF subjects may be influenced by gastro-oesophageal reflux. Although this appears to support a relationship between reflux and CF lung disease, it is not currently possible to conclude if this is a consequence of reflux aspiration. Gastric proteins, if present, will likely be at low abundance and more targeted approaches may be required for their detection. Using targeted analysis against a panel of gastric proteins, we aim to explore further the sputum proteome differences and identify specific biomarkers to diagnose reflux aspiration.

71

An anatomically resolved human heart transcriptome and proteome landscapes reveal molecular signatures and disease-relevant pathways of end-stage dilated cardiomyopathy

Ling Lin[1], Shanshan Liu[1], Juanjuan Xie[1], Zhangwei Chen[2], Mingqiang Fu[2], Juying Qian[2], Huali Shen[1, 3], Pengyuan Yang[1, 3]

1. Fudan University, Shanghai, SHANGHAI, China

2. Department of Cardiology, Shanghai Institute of Cardiovascular Diseases, Zhongshan Hospital, Fudan University, Shanghai, 20032, China

3. Institutes of Biomedical Sciences of Shanghai Medical School, Fudan University, Shanghai, China

Dilated cardiomyopathy (DCM) is characterized by the presence of left ventricular dilatation and contractile dysfunction [1-2]. Anatomically, the human heart is divided into four chambers, but the transcript and protein profiling of human myocardial tissue with DCM are not well understood [3-5]. Here, we performed high-throughput transcriptomic and proteomic analyses on human hearts from healthy control as well as end-stage DCM, which served as a systems-level perspective in characterizing cardiac biology and exploring DCM mechanism. Comparisons of the 9,219 identified proteins in four chambers of human hearts with deep sequencing data of the 14,880 transcripts indicated deep coverage of the proteome. The core proteome of healthy hearts was significantly enriched in biological functions including extracellular matrix organization, cytoskeleton organization, biosynthetic process signal transduction, ect. Comparing DCM hearts with healthy ones via iTRAQ-based quantitative proteomic analysis, the changing patterns in the left ventricle were more extensive and drastic compared to the other chambers in end-stage DCM, among which ECM organization, mitochondrial function and muscle contraction were the most enriched three GOBP terms. Furthermore, oxidative phosphorylation, calcium signal pathway, and carbohydrate metabolism were disturbed during DCM development. Our comprehensive proteomic and transcriptomic analyses presented precise quantification of potential signatures, signaling pathways, regulatory networks, and characteristic differences in each DCM chamber. Additionally, 8 secreted proteins were selected from the total 608 significant changed cardiac proteins, and further verified in a cohort of 53 DCM serums and 35 healthy serums via ELISA analysis (P<0.05). The combined use of Complement component C9, cathepsin B, dickkopf-related protein 3 and von Willebrand factor as serum biomarker panel showed high diagnostic potential in early detection of DCM. These integrated datasets provide diverse and rich resources for researches to investigate the molecular basis of heart physiology and pathology including cardiomyopathies.

1. Weintraub RG, Semsarian C, Macdonald P. Dilated cardiomyopathy. Lancet, 2017; 390(10092):400-414.
2. Dadson K, Hauck L, Billia F. Molecular mechanisms in cardiomyopathy. Clin Sci, 2017; 131(13): 1375-1392.
3. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, et al. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. Genet Med 2014; 16(8):601-608.
4. Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, et al. Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. Cell Rep, 2019; 26(7): 1934-1950.e5.
5. Doll S, Dreßen M, Geyer PE, Itzhak DN, Braun C, Doppler SA, et al. Region and cell-type resolved quantitative proteomic map of the human heart. Nat Commun, 2017; 8(1): 1469.

710

Identification of key regulators in Prostate cancer from gene expression datasets of patients

Md Zubbair Dr. Malik[1], R K Brojen Dr. Singh[1]

1. Jawaharlal Nehru University, New Delhi, DELHI, India

Publish consent withheld

711

The long and short of a Genome-Wide Association Study identified long non-coding RNA in prostate cancer

MohammedAdil Malik[1, 3, 2], Panchadsaram Janaththaani[1, 3, 2], Srilakshmi Srinivasan[1, 3, 2], Pawel Sadowski[4], Jyotsna Batra[1, 3, 2]

1. Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia

2. Australian Prostate Cancer Research Centre (APCRCQ), Brisbane, Queensland, Australia

3. Translational Research Institute Pty Ltd, Brisbane, Queensland, Australia

4. Central Analytical Research Facility, Queensland University of Technology, Brisbane, Queensland, Australia

Background

Through Genome-Wide Association Studies (GWAS), *IRX4* has been identified to be associated with prostate cancer (PCa) risk at chromosome 5p15. Interestingly, our group discovered a long non-coding RNA (*IRX4lncRNA*) in the anti-sense strand of *IRX4* and a novel Insertion-Deletion Polymorphism (INDEL) regulating androgen mediated *IRX4* and *IRX4lncRNA* expression. The potential of lncRNAs to encode regulatory small peptides with indispensable regulatory functions, called micropeptides (miPEPs) has recently been discovered. Our preliminary *in silico* analysis indicated that *IRX4lncRNA* can potentially encode for the miPEPs. We aim to determine the role of *IRX4lncRNA*/miPEP and validate its efficacy as PCa biomarker.

Methodology

Sequential Window acquisition of all theoretical fragment ion spectra mass spectrometry (SWATH-MS) and Multiple Reaction Monitoring (MRM) were performed to identify and validate micropeptides coded by *IRX4lncRNA* in prostate cancer cells. To determine the functional role of *IRX4lncRNA*/micropeptides in prostate cancer, in-vitro and in-vivo assays assays are being performed using overexpression and knockdown models. Further, identification of these peptides in clinical patient serum/tissue samples will determine its expression and utility as a biomarker.

Results/expected results

SWATH-MS and MRM analysis validated the expression of two peptides encoded by *IRX4lncRNA* in PCa cells (PC3, VCaP and LnCaP). Overexpression models of the ORF encoding these peptides resulted in increased proliferation of prostate cancer cells. Additionally, mass spectrometry analysis of the overexpression and knockdown models of the ORF identified changes in the pathways asscoaued with ubiquitin mediated proteolysis, enrichment in Rap1signalling pathway and changes in fatty acid metabolism. These pathways will be analysed in detail to determine specific function of the *IRX4lncRNA*/miPEP. Moreover, these peptides are being tested in clinical patient serum/tissue samples to determine its expression and utility as a biomarker

Conclusion

The findings from this study will elucidate the role of IRX4lncRNA/miPEP and their biological impact in PCa.

712

Multi-omics analysis of a nutrient transport protein required for full virulence in *Campylobacter jejuni*

Lok Man[1, 2], Joel A. Cain[1, 2], William Klare[1, 2], Nestor Solis[1], Pamela Soh[1], Zeynep Sumer-Bayraktar[1, 2], Paula Niewold[3, 2], Ashleigh Dale[1, 2], Stuart J. Cordwell[1, 3, 2, 4]

1. School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

2. Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

3. Discipline of Pathology, School of Medical Sciences, The University of Sydney, Sydney, NSW, Australia

4. Sydney Mass Spectrometry, The University of Sydney, Sydney, NSW, Australia

Campylobacter jejuni is the leading cause of bacterial gastroenteritis in the developed world. Infection is predominantly caused by the consumption of undercooked or poorly prepared poultry. *C. jejuni* exists mainly as a commensal within the intestines of chickens, but is pathogenic in humans. While the mechanism of this difference is unknown, factors such as motility and nutrient uptake are significant in the host-pathogen nexus. *C. jejuni* is considered asaccharolytic and primarily utilizes amino and organic acids as carbon sources, with only some strains able to utilize fucose. We conducted label-based LC-MS/MS proteomics of *C. jejuni* NCTC11168O to identify proteins associated with growth in environments that mimic host conditions (e.g. deoxycholate, iron limitation, presence of mucin). We quantified 1561 proteins, equating to ~93% of the predicted *C. jejuni* proteome. The most significantly induced protein (mean 4.6-fold induction) was the product of the *cj0025* gene, which has been annotated as a 'putative C4-dicarboxylate transporter'. Deletion of *cj0025* resulted in reduced *C. jejuni* motility, increased susceptibility to polymyxin B, and reduced biofilm formation. Human epithelial cell infection assays confirmed *Δcj0025 C. jejuni* demonstrated significantly reduced invasion. To determine the function of Cj0025, metabolomic profiles of media inoculated with *C. jejuni* wild-

type or *Δcj0025* were compared, with focus on the uptake of amino acids and Krebs cycle intermediates. These assays showed that *Δcj0025* was capable of utilizing all amino and organic acids commensurate with the wild-type. Sequence similarity to a family of bacterial cystine (Cys-Cys) transporters was determined and medium cystine levels were significantly maintained in *Δcj0025* mutants compared with wild-type. Growth supplemented with a toxic mimic of cystine, selenocystine, significantly inhibited wild-type growth, but did not affect *Δcj0025*. We confirm that Cj0025 is a cystine transporter, which we have named TcyP consistent with the nomenclature of homologous proteins.

713

Proteomics of cotton male gametophytes: heat stress dramatically impacts the early pollen developmental stage

Farhad Masoomi-Aladizgeh[1], Ullah Najeeb[2], Sara Hamzelou[3], Dana Pascovici[4], Ardeshir Amirkhani[4], Daniel K.Y. Tan[5], Mehdi Mirzaei[4], Paul A. Haynes[3], Brian J. Atwell[1]

1. Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia

2. Queensland Alliance for Agriculture & Food Innovation, The University of Queensland, Toowoomba, QLD 4350, Australia

3. Department of Molecular Sciences, Macquarie University, North Ryde, NSW 2109, Australia

4. Australian Proteome Analysis Facility, Macquarie University, North Ryde, NSW 2109, Australia

5. The University of Sydney, Plant Breeding Institute, Sydney Institute of Agriculture, School of Life and Environmental Sciences, Faculty of Science, Sydney, NSW 2006, Australia

As with all flowering plants, the formation of gametes in cotton (*Gossypium hirsutum*) is highly susceptible to heat. Substantial reduction in lint quantity and quality occur when developing flowers and bolls are exposed to sustained temperatures above 32 ºC, particularly as a result of the high vulnerability of plant male reproductive cells to heat stress. This study reports proteomic changes in the pollen of cotton (cultivar Sicot 71), aiming to identify heat-responsive proteins in the early and late male gametophyte. Two distinct phases of pollen development (tetrads and binucleate microspores) were exposed to 36/25 ºC (day/night) or 40/30 °C, each for 5 days. Subsequently, mature pollen grains were collected for quantitative label-free shotgun proteomic analysis. A total of 868 proteins was identified across all samples analysed. Differential expression analysis demonstrated that 48 proteins were up-regulated in mature pollen after the exposure of squares at the tetrad stage to 36/25 °C, while 56 proteins were down-regulated. Severe cell damage was observed when squares were exposed to 40/30 °C at the tetrad stage, resulting in failed dehiscence. Moderate heat (36/25 ºC) at the later binucleate microspore stage resulted in up-regulation and down-regulation of 47 and 24 proteins in the mature pollen grains, respectively, while extreme heat (40/30 °C) led to only ~30 proteins being differentially expressed. These differentially expressed proteins (DEPs) have functions related to metabolic process, catalytic activity, cellular process, binding, biological regulation, regulation of biological process, in addition to response to stimulus. The data revealed that heat shock proteins (HSPs) increased even long after heat stress. Interestingly, HSP70, one of the important HSPs, was up-regulated more than 100-fold under extreme heat stress.

714

Screening of optimal phase transfer surfactant cocktail for unbiased quantitative proteomics

Takeshi Masuda[1, 2, 3], Arisu Furukawa[1], Yuma Inamori[1], Shingo Ito[1, 2, 3], Sumio Ohtsuki[1, 2, 3]

1. Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, JAPAN

2. Faculty of Life Sciences, Kumamoto University, Kumamoto, JAPAN

3. AMED-CREST, Tokyo, JAPAN

Background

For giving highly sensitive and quantitative proteomics data, proteins are required to be completely extracted and digested into peptides. We have developed a sample preparation protocol using a phase transfer surfactant (PTS) cocktail containing sodium deoxycholate (SDC) and sodium lauroyl sarcosinate (SLS) (SDC/SLS). The protease activity and solubilization ability are highly enhanced in the PTS cocktail compared to other solubilizing agents. However, biased in digestion efficiency are still observed. The purpose of the present study was to identify a new PTS cocktail to minimize missed cleavage peptides toward unbiased quantitative proteomics.

Methods

Trypsin and lysyl endopeptidase activities in 31 PTS cocktails were measured by the absorbance of *p*-nitoraniline at 410 nm resulting from digestion of N-benzoyl-lysine-*p*-nitroanilide. For the evaluation of solubilization ability, proteins solubilized from human liver microsome fraction were quantified by a BCA assay. Proteins were digested at optimized condition and analyzed by using nanoLC-MS/MS.

Results and discussion

In the protease activity comparison, we selected 7 PTS cocktails showing comparable protease activity with SDC/SLS. The solubilization ability of these 7 PTS cocktails were greater than that of SDC/SLS. The protein solubilization ability of a PTS cocktail containing SDC and ursodeoxycholate (UDC) (SDC/UDC) was the greatest and 1.4-fold greater than those of SDC/SLS. Therefore, we used SDC/UDC in subsequent experiments. The number of missed cleavage sites was significantly decreased to 84% in SDC/UDC compared to SDC/SLS. Next, to evaluate whether the intensities of fully digested peptides are increased by the enhancement of digestion efficiency in SDC/UDC, we compared the intensities of these peptides between two PTSs. As expected, the number of peptides showing 2-fold higher in SDC/UDC was increased to 5.5-fold than those in SDC/SLS. These results suggest that the new PTS cocktail with SDC/UDC provides better unbiased quantitative proteomics data than the present PTS cocktail.

715

Comparative glycome analysis between exosomes, cell surface, and secreted glycoproteins with lectin microarray

Atsushi Matsuda[1], Maki Yoshida[2], Takanori Wagatsuma[2], Makoto Suematsu[1], Yasuaki Kabe[1], Hisashi Narimatsu[2], Atsushi Kuno[2]

1. Keio University School of Medicine, Shinjuku-ku, TOKYO, Japan

2. Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

Background: Exosomes are released from cells and circulated into the blood. Thus, exosomes are regarded as an attractive biological resource for the development of new cancer marker for liquid biopsy diagnosis. Exosomes generally covered with glycocalyx, which are controlled by a host cell glyco-synthetic machinery as similar to secreted and plasma membrane glycoproteins. Several exosomal subpopulations classified by tetraspanins have been investigated in relationship to diseases. However, their comparative analysis has never been attempted in terms of glycomics. In this study, we demonstratecomparative glycomic analysis between exosomes, secreted, and cell membrane glycoproteins derived from pancreatic cancer cells using lectin microarray system.

Methods: Pancreatic cancer culture cells derived secreted, plasma membrane glycoproteins, and exosomes were prepared, respectively. Exosomes were isolated and purified using commercially available exosome isolation kits. Further each CD antigens-positive exosomes were fractionated from total exosomes using specific antibodies against exosome marker (CD9, CD63, CD81) and were subjected to the comparative glycan profiling with lectin microarray.

Finding: As the result, it was found that the total exosome-derived glycoproteins contain abundant glycoproteins with specific glycan structures compared with secreted and membrane glycoproteins. Moreover, the multivariable analysis of the glycan profiling of the antigen positive exosomes indicated that each exosome fraction showed specific lectin signal patterns. The results suggest that the surface glycan structures vary depending on the exosome subpopulations.

Conclusion: In this study, we succeeded in establishing the protocol for rapid glycan profiling on the surface of small particles such as exosomes. This method is expected to be useful for the discovery of tumor-associated exosomal glycosylation.

716

Proteomic characterisation of treatment resistance in FLT3/MLL mutant paediatric acute myeloid leukaemia.

Tabitha E Mclachlan[1], Heather C Murray[1], David A Skerrett-Byrne[1], Ryan Duchatel[1], Océane Dubois[2], Julia Rossi[2], Anoop Enjeti[3], Frank Alvaro[4], Nicole M Verrills[1], Matthew D Dun[1]

1. University of Newcastle, Redhead, NEW SOUTH WALES, Australia

2. University Institute of Technology La Rochelle, La Rochelle, France

3. Calvary Mater Hospital, Newcastle , NSW, Australia

4. John Hunter Children's Hospital, Newcastle, NSW, Australia

Paediatric acute myeloid leukaemia (pAML) is the second most common form of leukaemia in children however, is responsible for the most leukaemia-associated deaths. The majority of children achieve an initial chemotherapeutic response however, 20% of patients will relapse. Receptor tyrosine kinase FLT3 is mutated in 20% of cases and is associated with increased likelihood of forming drug resistance. Histone lysine methyltransferase MLL is mutated in 10% of cases and is commonly co-mutated with FLT3. Cytarabine in combination with anthracycline based drugs are the long-standing treatment for pAML. However, multiple new therapies targeting FLT3 in relapsed pAML have been trialled but unfortunately have not increased overall survival. To identify more effective treatment targets, the mechanisms underpinning the development of resistance to cytarabine and anthracycline based chemotherapies requires characterisation.

We hypothesised, that analysis of protein epigenetic modifications, coupled with signalling pathways analysis would reveal mechanisms underpinning FLT3/MLL mutant pAML survival in cytotoxic conditions. A drug resistant FLT3/MLL mutant MV411 subline was created by culture in increasing concentrations of cytarabine and daunorubicin. Unbiased global analysis of the proteome of parental and drug resistant MV411 lines via label-free Data Dependent Acquisition (DDA) LC-MS/MS identified significantly increased activity of known and novel proteins associated with leukaemogenesis and therapy resistance such as IDH1,2 and 3A. The resistant cells also demonstrated significant changes in histone posttranslational modifications particularly methylation and acetylation, suggesting that altered epigenetic regulation is underpinning protein expression changes driving resistance.

Furthermore, ingenuity pathway analysis of differentially expressed proteins demonstrates the changes in cellular metabolism that occur in therapy-resistant AML cells. To validate these findings western blotting was used.

Our preliminary data, provides important new information on the cooperative mechanisms that underpin resistance to standard of care chemotherapies, providing us with a framework to investigate novel therapeutics to improve the treatment of children with AML.

717

Global analysis reveals the majority of deamidated HLA-bound peptides arise from deglycosylation via ERAD pathway

Shutao Mei[1], Nathan Croft[1], Jiangning Song[1], Anthony Purcell[1]

1. Monash University, Biomedicine Discovery Institute, Clayton, VIC, Australia

Publish consent withheld

718

End-to-end integration of known variants and modifications from PEFF into the Trans-Proteomic Pipeline for enriched MS/MS sequence determination

Luis Mendoza[1], Eric W Deutsch[1], Jimmy K Eng[2], Robert L Moritz[1]

1. Institute for Systems Biology, Seattle, WA, United States

2. University of Washington, Seattle, WA, USA

Background

The Proteomics Standards Initiative (PSI) Extended File Format (PEFF) allows for the specification of known mutations, post-translational modifications (PTMs), and other processing events of a given proteome in a unified, consistent format.

The Trans-Proteomic Pipeline (TPP) is a widely used and well-validated open source suite of software tools that facilitates and standardizes proteomics analysis. We describe recent enhancements and additions to TPP that enable full analysis, from raw file to the export of validated results and visualization, taking advantage of the rich information contained in PEFF.

Methods

TPP includes the latest version of the Comet search engine, which supports simple amino acid variants and mass modifications specified in PEFF files. The pepXML format was extended to allow it to represent search results that incorporate non-canonical sequence variants and known mass modifications. A new peptide-sequence-to-protein mapping mechanism was incorporated to exhaustively map detected peptide sequences to all possible protein variations. Various results and sequence viewers and interfaces have been updated to display and explore these results, with links to the relevant knowledge sources for further user verification.

Results

Naturally-occurring mutations resulting in sequence differences are present in most organisms, yet the majority of protein identification is made against a single canonical database of sequences that contain little knowledge of such variants. A further complication arises when post-translational modifications that are not specified in the search parameters are present in the sample, leading to incorrect search results for those spectra.

By automatically incorporating validated sequence variants and PTMs, a larger share of high-quality spectra are confidently identified, increasing sensitivity while decreasing the false discovery error rate. For instance, one of the highest-scoring spectra that was assigned to a decoy sequence in PeptideAtlas was found to map to a well characterized protein via a SAAV that has been observed in multiple experiments.

1. http://psidev.info/peff
2. http://www.tppms.org/
3. http://comet-ms.sourceforge.net/
4. http://www.tppms.org/tools/pm/

719

Pioneering translational aquaculture using proteomics

Omar Mendoza Porras[1], James Broadbent[1], Cedric Simon[1], Michelle Colgrave[1], Nick Wade[1]

1. CSIRO, St Lucia, QLD, Australia

Background: Global aquaculture production of black tiger shrimp *Penaeus monodon* was valued at US$4.9 billion in 2016. Blood chemistry often mirrors health status in living organisms, but highly abundant proteins often preclude identification or measurement of important markers. In shrimp, haemolymph is a blood-like fluid that could be used to identify health markers, but the multimeric and multi-functional protein haemocyanin (Hc) comprises 95% of the total protein content. Hc was initially identified as a copper-based oxygen transport protein, but roles in immunity and as a growth performance indicator in shrimp aquaculture have been recently reported. A large number of Hc isoforms have been previously reported; however, functional analysis of specific Hc isoforms are non-existent. **Methodology**: Transcriptomics-informed proteomics was used to resolve the complexity of Hc isoforms. Multiple reaction monitoring mass spectrometry (MRM-MS) was used to assess the suitability of six Hc isoforms as health and performance markers in 5 µl of haemolymph plasma from shrimp fed specific diets, or unfed. **Findings**: A total of 21 Hc isoforms were identified using proteomics, wherein six isoforms bore unique tryptic peptides suitable for MRM-MS. Shrimp weight gain and haemolymph total protein concentration were significantly higher in fed shrimp. Principal component analysis of MRM-MS Hc abundance clearly discriminated between fed and unfed shrimp. The abundance of the gamma isoform PmoHcG7 was significantly higher in the unfed group whilst PmoHcG2 was significantly higher in shrimp fed diets containing 10% microbial biomass and a control basal diet. **Conclusion**: Hc is a protein involved in shrimp health and growth that warrants investigation as a potential marker of shrimp performance. The application of proteomics in aquaculture can enable feed companies to test

functional diet formulations, and as a result, farmers will have the opportunity for better growth rates and optimised health in aquaculture shrimp whilst maintaining sustainability.

Keywords: Haemolymph, Haemocyanin, Shrimp, Aquaculture, proteomics

720

CDDO-Im, a Nrf2 inducer targets multiple amino acid residues on proteins via novel chemical mechanisms.

Xiaoli Meng[1], James C Waddington[1], B. Kevin Park[1], Michael B Sporn[2]

1. MRC Centre for Drug Safety Science, University of Liverpool, Liverpool, UK

2. Triterpenoids Therapeutics Inc., Lebanon, New Hampshire, USA

Background: The transcription factor Nrf2 plays an important role in the cellular defensive machinery against oxidative stress and inflammation and is a potential therapeutic target in human disease. The synthetic triterpenoids including CDDO and its methyl (CDDO-Me, bardoxolone methyl) and imidazolide (CDDO-Im) derivatives have been shown to enhance Nrf2-mediated antioxidant activity in a number of diseases. However, how triterpenoids activate Nrf2 is poorly understood. It has been suggested that triterpenoids are capable of reacting with thiols on Keap1, leading to disruption of the Keap1-Nrf2 association and upregulation of an Nrf2 response. Unlike CDDO-Me, CDDO-Im has an extra reactive site which could covalently bind to other amino acid residues than cysteine. This study aims to investigate how CDDO-Im covalently binds to proteins and the potential intracellular protein targets that may be vital for the molecular pharmacological actions of CDDO-Im.

Methods & results: Here, model proteins such as HSA and GSTP were used to determine the potential chemical reactivity of CDDO-Im towards proteins. Mass spectrometric analysis revealed that CDDO-Im, similar to CDDO-Me, formed covalent adduct with cysteine residues on HSA and GSTP through Michael addition. More importantly, for the first time, we have shown that CDDO-Im can covalently modify lysine, arginine, serine, and tyrosine residues on HSA via transacylation mechanisms. Cross-linking adducts derived from acylation and Michael addition to an adjacent cysteine residue and two different tryptic peptides were also detected.

Conclusion: This study provided new insights into the chemical mechanisms of action of CDDO-Im and paved the way for exploring potential novel targets for CDDO-Im.

721

Illuminating the Druggable Genome: an NIH Common Fund Program

Mehdi Mesri[1], Christine Colvis[1], Aaron Pawlyk[1]

1. National Institutes of Health (NIH), Bethesda, MD, United States

Existing clinical drugs only target a few hundred of the ~3000 “druggable genome,”, leaving a large subset within this druggable genome that remain largely understudied. Three protein families, the ion channels, G-protein-coupled receptors and protein kinases, have been identified to contain adequate numbers (~400 proteins) of understudied members and are well-established druggable families with high potential to impact human health.

To improve our scientific understanding of understudied members of these three protein families, the National Institutes of Health (NIH) Common Fund launched the Illuminating the Druggable Genome (IDG) Program in 2014. The Pilot Phase of the Program (2014-2017) established scalable technology platforms, and developed a website, Pharos (https://pharos.nih.gov/idg/index), that aggregates protein information from several sources, allowing researchers easy access to in depth protein data. Now in its Implementation Phase (2017-2024), the IDG Program aims to build on the knowledge and tools developed during the Pilot Phase and to generate, aggregate, analyze, and disseminate knowledge and tools around understudied proteins.

The kinase team, a working group of six laboratories, have constructed a heavy amino acid-labeled peptide library encoding tryptic fragments of dark kinases to quantitatively monitor their abundance in normal and perturbed cells using parallel reaction monitoring with stable isotope dilution (PRM-SID). The PRM-mass spec analysis is done in parallel with RNASeq, and data analyzed using network inference tools to provide insight into dark kinase integration in diverse cell types in response to genetic and pharmacological perturbations. The experimental work and reagents developed by the Kinase team is maintained in a comprehensive manner in the Dark Kinase Knowledgebase (DKK) that can be accessed at https://darkkinome.org/.

These new knowledge and tool sets are intended to equip the scientific community, including small businesses and the pharmaceutical industry, with the ability to explore previously understudied biology with the potential to rapidly impact human health.

722

Standardizing the performance of SWATH-MS software tools for label-free quantification using public repository spectral ion libraries

Mukul K. Midha[1], David S. Campbell[1], David D. Shteynberg[1], Ulrike Kusebauch[1], Michael R. Hoopmann[1], Guenther Kahlert[1], Robert L. Moritz[1]

1. ISB, Seattle, WASHINGTON, United States

The ISB produces a public repository for spectral ion libraries called SWATHAtlas (www.swathatlas.org) and the libraries available have been typically used to analyze the SWATH type mass spectrometry experiments using data-independent acquisition (DIA) methods. For confidence in protein identification and reporting quantitative results, the characteristics of a reference ion library are important when processing SWATH data. In this study, we provide an extensive demonstration of the effect of different libraries onto the performance of SWATH analysis tools and reported results that incorporate different chromatography flow rates,

different mass spectrometry instruments and the most widely used SWATH analysis software tools. For consistent assessment of reference ion libraries, we developed DIA-Lib QC, a library assessment tool to calculate the metrics of library correctness and its completeness, which ensures accurate library-based SWATH analysis. With our benchmarking hybrid datasets, we evaluated the characteristics and performance of different SWATH-MS tools such as OpenSWATH, PeakView, Skyline, Spectronaut, DIA-Umpire, and DISCO based on the accuracy of identification, the precision of quantitation, quantitative sensitivity, and relative performance of these applications. Conclusions include tools perform with higher fidelity with targeted reference libraries, library-based tools provide higher agreement in overall identifications with targeted libraries, reliable quantification performance, and robustness for label-free quantitative proteomics.

723

Towards Turnkey Targeted Proteomics Solutions Using Internal Standard Triggered Acquisitions on Next Generation Orbitrap Mass Spectrometers

Sebastien Gallien[1, 2], Aaron S Gajadhar[3], Bhavin Patel[4], Markus Kellmann[5], Tabiwang N Arrey[5], Christian Thoeing[5], Alexander Harder[5], Romain Huguet[3], Graeme McAlister[3], Derek Bailey[3], Shannon Eliuk[3], Emily I Chen[2], Yue Xuan[5], Andreas Huhmer[3], Ken Miller[3]

1. Thermo Fisher Scientific, Paris, France

2. Thermo Fisher Scientific, Precision Medicine Science Center, Cambridge, MA, USA

3. Thermo Fisher Scientific, San Jose, CA, USA

4. Thermo Fisher Scientific, Rockford, IL, USA

5. Thermo Fisher Scientific, Bremen, Germany

Targeted quantitative proteomics based on high resolution parallel reaction monitoring (PRM) technique benefits from improved measurement selectivity, allowing more sensitive endogenous peptides quantification in complex samples. Here we introduce an extension of PRM, called “SureQuant” method, which uses spiked-in internal standards (IS) to dynamically control the acquisition process and to maximize its productivity. This novel method has been implemented in the native instrument control software of next-generation Orbitrap instruments, Thermo Scientific™ Orbitrap Exploris™ 480 and Eclipse™ Tribrid™ mass spectrometers, to enable a broad access. The SureQuant IS targeted protein quantification method has been adapted from the IS-PRM method in order to improve its usability and robustness (especially against chromatographic variations). Its ability to deliver high-density, ultra-sensitive measurements has benefited to a variety of applications. Applied to the monitoring of signaling pathways in cell lines and tissues specimens supplemented with stable isotopically labeled (SIL) peptides (30-150 IS, including Thermo Scientific™ Pierce™ SureQuant™ kits), the method enabled systematic quantification of endogenous peptides with high precision (<5%-CV for the majority of peptides) and short analysis time (10-40 min LC gradient). In a larger scale application of the method, non-depleted plasma samples supplemented with 804 SIL peptides (Biognosys™ PQ500™ kit) were analyzed with a 70-min LC gradient for global plasma proteome quantification. More than 550 endogenous peptides, surrogates of around 400 plasma proteins, were reproducibly quantified over a 6 orders of magnitude range. This proteome coverage compared favorably with that of profiling methods, while still benefiting from the enhanced data quality of targeted measurements (including peptide quantification in the low amol range). The multiple analytical benefits of the SureQuant method combined with the ability to embed pre-set (optimized) methods, associated with predefined kits of IS peptides, directly into the instrument control software represents a decisive step towards the provision of turnkey targeted proteomics solutions.

724

Simple and efficient HLAp-ligandomics by using high-field asymmetric waveform ion mobility spectrometry (FAIMS)

Yuriko MINEGISHI[1], Kensaku Nemoto[1], Kazuma Kiyotani[1], Koji Ueda[1]

1. Japanese Foundation for Cancer Research (JFCR), Koto-ku, TOKYO, Japan

Human leukocyte antigen (HLA) class 1-binding peptide (HLAp) with somatic mutation, recognized as a neoantigen, is as an optimal target for cancer vaccine treatment. Though identifying actually-presented neoantigens is crucial for realization of cancer precision medicine, comprehensive analysis of neoantigens (HLAp-ligandomics) remains difficult due to insufficient HLAp isolation yield, undefined mass spectrometric parameters for non-tryptic digest peptides, and lack of bioinformatics pipelines to automatically construct “individual” proteome database from genetic sequence dataset.

To overcome these difficulties, we optimized immuno-capturing procedures for HLAp isolation, employed high-field asymmetric waveform ion mobility spectrometry (FAIMS) for maximization of HLAp identification efficiency, and finally established a pipeline to construct proteome databases seamlessly from whole exome sequencing analysis. Particularly, FAIMS technology enables unique ion fractionation in vapor phase based on ion mobilities under different electric fields, allowing more sensitive and time-saving peptide identification in a single LC/MS run.

The data was acquired with Orbitrap Fusion Lumos mass spectrometer installed with or without FAIMS Pro (Thermo Fisher Scientific). The compensation voltages (CVs) were set to -40/-60/-80 for targeting charge states 2 & 3, while they were set to -10/-20/-30 for targeting charge state 1. The HCT116-specific mutant proteome (mutanome) database was established by the in-house pipeline. As a result, the number of identified HLAp from 1.0E+7 HCT-116 cells (1,443 peptide groups) surpassed around twice as much as that without FAIMS (702 peptide groups). Majority of the identified HLAp (86.2%) had 8-10 amino acid length. Gibbs Cluster 2.0 revealed the dominance of A*01:01 and B*45:01-assigned peptides in HCT116. Importantly, a couple of neoantigens were successfully identified from the database search analysis using HCT116-specific mutanome.

In the future, if high-performance HLAp profiling technology could be used in routine clinical practice, application of personalized cancer immunotherapy would be further accelerated.

725

Can retina as part of the CNS unfold the common pathophysiology underlying the neurodegenerative diseases?

Mehdi Mirzaei[1, 2], Liting Deng[2], Kanishka Pushpitha[3], Yunqi Wu[1, 2], Dana Pascovici[1, 2], Jemma Wu[1, 2], Matthew J McKay[1, 2], Paul A Haynes[2], stuart L Graham[3], Vivek Gupta[3]

1. Australian Proteome Analysis Facility, Macquarie University, Sydney, NSW, Australia

2. Molecular Science , Macquarie University, Sydney, NSW, Australia

3. Clinical Medicine , Macquarie University, Sydney, NSW, Australia

The concept of "the eye as a window to the brain" is increasingly being explored to examine brain function in areas of health and disease. During embryonic development, the retina and optic nerve project from the brain, and are thus considered part of the central nervous system (CNS). Retinal tissue demonstrates similarities with the brain in terms of neural anatomy and cellular and biochemical responses to injury. In fact, similar to the blood-brain barrier, retinal tissue maintains a distinct immune environment that is facilitated by the blood-retinal barrier.

Retina is increasingly being investigated to assess the effects of brain related pathological conditions such as Alzheimer's disease (AD). These molecular changes are reported to exhibit pathological similarities with retinal neurodegenerative disorders including glaucoma. AD is now understood to be associated with glaucoma with higher rates of incidence of glaucoma being reported in some cohorts of Alzheimer's patients. The exact molecular basis of this association remains poorly defined.

The study aims to decipher the common molecular mechanisms underlying retinal pathology in response to various neurodegenerative disorders affecting the retina, particularly in glaucoma and AD. To achieve this, we employed quantitative proteomics and biochemical approaches coupled with bioinformatics tools to analyse retinal glaucoma tissue from human post-mortem subjects and an animal model of experimental glaucoma. The findings were correlated with data obtained from retinal and brain tissues from APP/PS1 double transgenic mutant mouse model of AD. Our results reveal several novel candidate markers and biochemical pathways that were commonly affected in the two diseases. Hallmark proteins of AD such as amyloid beta and tau were identified to be modulated in these two distinct neural tissues of the CNS. This study provides mechanistic insights into the common pathophysiology underlying neurodegenerative diseases in the retina and brain.

726

Monitoring of *N*-linked Glycans in Endometrial Cancer

Parul Mittal[1], Manuela Klingler-Hoffmann[2], Matthew Briggs[2], Martin Oehler[3], Peter Hoffmann[2]

1. Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adealaide, SA, Australia

2. University of South Australia, Mawson Lakes, SA, Australia

3. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

Protein glycosylation is a posttranslational modification, which can result in functional changes of the glycoproteins and can play a key role in cancer progression and treatment. Here, we have analysed formalin fixed paraffin embedded (FFPE) endometrial cancer and adjacent tissues to explore tissue specific N-linked glycan abundance using glycan matrix assisted laser desorption ionisation mass spectrometry imaging (MALDI MSI). We have used single sections and tissue microarrays (TMA) in our analysis. Identification of discriminative *m/z* values was achieved using porous graphitized carbon liquid chromatography (PGC-LC) and collision induced electrospray negative mode MS fragmentation analysis (ESI-MS/MS). High mannose glycan structures were predominately detected in the tumour regions, while complex bi- and tri-antennary structures were observed in the normal/adjacent regions, indicating a biological role for these modifications. In summary, tumour and normal regions were clearly distinguished based on their *N*-glycan distributions.

727

High-throughput discovery of functional protein modifications by hotspot thermal profiling

Raymond Moellering[1]

1. University of Chicago, Chicago, IL, United States

Mass spectrometry has revolutionized the ability to study posttranslationally modified proteoforms from biologic samples, yet we still lack methods to systematically predict, or even prioritize, which modification sites may perturb protein function. This talk will describe a novel proteomic method to detect the effects of site-specific protein phosphorylation on the thermal stability of thousands of native proteins in live cells. This massively parallel biophysical assay revealed surprising shifts in overall protein stability in response to site-specific phosphorylation sites, as well as trends related to protein function and structure. This method, hotspot thermal profiling (HTP), detects both intrinsic changes to protein structure as well as extrinsic changes to protein-protein, and protein-metabolite interactions resulting from the diminutive introduction of a phosphate onto large proteins. Finally, I will discuss how functional "hotspot" protein modification sites can be discovered and prioritized for study in a high-throughput and unbiased fashion, and characterized kinetically in response to k cellular signaling events. This approach should be applicable to diverse organisms, cell types and posttranslational modifications.

728

Intrinsic reactivity in cell signaling: A feature not a flaw of metabolic regulation

Raymond Moellering[1]

1. University of Chicago, Chicago, IL, United States

Prevailing models on the role(s) of inherently reactive cellular metabolites suggest that these molecules are limited to generating non-specific biomolecular damage in cells. Work from our lab and others has shown that the posttranslational modifications formed from these metabolites are more restricted than expected, and that functional modification sites are enriched and conserved within the proteome. This raises a fundamental question: Do intrinsically reactive metabolites and their non-enzymatic modifications on proteins comprise conserved signaling roles in cellular signaling? In this talk I will discuss our recent discovery of a new intracellular communication mechanism integrating glycolysis with the KEAP1-NRF2 pathway through a novel reactive metabolite-induced posttranslational modification. Pharmacologic regulation of this intrinsic feedback loop with a new small molecule inhibitor of central glycolysis enacts cytoprotective cellular responses in NRF2-responsive cellular and *in vivo*models of disease. Implications for the role of reactive metabolites, as well as their pharmacologic manipulation, in cellular signaling and disease will be discussed.

729

Multiplexed glycan profiling of extracellular vesicles using lectin-surface enhanced raman spectroscopy (lectin-SERS)

Edward S.X. Moh[1], **Wei Zhang**[1], **Mingkai Wu**[1], **Yuling Wang**[1], **Nicolle H Packer**[1]

1. ARC Centre of Nanoscale BioPhotonics, Macquarie University, Sydney, NSW, Australia

Aberrant cellular glycosylation has been a reported feature of multiple disease states with high-resolution separation and mass spectrometric methodologies being key to these discoveries. However, clinical translation of these discoveries is hard to implement due to reliance on instrumentation and expert data interpretation. The only glycan related clinical example to-date uses the lectin, *Lens culinaris* agglutinin, to detect fucosylated alpha-fetoprotein-L3 as a marker for hepatocellular carcinoma progression. Differential lectin binding to specific glycan structures can simplify the complexity of detailed mass spectrometric profiling and aid in the clinical application of glyco-oriented disease monitoring and diagnosis.

Multi-lectin profiling provides an advantage, as differences between the glycome of different diseased cells can be subtle and require a panel of lectins for differentiation. Multiplexed fluorescence detection of lectins is limited by the overlapping spectrometric properties of fluorophores, and existing lectin arrays cannot be considered true multiplexing but more about testing lectin binding individually in a standardised or high throughput fashion. Surface enhanced Raman spectroscopy (SERS) using reporter coated gold nanoparticles provides an effective way for multiplexing, as a single excitation wavelength can read out multiple unique Raman reporter signals from different SERS particles coated with diverse targeting probes.

In this work, we developed multiple lectin conjugated SERS-particles and demonstrate the potential of multiplexed lectin profiling of CD63+ extracellular vesicles *in vitro* and in body fluids, with detection using a handheld Raman reader. We show that enzymatic desialylation of the vesicles reduces the WGA-SERS signal that corresponds to sialic acid binding, with a concurrent increase in PHA-L-SERS signal corresponding to an increase in the exposed galactose terminal sugars; validating that differences in vesicle glycosylation can be detected using this method. Further optimisation and testing of this method will lead to rapid profiling of clinical samples as a pre-screening method for more detailed diagnostic tests.

73

Food safety assessment in genetically engineered canola – LC-MS/MS as an alternative to antibody-based approaches

Michelle Colgrave[1], **Keren Byrne**[1], **Joanne Caine**[2], **Lukasz Kowalczyk**[2], **Sapna Vibhakaran Pillai**[3], **Bei Dong**[3], **Susan MacIntosh**[4], **Judith Scoble**[2], **James Petrie**[3], **Surinder Singh**[3], **Xue-Rong Zhou**[3]

1. Agriculture & Food, CSIRO, St Lucia, QLD, Australia

2. Manufacturing, CSIRO, Parkville, VIC, Australia

3. Agriculture & Food, CSIRO, Black Mountain, ACT, Australia

4. NuSeed Americas, Alsip, IL, USA

Omega-3 oils are highly sought after for their human health benefits such as cardiovascular protection and improved cognitive development and function. Originally sourced from fish, growing concerns with food security, fish stocks, and industrial fishing have led to increased interest in alternative sources of omega-3 oils. Using gene technology, the ability to produce omega-3 oils from lower plants was transferred into canola, by introducing seven transgenes, representing the most complex piece of metabolic engineering so far achieved in plants.

For genetically modified products, food/feed risk assessment involves evaluation of protein stability and plant expression levels. Protein digestibility is considered a determinant of potential allergenicity that has been traditionally measured using antibodies. Yet membrane proteins are extremely difficult to express in heterologous systems making raising antibodies impracticable. In this study a novel mass spectrometric approach was developed that enabled the investigation of the *in vitro* digestibility of the seven transmembrane proteins. The proteins and their peptide fragments were identified by high resolution LC-MS/MS and subsequently LC-MRM-MS was employed for specific peptide quantitation.

A two-stage digestion strategy involving simulated gastric fluid followed by trypsin enabled the measurement of protein digestibility *in vitro*. Tryptic peptide markers spanning the length of each desaturase protein were monitored and showed that all regions were readily degraded (>95% within 5 min) and highlighted regions of the elongase enzymes showing limited resistance to gastric digestion. Traditional gel-based and Western blotting analysis of ω3-desaturase and Δ6-elongase revealed rapid protein

hydrolysis and no fragments (>3 kDa) remaining after 60 minutes, complementing the novel approach. The LC-MS approach was sensitive, selective and did not require the use of purified proteins and is widely applicable to food/feed and environmental safety assessment.

LC-MS innovations in food safety assessment have facilitated the deregulation of a sustainable plant resource capable of producing omega-3 oils.

730

Potential early clinical stage colorectal cancer diagnosis using a proteomics blood test panel

Abidali Mohamedali[1], Seong Beom Ahn[2], Samridhi Sharma[2], Sadia Mahboob[2], William J Redmond[2], Dana Pascovic[3], Jemma X Wu[3], Thiri Zaw[3], Subash Adhikari[2], Vineet Viabhav[2], Ed C Nice[4], Mark S. S Baker[5]

1. Molecular Sciences, Faculty Of Science and Engineering, Macquarie University, Sydney

2. Biological Sciences, Faculty of Medicine and Health Sciences , Macquarie University, Sydney

3. Australian Proteome Analysis Facility, Macquarie University, Sydney

4. Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC

5. Biomedical Sciences, Faculty of Medicine and Health Sciences, Macquarie University, Sydney, NSW, Australia

Background: One of the most significant challenges in colorectal cancer (CRC) management is the use of compliant early stage population-based diagnostic tests as adjuncts to confirmatory colonoscopy. Despite the near curative nature of early clinical stage surgical resection, mortality remains unacceptably high - as most patients diagnosed by faecal haemoglobin followed by colonoscopy occur at latter stages. Additionally, current population-based screens reliant on fecal occult blood (FOBT) have low compliance (~40%), and tests suffer low sensitivities. Therefore, blood-based diagnostic tests offer survival benefits from their higher compliance (>97%), if they can at least match the sensitivity and specificity of FOBTs. However, the discovery of low abundance plasma biomarkers is difficult due to high abundance plasma proteins.

Methods: A combination of ultradepletion (e.g., MARS-14 and an in-house IgY depletion columns) strategies, extensive peptide fractionation methods (SCX, SAX, High pH and SEC) and SWATH-MS were utilized to uncover protein biomarkers from a cohort of 100 plasmas (i.e., pools of 20 healthy and 20 stages I-IV CRC plasmas). The differentially expressed proteins were analyzed using ANOVA and pairwise t-tests ($p<0.05$; fold-change>1.5), and further examined with a neural network classification method using *in silico* augmented 5,000 patient datasets.

Results: Ultradepletion combined with peptide fractionation allowed for the identification of a total of 513 plasma proteins, 8 of which had not been previously reported in human plasma (based on PeptideAtlas database). SWATH-MS analysis revealed 37 protein biomarker candidates that exhibited differential expression across CRC stages compared to healthy controls. Of those, seven candidates (CST3, GPX3, CFD, MRC1, COMP, PON1 and ADAMDEC1) were validated using Western blotting and/or ELISA. The neural network classification narrowed down candidate biomarkers to 5 proteins (SAA2, APCS, APOA4, F2 and AMBP) that maintained accuracy that could discern early (I/II) from late (III/IV) stage CRC.

731

MHC peptidome analysis to identify and characterize cancer neo-antigens

Sonali Mohan[1, 2], Keshava K. Datta[1], Harsha Gowda[1, 3, 2]

1. Cancer Precision Medicine Group, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

2. Faculty of Medicine, The University of Queensland, Herston, Brisbane, QLD, Australia

3. School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia

Advent of immunotherapies has revolutionized cancer treatment. Recent success with immunotherapy is predominantly due to checkpoint inhibitors that block inhibitory signals and enable T cell activation that target cancer cells. Other strategies including adoptive cell transfer and cancer vaccines are being investigated in parallel to increase available arsenal for immune therapy. Cancer genome sequencing has revealed thousands of somatic mutations across various cancers. It is known that some proteins encoded by mutated genes are processed and presented on the cell surface. These MHC presented mutant peptides serve as neo-antigens that are recognized by T cells. Identification of such neo-antigens can strengthen cancer immunotherapy efforts and provide a set of antigens that can be targeted. Currently, various prediction programs are employed to predict potential neo-antigens based on cancer genome sequencing data. However, this approach can result in significant number of false positives and false negatives. Recent studies have demonstrated mass spectrometry based unbiased approaches to identify MHC bound peptides. We have combined whole-exome sequencing analysis with MHC peptidome mass spectrometry to identify potential neo-antigens from melanoma cell lines. By analyzing our MHC peptidome dataset and publicly available datasets, we have identified sequence features and other rules that potentially determine MHC presentation of peptides. These observations can prove useful for developing better experimental strategies and prediction tools to identify potential cancer neo-antigens. Reliable identification of cancer neo-antigens can accelerate development of novel therapeutic approaches that can exploit host immune system to treat cancers.

732

Identification of serum-based biomarkers for the onset of oral cancer among tobacco users

Varshasnata Mohanty[1], Riaz Abdulla[2], Santosh K. Behera[1], Sneha M. Pinto[1], Samapika Routray[3], Neeta Mohanty[4], Bipin Nair[5], T. S. Keshava Prasad[1], Mandakulutur S. Ganesh[6], Arnab Pal[7], Jay G. Ray[8], Yashwanth Subbannayya[1], Harsha Gowda[10, 9, 1], Aditi Chatterjee[10, 9, 1]

1. Center for Systems Biology and Molecular Medicine, Yenepoya Research Centre, Yenepoya (Deemed to be University), Mangalore, Karnataka, India

2. Department of Oral Pathology, Yenepoya Dental College, Yenepoya (Deemed to be University), Mangalore, Karnataka, India

3. Department of Dental Surgery, All India Institute of Medical Sciences, Bhubaneswar, Odisha, India

4. Department of Oral Pathology & Microbiology, Institute of Dental Sciences, Siksha'O'Anusandhan University, Bhubaneswar, Odisha, India

5. Amrita School of Biotechnology, Amrita University, Kollam, Kerala, India

6. Department of Surgical Oncology, Vydehi Institute of Oncology, Bangalore, Karnataka, India

7. Department of Biochemistry, Postgraduate Institute of Medical Education and Research, Chandigarh, Punjab, India

8. Department of Oral Pathology, Burdwan Dental College and Hospital, Burdwan, West Bengal, India

9. Institute of Bioinformatics, Bangalore, Karnataka, India

10. Manipal Academy of Higher Education, Manipal, Karnataka, India

Oral squamous cell carcinoma (OSCC) is one of the leading causes of cancer-related deaths worldwide and has no effective screening strategy. It arises as a primary lesion originating in any of the oral tissues, or as an extension from a neighboring anatomic structure, and can also arise by metastasis from a distant site of origin. Use of tobacco is one of the predominant risk factors for the development of oral cancer. Currently, there are no markers identifying individuals at a higher risk of developing the disease. Therefore, successful identification and translation of candidate molecules could aid in clinical knowledge.

Recent developments in quantitative proteomics have enabled the discovery of blood-based markers for various cancers. In this study, we employed a tandem mass tag (TMT)-based quantitative proteomics approach to study alterations in serum proteomes of oral cancer patients with tobacco habit as compared to healthy individuals. The analysis on Fourier transform Orbitrap Fusion Tribrid mass spectrometer resulted in the identification of ~1200 proteins of which ~200 proteins found to be dysregulated (p-value ≤ 0.05). Enrichment analysis revealed significant perturbation of signaling pathways pertaining to angiogenesis, metabolic process, immune response and regulation of MAPK cascade pathways. Proteins mediating these pathways including ADA2, CLEC3B, PDLIM1, PGLYRP2, and NKPD1 were significantly altered in OSCC patients who are tobacco users.

In summary, the current study identified several serum-based proteins, of which ADA2, SERPINA6, and SERPINF1 are few key molecules altered in OSCC patients with tobacco use and may be potential biomarkers to diagnose and monitor the disease. This needs to be validated in larger cohorts to help identify individuals who are at higher risk of developing oral cancer.

733

Quantitative proteomics reveals docosahexaenoic acid-mediated neuroprotective effects in lipopolysaccharide-stimulated microglial cells

Bo Yang[1], Brian P Mooney[1], Michael Greenlief[1], Runting Li[1], Xue Geng[2], Kevin L Fritsche[1], David Q Beversdorf[1], James C Lee[2], Grace Y Sun[1]

1. University of Missouri, Columbia, Columbia, MISSOURI, United States

2. University of Illinois at Chicago, Chicago, IL, USA

The high level of docosahexaenoic acid (DHA) in phospholipids in the brain has spurred significant interest in its role in brain health and diseases. Recent studies provided evidence for DHA to play a role in inhibiting inflammatory responses in microglial cells, although the mechanism(s) remains elusive. In this study, a global proteomic approach was used to examine effects of DHA on microglial cells stimulated with lipopolysaccharides (LPS). Deep proteome coverage was achieved using the parallel accumulation serial fragmentation (PASEF) method in a hybrid trapped ion mobility spectrometry (TIMS) – quadrupole time-of-flight mass spectrometer (TIMS–ToF). Using label-free quantitative proteomics, a total of 2858 protein groups (with more than one unique peptide) were confidently identified and quantified in BV-2 microglial cells. Treating the cells with LPS and/or DHA altered cell morphology and expression of 43 proteins with a differential abundance (greater than 2.0-fold change). Bioinformatic analyses further indicated that these differentially abundant proteins were involved in the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) pathway, fatty acid metabolism, mitochondrial activity, response to bacterium, cytoskeleton, DNA binding, and ribosome biogenesis. Quantitative analyses of cell viability, tumor necrosis factor alpha, phospho-NF-κB p65, inducible nitric oxide synthase and prostaglandin E2 expression levels were consistent with the biological outcomes of the altered protein concentrations. Together, these data indicate for the first time multiple mechanisms of DHA-mediated protective effects in LPS-stimulated BV-2 microglial cell at the global proteome level. This information may shed light on a new basis for therapeutic strategies of DHA against inflammation.

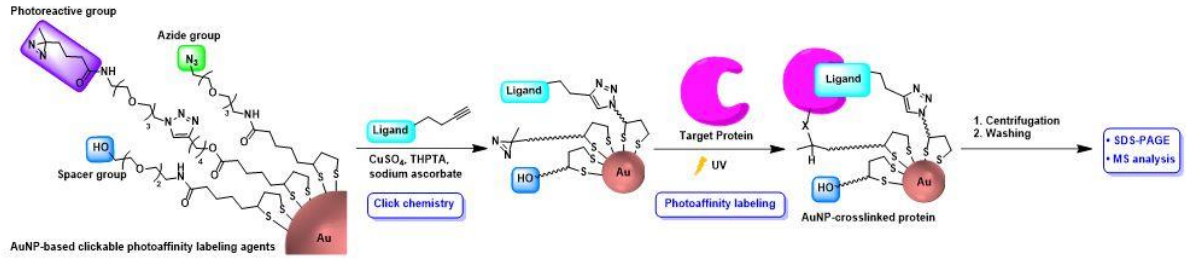
734

Development of gold nanoparticle-based clickable photoaffinity probes for target protein identification

Kanna Mori[1], Kaori Sakurai[1]

1. Tokyo University of Agriculture and Technology, Graduate School of Engineering, Koganei City, TOKYO, Japan

Target protein identification is the first critical step for the elucidation of the mechanism of action of bioactive small molecules and in the field of drug discovery. Photoaffinity labeling has been useful for target identification studies because it enables covalent crosslinking of target proteins in live cells. However, its routine application is hampered by the need to optimize probe design, high background labeling, and a multistep scheme. We have previously developed gold nanoparticle (AuNP)-based photoaffinity labeling probes and successfully improved probe synthesis and labeling efficiency by taking advantage of the multivalent effect and the unique property of AuNP. In this study, we devised new clickable photoaffinity probes based on AuNP to expand the scope of their utility. We designed and synthesized AuNP probes displaying an azide group, a photoreactive group, a hydrophilic spacer group. Any small-molecule ligand of interest with an alkyne tag can be immobilized on these clickable probes in one step by click chemistry. Using a bovine carbonic anhydrase (BCA Ⅱ) inhibitor with an alkyne tag, we demonstrated that both copper-promoted azide-alkyne cyclization and photoaffinity labeling can be performed at nanomolar concentrations of AuNP probes. The effectiveness of our new clickable photoaffinity probes were evaluated by SDS-PAGE following the thiol-exchange reaction and by direct application of MALDI-MS.



735

High-throughput single cell proteomics analysis with nanodroplet sample processing, multiplex TMT labeling, and ultra-sensitive LC-MS

Khatereh Motamedchaboki[1], Maowei Dou[1], Yongzheng Cong[2], Romain Huguet[1], Aaron M Robitaille[1], Yufeng Shen[3], Daniel Lopez-Ferrer[1], Ryan T Kelly[2, 4], Ying Zhu[4]

1. Thermo Fisher Scientific, San Jose, CALIFORNIA, United States

2. Brigham Young University, Provo, UT, USA

3. CoAnn Technologies LLC, Richland, WA, USA

4. Pacific Northwest National Laboratory, Richland, WA, USA

Proteomics analysis with large number of cells represents the average protein content of a heterogenous mixtures of cells. Understanding heterogeneity at single cell level is of great interest for biomedical research. MS-based proteomics is a promising technique for single cell analysis by enabling identification and quantification of thousands of proteins in unbiased manner. However, due to inefficient single cell isolation, large sample losses during sample preparation and low throughput, the extension to single cell studies has been largely ineffective. To address these challenges, we combined nanoPOTS (Nanodroplet Processing in One-pot for Trace Samples) technology with tandem mass tag (TMT™) isobaric labeling to efficiently process and analyze single mammalian cells containing <0.2 ng total proteins on new Orbitrap Eclipse™ Tribrid™ Mass Spectrometers with real time search and FAIMS Pro™ to improve single cell proteome coverage and enhance quantification accuracy. Single cells were isolated from cultured murine and HeLa cells via fluorescence-activated cell sorting and samples were processed on nanoPOTS chip. The UltiMate™ 3000 RSLCnano system was used with 20 to 30 µm i.d. columns coupled to Orbitrap Eclipse. Proteome Discoverer™ 2.4 software was used for data analysis. Both label-free and isobaric labelling (TMT10plex™)-based protein quantifications were evaluated with a focus on reproducibility in quantification and throughput. The label-free workflow resulted in identification of an average of 465 protein groups from single Hela cells and 802 protein groups from a large HeLa cell with high-confidence MS/MS spectra. The TMT10plex analysis of the three cultured murine cell populations (C10, SVEC and Raw cells) enabled identification of 2346 proteins and 1300 quantifiable among 40 single cells. We have demonstrated that single cell proteome can be quantified using label-free or TMT workflows by combing nanoPOTS with Orbitrap Eclipse Tribrid mass spectrometer, enabling researchers to investigate cell heterogeneity as well as rare cells.

736

Addressing throughput and detectability in targeted analysis of serum proteins

Robert Moulder[1], Santosh D Bhosale[1], Riitta Lahesmaa[1]

1. Turku Bioscience, University of Turku & Åbo Akademi University, Turku, Finland

Background: Throughout the proteomics era, the analysis of serum as a source of biomarkers for disease risk and status has retained interest. For subsequent validations, targeted proteomics provides a powerful alternative to antibody based methods. Nevertheless, the throughput of the methodology can encumber its scope for application with large cohorts.

Data from analyses aimed towards increasing the throughput of targeted serum protein measurement are presented. Selected Reaction Monitoring (SRM) and Parallel-Reaction Monitoring (PRM) with a conventional capillary LC (75 µm) and an EvoSep system were used.

Methods: Serum was prepared for proteomics without depletion. Synthetic analogues for proteotypic peptides, including the PQ500 panel (Biognosys) and retention time standards, were used to evaluate separation and scheduling methods in terms of targets detected and analysis time. For library creation and PRM, a Q Exactive Orbitrap HF mass spectrometer was used. A TSQ triple quadrupole MS was used For SRM measurements. An EasyNano and an EvoSep LC instrument were evaluated.

Endogenous peptides and isotopically labelled synthetic peptides were measured throughout to confirm targets and control peak integration, and two stable unique peptides per protein were measured where possible.

Results: SRM of multiple targets was restricted by the number of concurrent transitions. However, evaluating several peptides per protein and maximizing use of the chromatographic time window reduced this limitation. PRM provided selective detection for a range of common protein targets with sensitivity comparable to SRM and excellent selectivity.

Comparing the LC systems, the EvoSep system enabled time saving for the moderate sized panels of targets, reducing loading and equilibration times.

Conclusions: With SRM and PRM measurements, in the order of 120 protein targets was achieved whilst maintaining two or more peptides per protein and recording multiple transitions or PRM spectra. In particular, the use of peptides across the chromatographic scale was an important advantage.

737

The proteome landscape of the kingdoms of life

Johannes B. Mueller[1], Philipp E. Geyer[1, 2], Ana R. Colaço[2], Peter V. Treit[1], Sophia Doll[1], Sebastian V Winter[1], Jakob Bader[1], Niklas Köhler[3], Fabian Theis[3], Alberto Santos[2], Matthias Mann[1, 2]

1. Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Germany

2. Faculty of Health Sciences, University of Copenhagen, NNF Center for Protein Research, Copenhagen, Denmark

3. Institute of Computational Biology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

Improvements in sequencing technologies have led to vast amounts of genomic data, including full genome sequences of a large number of organisms. This opens up the possibility for parallel exploration of their proteomes; however, so far there has been no large-scale effort in this direction. Moreover, quantitative proteomics data exist almost exclusively for the most common model organisms. Here, we set out on a proteomics exploration by selecting 100 sequenced organisms from the entire tree of life. We analyzed their proteomes by state of the art label-free quantitation methods. To make our proteome data extendable to other research laboratories, we took advantage of a novel reversed-phase chromatographic column – the µPAC (PharmaFluidics). It has a lithographically etched pillar structure that does not suffer from typical drawbacks of bead-based columns, yielding a highly reproducible retention time of molecules between measurements and columns. With about 340,000 proteins from more than two million sequence unique tryptic peptides, our proteomic map of the tree of life represents by far the largest proteome dataset. It confirms the existence of a very large number of predicted proteins, increasing the total known to the research community by about 50%. This extensive dataset is suited for deep learning algorithms for prediction of technical as well as biological parameters, which we demonstrate on peptide retention times for LC-MS measurements. Quantitative comparison of protein expression and functional abundances across species in a graph database allows exploration of proteins, pathways and organelles across phyla. In particular, we visualize the conservation of protein groups that have distinct cellular functions by quantitative protein levels, which allows us to rank those functions within and between organisms. To make this data easily accessible, our graph database will be available via an interactive web site, which enables browsing of structural information, functional annotations and protein homologies.

738

In-depth proteomics at a single glomerulus level in human kidneys using MALDI Imaging Mass Spectrometry and shotgun proteomics on paraffin embedded biopsy tissue section

Yume Mukasa[1], Takashi Nirasawa[2], Ryo Kajita[2], Marion Rabant[3], Jean-Paul Duong Van Huyen[3], Patrick Bruneval[4], Hatsue Ishibashi-Ueda[5], Nobuto Kakuda[1], Masaya Ikegawa[1]

1. Department of Life and Medical Systems, Doshisha University, Kyotanabe, Kyoto, Japan

2. Bruker Japan K. K., Yokohama, Kanagawa, Japan

3. Necker-Enfants malads Hospital, Paris, France

4. Georges-Pambidou European Hospital, Paris, France

5. National Cerebral and Cardiovascular Center Research Institute, Suita, Osaka, Japan

The molecular characterization at the level of individual glomerulus in kidneys remains a technological challenge that needs to be addressed in order to better understand pathological mechanisms. Here, we developed and applied a mass spectrometry-based methodology to investigate heterogeneity of proteomes from in situ tissue sections from human biopsied samples with or without kidney disease.

In this study, we adopted paraffin-embedded tissue sections fixed in formalin-acetic acid-alcohol (FAA) fixative, using matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS). Furthermore, we established a workflow that integrates MALDI-IMS and shotgun analysis (timsTOF Pro system). FAA fixed paraffin embedded tissue samples were from Necker Children's Hospital. After HE staining, serial slide thickness of 10 µm was made from the same block of the sample on

ITO coated slide glass. Dewaxing, antigen retrieval, pH adjustment on tissue samples and on-tissue digestion with trypsin, deposition of α-cyano-4-hydroxycinnamic acid as a matrix using TM-Sprayer. MALDI-IMS was done by rapifleX with a spatial resolution of 20 or 50 *μ*m. Ions were detected in mass range of *m/z* 800 to 3000. Statistical analysis was performed with SCiLS Lab 2019 software. Shotgun Proteomics from serial sections of MALDI-IMS were attempted using timsTOF Pro with nanoElute system.

Here we have established an integrated workflow using MALDI-IMS method and shotgun proteomics analysis from FAA-fixed kidney biopsies from non-pathological cases. Of note, we have succeeded in obtaining IMS data with 20 *m*m resolution and was able to obtain MS ion images at a single glomerulus level from FAA-fixed renal biopsies. Thousands of proteins from the serial sections with shotgun proteomics were also annotated which were analyzed with Protein ID and Proteinscape ver 3.0. We will further apply this method to elucidate and validate its efficacy with a variety of renal diseases such as diabetic nephropathy and renal amyloidosis.

739

Global & targeted proteomics reveal major dysregulated networks in meningiomas

Shuvolina Mukherjee[1], Sanjeeva Srivastava[1], Deeptarup Biswas[1], Rucha Gadre[1], Pooja Jain[2], Nelofer Syed[3], Tzouliana Stylianou[3], Graham Ball[4], Epari Sridhar[5], Aliasgar Moiyadi[6], Prakash Shetty[6]

1. Indian Institute of Technology (IIT) Bombay, Mumbai, MAHARASHTRA, India

2. Centre for Integrative Systems Biology and Bioinformatics (CISBIO), Imperial College, London, Uk

3. Division of Brain Sciences, Department of Medicine, Imperial College, London, UK

4. School of Science and technology, Nottingham Trent University, Nottingham, UK

5. Department of Pathology, Tata Memorial Center, Mumbai, Maharashtra, India

6. Department of Neurosurgery, Tata Memorial Center, Mumbai, Maharashtra, India

Introduction: Meningiomas are tumors originating from the outer layering of the brain and comprise of nearly 30% of CNS tumors of the brain. In certain cases, the tumor location is unamenable for complete resection leading to comorbidities; however, there are very limited effective therapeutic strategies. The current study investigates the patient specific global proteome of meningiomas resulting in identification of 4639 proteins (1% FDR, ≥ 2 unique peptides) Investigation of proteomic alterations that corroborate with radiology and histopathology was also done. An in-depth analysis revealed upheaval of Focal adhesion, PI3-Akt pathways which were further probed using inhibitor against Integrin Linked Kinase, a major influencer for the key pathways. Materials & Methods: The study comprised of tissue extraction from surgically resected biospecimens of meningiomas. LC MS/MS analysis was performed using LFQ approach using Q-Exactive platform (Thermo Fisher Scientific) & Proteome Discoverer 2.2 was used for data analysis. In silico analysis was done using GO term-based enrichments & Machine learning approaches Conclusion: We have identified candidates that are differentially abundant based on the radiological location of the tumor like FBLN5 and NUP210. IHC-based WHO grade analysis suggested potentially new subtype in meningioma Grade-I. Additionally, in silico analysis points out prominent alterations in Focal Adhesion & NF kappa B pathways in meningioma patients. We observed perturbations in several of the pathway components via Global proteomic profiling in treated meningioma cell line (Ben-Men1) which substantiates potential of ILK inhibition as a therapeutic adjunct for recurrent and non-resectable meningioma cases.

74

Bottoms Up Proteomics! The Dynamic Beer Proteome

Ben Schulz[1], Edward Kerr[1], Toan Phung[1], Glen Fox[1]

1. University of Queensland, St Lucia, QLD, Australia

Beer is one of humankind's oldest biotechnologies, and remains a key industry in Australia and internationally. The process of beer production involves agricultural ingredients with substantial varietal and environmental variability, and a series of bioprocessing steps. We have used mass spectrometry proteomics to investigate the complexity of the proteome throughout the course of beer production. This has uncovered a highly dynamic proteome with a wealth of post-translational modifications including proteolysis, glycosylation, and glycation that together with the bioprocess parameters control the final beer proteome. A key step in beer production is fermentation, where yeast convert sugars to ethanol and carbon dioxide. We have studied the diversity in the cell wall and global proteomes of standard brewing yeasts and wild Australian yeasts to better understand their performance in commercial brewing settings. We see great potential for the use of MS proteomics and related systems biology approaches in understanding and improving the ancient art of beer making.

740

Pathways identification by phosphoproteomic profiling identifies DNA-PK as a novel therapeutic target in Acute Myeloid Leukaemia (AML)

Heather Murray[1], Anoop Enjeti[2], Richard Kahl[1], Hayley Flanagan[1], Jonathan Sillar[1, 2], David Skerrett-Byrne[1], Juhura (Trisha) Al-mazi[1], Nathan Smith[3], Gough Au[4], Kathryn Evans[5], Richard Lock[5], Charles DeBock[6], Martin Larsen[7], Nicole Verrills[1, 8], Matthew Dun[1, 8]

1. Cancer Research Program, Hunter Medical Research Institute, and School of Biomedical Sciences and Pharmacy Faculty of Health and Medicine, University of Newcastle, Callaghan, NSW, Australia

2. Calvary Mater Hospital, Newcastle, NSW, Australia

3. Analytical and Biomolecular Research Facility Advanced Mass Spectrometry Unit, University of Newcastle, Newcastle, New South Wales, Australia

4. Viralytics, Hunter Medical Research Institute, and School of Biomedical Sciences and Pharmacy Faculty of Health and Medicine, Newcastle, New South Wales, Australia

5. Children's Cancer Institute, Lowy Cancer Research Centre, University of New South Wales , Sydney, New South Wales, Australia

6. VIB Center for the Biology of Disease, and KU Leuven Center for Human Genetics, Leuven, Belgium

7. Department of Molecular Biology and Biochemistry, Protein Research Group, University of Southern Denmark, Odense, Denmark

*8. *these authors contributed equally to this work, ,*

Acute Myeloid Leukaemia (AML) is the most aggressive form of acute leukaemia, with a 5-year survival rate of 24%. Activating mutations in the receptor tyrosine kinase FLT3 are the most common driver mutations in AML (25-30% of patients). Inhibiting the FLT3 receptor as a mono-therapeutic strategy in AML has proven difficult however, due to the development of treatment resistance and relapse. Characterisation of the oncogenic signalling pathways downstream of FLT3 is required to identify improved therapeutic approaches for AML. To this end, we have performed phosphoproteomic analysis of primary blasts from 7 AML patients (4 FLT3-mutant, 3 FLT3-wildtype). Tryptic peptides were labelled with isobaric tags prior to multistage phosphopeptide enrichment using titanium dioxide, sequential elution from IMAC, and offline HPLC fractionation, followed by analysis on an Orbitrap Velos. 6,303 phosphopeptides were identified across the 7 AML patient blast samples. Analysis of differentially expressed phosphoproteins in FLT3-mutant versus FLT3-wildtype AML patients revealed dysregulation of DNA double strand break repair pathways, with increased phosphorylation of Non-Homologous End joining (NHEJ) proteins; indicating NHEJ pathway activation. Kinase enrichment analysis predicted increased activity of the NHEJ core kinase, DNA-PK, in FLT3-mutant samples. Accordingly, cell viability assays revealed that FLT3-mutant cell lines were sensitive to pharmacological inhibition of DNA-PK. Inhibition of DNA-PK kinase activity combined with inhibition of the FLT3 receptor led to synergistic induction of cell death, selectively in FLT3-mutant cell lines and in FLT3-mutant primary AML patient samples *ex vivo*. Furthermore, DNA-PK inhibitor therapy combined with FLT3 inhibition significantly prolonged survival compared to either monotherapy in an orthotopic human xenograft mouse model of AML. In conclusion, phosphoproteomic analysis of primary AML samples has enabled identification of novel therapeutic avenues. Targeting DNA-PK in combination with standard therapeutic agents has the potential to improve outcomes for this poor prognosis cancer.

741

Insights into nanoparticles interaction with soybean and wheat at the proteomics level

Ghazala MUSTAFA[1, 2], Farhat YASMEEN[3, 2], Setsuko Komatsu[2]

1. Department of Plant Sciences, Quaid-I- Azam University, Islamabad, Pakistan, Islamabad, Pakistan

2. Department of Environmental and Food Sciences, Fukui University of Technology , Fukui, Japan

3. Department of Botany, Women University Swabi, Swabi, Pakistan

The recent trends of nanoparticles (NPs) usage in agriculture sector is captivating. To investigate the interactions of soybean and wheat with NPs, label-free proteomic technique was used. Al_2O_3 and Ag-NPs under flooding stress enhanced soybean growth compared to ZnO-NPs. Among the various sizes and concentrations, 15-nm Ag-NPs promoted soybean growth. Fe NPs have stimulatory effects on wheat germination and plant growth. The physiological parameters of wheat varieties were increased on 10 ppm Cu-NPs exposure. The Ag-NPs treated soybeans have experienced less oxygen-deprivation stress through the production of less cytotoxic by-products of glycolysis. The different sizes of Ag-NPs affected the soybean growth under flooding by regulating proteins related to amino-acid synthesis and wax formation. The 30-60 nm Al_2O_3-NPs improved the soybean growth compared to other sizes and concentrations. Al_2O_3-NPs responsive proteins were related to protein synthesis/degradation, glycolysis, and lipid metabolism. Al_2O_3-NPs promoted the soybean growth under flooding by regulating energy metabolism and cell-death. Mitochondrion was the target organelle of Al_2O_3-NPs. Mitochondrial proteomics revealed the increase in abundance of voltage-dependent anion channel with 135 nm Al_2O_3-NPs. Al_2O_3-NPs of various sizes affect mitochondrial proteins under flooding by regulating membrane permeability and tricarboxylic acid cycle activity. During recovery-stage of Al_2O_3-NPs, *S*-adenosyl-*L*-methionine-dependent-methyltransferases and enolase were involved in mediating recovery responses. The number of proteins related to photosynthesis and protein metabolism was decreased and increased in wheat varieties treated with Fe-NPs compared to untreated plants. Fe- and Cu-NPs mend the seedling growth of wheat, which might be concomitant with the enhancement of protein abundance related to photosynthesis and glycolysis/TCA in wheat varieties. These results suggested that NPs interaction with soybean mediated the energy regulation and membrane permeability. On the other hand, NPs interaction with wheat mediated the photosynthesis and TCA.

1. F. Yasmeen et al., Plant Molecular Biology Reporter, 36, 326–340 (2018).
2. G. Mustafa and S. Komatsu, J Proteome Res., 15, 4464-4475 (2016).
3. F. Yasmeen et al., J Proteomics, 143, 136–150 (2016).
4. F. Yasmeen et al., Biochem Biophys Acta, 1864, 586–1598 (2016).
5. G. Mustafa et al., J Proteomics, 148, 113-125 (2016).
6. G. Mustafa et al., J Proteomics, 128, 280-297 (2015).
7. G. Mustafa et al., J Proteomics, 122, 100-118 (2015).

742

Fast and fully automated analysis of HDX-MS data with deMix through robust peptide feature comparison

Seungjin Na[1], Jae-Jin Lee[2], Jong Wha J Joo[3], Kong-Joo Lee[2], Eunok Paek[1]

1. Hanyang University, Seoul, SEOUL, South Korea

2. Ewha Womans University, Seoul, South Korea

3. Dongguk University-Seoul, Seoul, South Korea

Characterization of protein structural changes in response to protein modifications, ligand or chemical binding, or protein-protein interactions is essential for understanding protein function and its regulation. Hydrogen/Deuterium Exchange (HDX) coupled with mass spectrometry (MS) is one of the most powerful tools for characterizing the protein dynamics and changes of protein conformation. With the advent of high throughput technologies, the data size grows everyday and an automated tool is essential for the analysis. Here, we introduce fully automated software, called *deMix* (decode deuterated mixture) that performs binomial fitting of deuterium contribution and deals directly with deuterated isotopic distributions. deMix gets robust over noise interfering with deuterated distributions through two procedures 1) dynamically determining the elution time spans for candidate peptide masses and aggregated isotopic distributions of the same peptide feature over elution time and 2) measuring how many peaks are matched in terms of abundance between theoretical and experimental isotopic distributions (referred to as *Matched Peaks Count*). The proposed method also has strength in analyzing bimodal deuterated distributions, arising from EX1 behavior or heterogeneous peptides in conformational isomer proteins. In an HDX-MS analysis of Nm23-H1, a tumor metastasis suppressor, deMix clearly showed that the HDX rates were increased under an oxidative condition (compared with a native condition). Notably, deMix discovered bimodal deuteration behaviors indicating two conformational states by stepwise oxidation of Nm23-H1. The software is freely available at https://prix.hanyang.ac.kr.

743

Enhancing middle-down proteomics data analysis of heavily modified peptides

Seungjin Na[1], Eunok Paek[1]

1. Hanyang University, Seoul, SEOUL, South Korea

Middle-down proteomics, as an intermediate between bottom-up and top-down proteomics, has a great potential in specific applications such as chromatin biology and recombinant monoclonal antibodies. In chromatin biology studies, for instance, the analysis of intact histone N-terminal tails using middle-down mass spectrometry (MS) enables the detection of co-occurring multiple post-translational modifications (PTMs) or combinatorial histone PTMs involving biological states of chromatin. Widely used computational approaches, however, have been developed to deal with short tryptic peptides in bottom-up MS, and are thus neither efficient nor optimal for the identification of relatively long peptides in middle-down MS as well as the characterization of combinatorial PTMs. Here, we introduce a workflow that can accurately detect peptide features, perform deisotoping of MS/MS spectra, and allow any number of modifications at multiple sites per peptide, enhancing both the speed and accuracy of middle-down MS data analysis. We propose a new procedure to fast decide charge states of peptide features. For efficient identification of modified peptides with no limitation, we utilized a spectral alignment algorithm based on multiple sequence tags and dynamic programming. Our workflow can support more enhanced analysis given prior knowledge about modifications (e.g., prevalent acetylations and methylations in histones) of analyzing datasets. On experimental middle-down MS datasets of recombinant Bovine histones, our workflow improved the identifications by 2-fold over existing approaches. We could identify 50 differently modified peptides from H3 N-terminal tail in terms of the unique peptide mass and also identify hundreds of peptides with high charge states over 10+. Beyond middle-down MS data analysis, we argue that our method will be also applicable in identifying peptides in various approaches such as degradomic-peptidomic analyses.

744

Glycomic and glycoproteomic approaches for development of novel glyco-biomarkers of cardiac fibrogenesis using a mouse model of dilated cardiomyopathy

Chiaki Nagai-Okatani[1], Mitsuhiro Nishigori[2], Takashi Sato[1], Naoto Minamino[2], Hiroyuki Kaji[1], Atsushi Kuno[1]

1. Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

2. Omics Research Center, National Cerebral and Cardiovascular Center, Suita, Japan

Cardiac fibrosis is a typical phenomenon in failing hearts for most cardiac diseases, including dilated cardiomyopathy (DCM), and its specific detection and quantification is crucial for the analysis of cardiac remodeling. Since cardiac fibrosis is characterized by extensive remodeling of the myocardial extracellular matrix (ECM), in which glycoproteins are major components, we assumed that fibrosis-related alterations in the cardiac glycome and glycoproteome would be suitable targets for the detection of cardiac fibrosis. In the present study[1)], we compared protein glycosylation between formalin-fixed paraffin-embedded heart sections of normal and DCM model mice by laser microdissection-assisted lectin microarray[2,3)]. Among 45 lectins, *Wisteria floribunda* agglutinin (WFA) was selected as the most suitable lectin for staining cardiac fibrotic tissues. The extent of WFA staining was highly correlated ($r > 0.98$) with that of picrosirius red staining, a common collagen staining method. However, the fibrosis-specific WFA staining did not overlap with staining signals of major collagen fiber components, indicating that collagen fibers themselves would not be WFA ligands. Further histochemical analysis with *N*-glycosidase revealed that the fibrosis-specific WFA staining was attributable to the binding of WFA to *N*-glycoproteins. Using a mass spectrometry-based approach[3)], we identified WFA-binding *N*-glycoproteins expressed in DCM hearts, many of which were fibrogenesis-related ECM proteins, as expected. Additionally, the identified glycoproteins carrying WFA-binding *N*-glycans were detected only in DCM hearts, suggesting their cooperative glycosylation alterations with disease progression. Collectively, these results indicate that WFA

staining is more suitable for the quantitative assessment of cardiac fibrogenic activity than current collagen staining methods. Furthermore, given that plasma WFA-binding glycoprotein levels were significantly correlated with the echocardiographic parameters for left ventricular remodeling, cardiac WFA-binding glycoproteins are candidate circulating glyco-biomarkers for the quantification and monitoring of cardiac fibrogenesis.

1. Nagai-Okatani C et al. Lab Invest (in press).
2. Zou X*, Yoshida M*, Nagai-Okatani C* et al. Sci Rep 7:43560 (2017).
3. Narimatsu H et al. J Proteome Res 17:4097-4112 (2018).

745

Chitosan remodels extracellular matrix integrity and regulate stomatal function leading to immunity against vascular wilt

Kanika Narula[1], Eman Elagamey[1], Magdi A.E. Abdellatef[1], Arunima Sinha[1], Sudip Ghosh[1], Niranjan Chakraborty[1], Subhra Chakraborty[1]

1. National Institute of Plant Genome Research, NewDelhi, DELHI, India

Background

Extracellular matrix (ECM) acts as a physical scaffold preventing recognition and entry of phyto-pathogens, while guard cell perceives and integrates signals metabolically. Wilt disease caused by *Fusarium oxysporum* is a major impediment for crop productivity. Although, chitosan and its acetylated form chitin are known MAMPs implicated in plant defense, the precise mechanism of chitosan triggered immunity (CTI) that leads to resistance against pathogens remains unknown. To understand the role of ECM and CTI against wilt disease, quantitative ECM proteome, metabolome along-with histological and biochemical analyses were performed in wilt resistant and susceptible chickpea (*Cicer arietinum* L.) cultivars.

Methods

Patho-stress was imposed on untreated and chitosan treated three-week-old chickpea seedlings and tissues were harvested at different post-infection time points. Stomatal mechanics were measured using GFS3000. ECM components were analysed in scanning electron microscopy and raman spectroscopy. Temporal ECM proteome and metabolome was developed using iTRAQ coupled Triple-TOF/MS and GC-MS analyses, respectively. Integrated global network was built and qRT-PCR analysis was performed to validate datasets.

Results

Morpho-histological examination revealed stomatal closure, reduced stomatal conductance and transpiration rate in chitosan-treated compared to untreated seedlings upon vascular fusariosis. ECM showed fortification leading to oligosaccharide signalling as documented by increased galactose, pectin and secondary carbohydrates. Multiomics using ECM proteomics and metabolomics identified 325 chitosan-triggered IRPs (CTIRPs), 65 CTIRMs that includes LysM domain protein, RLKs, sugars, organic acids and amino-acids linking ROS production, stomatal movements, root architecture. Immune-related correlation network identified functional hubs in CTI pathway. Data provide evidence that ROS, NO and eATP governs CTI. Induction of PR proteins, CAZymes and PAL were observed during CTI.

Conclusions

The study led us to discover the chitosan regulatory networks causing significant ECM and guard cell remodelling that translate ECM cues into cell fate decisions during fusariosis. ECM signaling and stomatal immunity plays pivotal role in CTI.

746

Phosphoproteomics of acute cell stressors targeting exercise signalling reveals drug interactions regulating protein secretion

Elise J Needham[1], Sean J Humphrey[1], Kristen C Cooke[1], Daniel J Fazakerley[1, 2], Xiaowen Duan[1], Benjamin L Parker[1, 4, 3], David E James[5, 1, 4]

1. School of Life and Environmental Sciences, The University of Sydney, Camperdown, NSW, Australia

2. Institute of Metabolic Science, University of Cambridge, Cambridge, United Kingdom

3. School of Biomedical Sciences, The University of Melbourne, Melbourne, VIC, Australia

4. Corresponding author,

5. School of Medicine, The University of Sydney, Sydney, NSW, Australia

Exercise triggers skeletal muscle signalling pathways that modulate the release of circulating factors to cause systemic health benefits. Understanding the structure of this network could lead to better strategies for treating cardiometabolic diseases. We previously showed that acute exercise induces >1000 changes in protein phosphorylation in human muscle. Here we employed a strategy to deconvolute this network by analysing phosphoproteomes of myotubes treated with small molecules that target different aspects of exercise signalling.

Small molecules were selected based on their potential ability to activate various branches of the exercise network. We tested 21 compounds and selected nine treatments to measure the entire phosphoproteome, as these covered the relevant positive controls with the greatest diversity. Single-run LC-MS/MS quantified a total of 20,249 Class I phosphopeptides, of which 24.7% were regulated in at least one treatment. Comparative analysis provided a valuable resource indicating which treatments regulate specific exercise-regulated phosphosites for future kinase-substrate relationship analysis and functional investigation. In light of the important role of AMPK in exercise, we compared a panel of six AMPK activators to determine novel potential AMPK substrates and investigate differences in the mechanism of AMPK activation. Bioinformatics suggested that combining β-adrenergic and calcium agonists would yield a phosphoproteome most closely resembling exercise. Experimentally measuring

these phosphoproteomes supported this and also revealed extensive interactions between the treatments. Dual stimulation promoted multisite phosphorylation of SERBP1, a regulator of *Serpine1* mRNA stability, a pro-thrombotic and fibrotic secreted protein. Secretomic analysis of L6 myotubes treated with β-adrenergic and calcium agonists revealed a significant decrease in SERPINE1 secretion and other deleterious secretory factors.

This work provides a novel approach to dissect exercise signalling, which may help to determine the mechanisms of beneficial effects of exercise. We demonstrate an underappreciated effect of exercise to reduce the circulating levels of certain factors, developing a framework for new insights into exercise benefits.

747

Resolving intercellular communication within the prostate cancer microenvironment through mass spectrometry-based proteomics

Elizabeth V Nguyen[2, 1], Brooke A Pereira[3], Birunthi Niranjan[2, 4], Yunjian Wu[2, 1], Mitchell G Lawrence[2, 4, 5], Thomas R Cox[3, 6], Gail P Risbridger[2, 4, 5, 7], Natalie L Lister[2, 4], Renea A Taylor[2, 5, 8], Roger J Daly[2, 1]

1. Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia

2. Cancer Program, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

3. Garvan Institute of Medical Research and the Kinghorn Cancer Centre, Sydney, NSW, Australia

4. Anatomy and Developmental Biology, Monash University, Clayton, Victoria, Australia

5. Cancer Research Division, Peter MacCallum Cancer Centre, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia

6. Faculty of Medicine, St Vincent's Clinical School, Sydney, NSW, Australia

7. Sir Peter MacCallum dept of Onocology, University of Melbourne, Parkville, Victoria, Australia

8. Physiology, Monash University, Clayton, Victoria, Australia

Introduction: In prostate cancer, cancer-associated fibroblasts (CAFs) promote tumor progression whereas non-malignant prostate fibroblasts (NPFs) do not. Resolving distinct signalling between CAFs and their NPF counterparts along with the dynamic interplay of CAFs with prostate tumour epithelium offers the potential to identify novel antineoplastic therapeutic targets.

Methods: The proteomes of patient-matched CAFs and NPFs (n=4) were analyzed by LC-MS/MS, with a hyper-reaction monitoring data-independent acquisition (HRM-DIA) workflow, to identify discriminating proteomic signatures [1]. To investigate the intercellular communication between NPFs or CAFs and co-cultured prostate epithelial (PE) cells, cell type-specific labelling with amino acid precursors (CTAP) was used in conjunction with a synchronous precursor selection tandem MS/MS/MS (SPS-MS^3) workflow. Specific culture parameters (i.e., cell types, monoculture, co-culture) were distinguishable with isobaric-labelled tandem mass tags (TMTs).

Major Findings: The CAF proteome exhibited a prominent interaction hub containing the fibrillar collagens COL1A1/2 and COL5A1; DDR2, a receptor for fibrillar collagens; and LOXL2, which promotes collagen crosslinking. Pharmacological inhibition of CAF-derived LOXL2 perturbed extracellular matrix (ECM) organization, decreased cell migration, and significantly impaired the motility of co-cultured RWPE-2 prostate tumor epithelial cells. Further interrogation of intercellular communication between NPFs or CAFs and BPH-1 immortalized PE cells revealed proteins exhibiting enhanced expression in BPH-1 cells upon co-culture with CAFs. These included the ECM protein FN1, the integrin family member ITGB4 along with its laminin ligands LAMC2 and LAMB3, cytoskeletal regulators CDC42 (a Rho family GTPase), TAGLN (an actin-binding protein) and VIM (an intermediate filament protein). Consistent with enhanced expression of CDC42, phosphoproteomics revealed perturbed phosphorylation of downstream PAK kinases in BPH-1 cells co-cultured with CAFs.

Conclusion: CAF-derived LOXL2 along with PE cell-expressed CDC42 and PAK kinases were indicated as important mediators of intercellular communication. These candidates may serve as an avenue for potential therapeutic strategies.

1. Nguyen EV, Pereira BA, Lawrence MG, Rebello RJ, Chan H, Niranjan B, Skhinas JN, Cox TR, Horvath LG, Risbridger GP, Taylor RA, Lister NL, Daly RJ (2019). Proteomic profiling of human prostate cancer-associated fibroblasts (CAF) reveals LOXL2-dependent regulation of the tumor microenvironment. Mol Cell Proteomics. Accepted. Doi:10.1074/mcp.RA119.001496

748

In Vitro Sialylation of Recombinant Alpha-1-Antitrypsin using α2,6 Sialyltransferase from *Photobacterium Damselae* Produces Disialylgalactose *N*-Glycans

Edward Pallister[1, 2], Matthew Choo[1], Jien-Nee Tai[1], Dawn Leong[1], Wen-Qin Tang[1], Kun Huang[2], Andrea Marchesi[2], Peter Both[2], Christopher Gray[2], Say Kong Ng[1], Pauline Rudd[1], Sabine Flitsch[2], Terry Nguyen-Khuong[1]

1. Bioprocessing Technology Institute, Singapore, SINGAPORE

2. School of Chemistry & Manchester Institute of Biotechnology, University of Manchester, Manchester, United Kingdom

The genetic disease AAT deficiency (AATD) is caused by a lack of Alpha-1-antitrypisn (AAT) in the body as a result of causes a number of complications ranging from chronic obstructive pulmonary disease to liver cirrhosis. Augmentation therapy of severe AATD sufferers involves human serum plasma AAT. Unfortunately, this treatment is expensive so a means of reducing the costs and improving the treatments efficiency are of significant interest. The half-life of alpha-1-antitrypsin is increased in vivo markedly by increasing the level of sialylation. Here we describe how the unique disialylgalactose sialylation (DSG) activity of a-2,6-sialyltransferase from Photobacterium Damselae. This disialylgalactose epitope is a unique a2,3 and a2,6 sialic acid sialylation

of free end galactose residues. DSG can be used in tandem with CHO produced alpha-1-antitrypsin to markedly increase sialylation. We used numerous glycomic and glycoproteomic techniques such as permethylation, ethyl esterification together with MS methods such as glycomic methods to such as MALDI-MS, Orbitrap-MS and ion-mobility mass spectrometry (IM-MS) to characterise the unique DSG epitope and identify the N-Glycans that contain the previously uncharacterised disialylgalactose motif. This is also the first example of producing the unique disialylgalactose sialylation on a CHO produced biotherapeutic glycoprotein.

749

Kinetics of protein precipitation: optimizing recovery, purity, and throughput using the ProTrap XG

Jessica L. Nickerson[1], Alan A. Doucette[1]

1. Dalhousie University, Halifax, NS, Canada

Introduction

Efficient sample preparation is key to any proteomics workflow, with three priorities at the forefront: protein recovery, purity, and sample throughput. Acetone precipitation is a common method of purifying proteins from a complex sample prior to LC-MS/MS analysis, but is sometimes set aside due to seemingly variable yields.[1] Our group previously established that the addition of 1-30 mM salt with 80% *v/v* acetone facilitates recoveries >95%.[2] We later developed a filtration cartridge (ProTrap XG) which automates precipitation and SDS depletion.[3] In order to maximize recovery, purity and sample throughput, we investigated the kinetics of protein precipitation with respect to structural properties of proteins, including molecular weight, hydrophobicity, and charge.

Methods

Extracted yeast lysates were combined with various concentrations of salt and 80% *v/v* acetone at a defined temperature (-20 to +37 °C), within the ProTrap XG. The samples were incubated for varying times, followed by centrifugation and pellet re-solubilization. Protein recovery was quantified by the bicinchoninic acid assay, followed by a bottom-up proteomics workflow, which was used to identify precipitated proteins after varying incubation times. Structural properties of the recovered proteins were compared across the various incubation conditions. Purity was assessed using the methylene blue active substances assay to quantify residual SDS.

Results

By incorporating ≥10 mM NaCl, acetone precipitation recovered >95% of protein following 2 minutes incubation at room temperature. Protein identifications made by bottom-up LC-MS/MS analysis indicate no difference in the structural properties of proteins recovered after 2 minutes, 60 minutes and 24 hours incubation.

Conclusions

Under optimal conditions, acetone precipitation quantitatively recovers all types of protein within a sample in 2 minutes while depleting >99% of the initial SDS.

1. Zhang et al. (2015) Analytical Biochemistry, 484, 40-50.
2. Doucette et al. (2015), Journal of Proteomics, 118, 140-150.
3. Doucette et al. (2013), Analytica Chimica Acta, 796, 48-54.

75

Insights in food digestion: tracking peptides from meat and meat hydrolysates by simulated gastrointestinal digestion and label-free proteomics

Evelyne Maes[1], Santanu Deb-Choudhury[1], Jessica Gathercole[1], Stephen Haines[1], Anita Grosvenor[1], Erin Lee[1], Ancy Thomas[1], Stefan Clerens[1, 2, 3]

1. AgResearch Limited, Lincoln, New Zealand

2. Biomolecular Interaction Centre, University of Canterbury, Christchurch, New Zealand

3. Riddet Institute, Massey University, Palmerston North, New Zealand

To increase our insights in the nutritional and functional benefit of meat proteins, knowledge on its digestion after consumption are crucial. Information about protein truncation during digestion is critical to track the relative abundance of the released peptides during both gastric as well as intestinal digestion stages. Apart from the protein breakdown, digestion may also release peptides that display functionality in addition to their nutritional value, such as bioactivity. Bioactive peptides are known to display a variety of health-promoting activities including anti-hypertensive, anti-oxidant and anti-inflammatory properties. The digestion process may release these bioactive peptides which are encrypted in larger protein sequences. However, passage through the gastrointestinal tract drastically alters peptide profiles and thus may enhance of limit the amount of bioactive sequences found.

In this study, we compared the peptide compositions after simulated gastrointestinal digestion of powdered meat samples from 18 month old steers and of enzymatically produced meat hydrolysates. These meat protein hydrolysates are important in the food industry due to their versatility of use, from flavour enhancers to functional ingredients. To evaluate the differences in *in vitro* digestion between meat powder vs meat hydrolysate, we performed a label-free quantitative proteomic analysis. Data were acquired via data-dependent acquisition on an Impact HD Q-TOF. Protein and peptide identification was performed using Peaks

Studio software. Further data analysis via principal component analysis displayed a clear distinction in the number and the type of peptides produced in meat powder vs. hydrolysed meat. Additional bioactivity profiling using our in-house database showed that numerous bioactives were generated during simulated digestion of the meat hydrolysate, while fewer hits were found in the meat powder. This indicates the potential of meat hydrolysates as functional ingredients in foods.

750

Comprehensive analysis of the effects of high-fat diet on protein abundance in the mouse liver using SWATH acquisition method.

Tomoya Nishibata[1], Sayaka Yamawaki[2], Naruha Misawa[2], Rinko Nagamine[2], Tomoki Awaji[2], Yuki Oosedo[2], Ayaka Sakaue[2], Kunihiro Kishida[1, 2], Kouhei Nagai[1, 2]

1. Graduate School of Biology-Oriented Science and Technology, Kindai University, Kinokawa, Wakayama, Japan

2. Faculty of Biology-Oriented Science and Technology, Kindai University, Kinokawa, Wakayama, Japan

[Introduction] Liver is a center of carbohydrate and lipid metabolism and deeply involved with the development of metabolic syndrome. If a method to comprehensively capture the biological and pathological changes in the liver caused by obesity could be established, it would contribute to the development of drugs or functional food ingredients to prevent lifestyle-related diseases. In this study, we performed SWATH-mass spectrometry of livers obtained from high-fat diet (HD)-induced obese mice. [Methods] Six weeks old C57/BL/6J male mice were fed with HD, or normal diet (ND) for 5 weeks. On the day 35, livers were collected after 16 h of fasting. Proteins extracted from the liver tissues were digested by trypsin, and the resulting peptides were analyzed by a nanoLC-MS/MS (SCIEX, TripleTOF5600+ system). Ion library of liver protein digests was built by data dependent analyses （DDA） and database search were performed against Uniprot mouse protein database. The SWATH measurements were performed under three conditions in which the amount of protein (3 ug or 5 ug), the Q1 window width (25 Da or variable) and number (25 or 50), and the MS/MS accumulation time (50 or 100 msec) were changed. [Results and Discussion] In total, 1,524 proteins were successfully quantified. Among them, 121 and 107 proteins increased by more than 1.2 times in the HD and ND, respectively (p <0.05). Compared to the ND, the amount of many proteins involved in detoxification and respiratory chain decreased in the HD group, some of which were consistent with previous reports, while the amount of proteins involved in cell-autonomous immunity, lipid metabolism, and Fe metabolism increased in the HD group. This results suggested the possibility that high-fat diet intake and increased serum cholesterol may be linked through abnormalities in iron metabolism.

751

Glycoproteomic analysis of human colorectal adenoma tissue

Krishnatej Nishtala[1], Howard CH Yim[2], Emad El-Omar[2], Nicolle H Packer[1]

1. Dept of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. St George & Sutherland Clinical School, UNSW, Sydney, NSW, Australia

Colorectal cancer develops over a few decades with genetic mutations occurring in the mucosal cells leading to their autonomous proliferation into premalignant adenomatous polyps which further become cancerous. Hence our goal was to characterise the disease associated changes in the glycoproteome of colorectal adenoma tissue using mucin-depleted human colon polyp proteins.

Tissue lysates from biopsies of adenomatous polyps and matched adjacent non-adenomatous tissue from the colon of patients diagnosed with dysplasia were precipitated by the addition of guanidine hydrochloride (6M GuHCl), which is the reported method of choice for MUC 2 enrichment (Larsson et al., Glycobiol., 2009). The pellet and the supernatant were analysed for their proteome. An equal fraction of the tissue was processed without GuCl precipitation to capture the whole proteome. Affinity enrichment by ZIC-HILIC was carried out on tryptic digests of the total tissue and the GuHCl supernatant for capturing the non-mucin glycopeptides. The samples were analysed on a Thermo QExactive HFX Orbitrap mass spectrometer. The data were analysed using Proteome Discoverer 2.2 and Byonic v3.3 for glycopeptide identification.

Proteomic analysis of the 6M GuHCl precipitated proteins from the tissue lysates identified MUC2 (104 unique peptides) as the expected predominant colonic mucin but MUC5AC, MUC5B, Microfibril associated glycoprotein 4, Transmembrane glycoprotein NMB and Vitronectin were also abundant. These proteins were either not identified or of very low abundance in the whole tissue proteome. Cytosolic proteins comprised the majority of the GuCl supernatant fraction indicating efficient total mucin enrichment by GuCl precipitation. Byonic analysis of the ZIC-HILIC enriched fraction of the trypsin-digested GuHCl supernatants identified around 550 glycopeptides (*O*-glycosylated: 145, *N*-glycosylated: 320 and *N & O* glycosylated: 85) corresponding to about 200 proteins in the colorectal polyp. This approach was used to characterise the non-mucin related glycoproteome of colorectal adenoma dysplasia compared to matched normal tissue.

752

Proteogenomic approach to kinase regulation in osteosarcomas with different original sites: Report by ICPC JAPAN team

Rei Noguchi[1], Emi Hattori[1], Akihiko Yoshida[2], Akira Kawai[3], Tadashi Kondo[1]

1. Division of Rare Cancer Research, National Cancer Center Research Institute, Tsukiji, Tokyo, Japan

2. Department of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tsukiji, Tokyo, 2974049

3. Division of Musculoskeletal Oncology, National Cancer Center Hospital, Tsukiji, Tokyo, Japan

Aberrant regulations of kinase activity play a crucial role in the carcinogenesis and cancer progression. As therapeutic targets and biomarkers for companion diagnosis, the mutations in the activity domain of kinases are extensively investigated. In this study, we examined the utility of the multi-omics approach toward the comprehensive understanding of aberrant regulation of kinase. We investigated two types of osteosarcomas; one was originated from bone, and another from soft tissue. These two osteosarcomas exhibit the distinct clinical features, and we studied the mutations and activity of kinases, and the response to kinase inhibitors in their patient-derived cell lines. Mutation status of 27 kinases was examined by NCC Oncopanel, which is based on the next-generation sequencing technology. The activities of 100 kinases were monitored by the PamStation 12 platform, which is based on the in vitro kinase assay. In addition, the anti-proliferative effects of 30 FDA-approved kinase inhibitors were also examined in the cell lines. We found that the two types of osteosarcoma cells showed the remarkable differences in the activities of FGFRs1-4. Corresponding to the kinase activities, the inhibitors against them showed the considerable anti-proliferative effects. We found the mutations in PIK3CA in the osteosarcoma from soft tissue, and the amplification in EGFR in the osteosarcoma from bone. However, the inhibitors for the PI3K-AKT pathway or EGFR did not show the significant effects on their cell lines. Our result suggests that we need to investigate the kinase activities in addition to genetic mutations, to predict the effects of kinase inhibitors. The different status of kinase mutations, activities, and response to inhibitors should be considered in an integrative way. Overall, the multi-omics experiments and data integration will be a crucial approach to understand the mechanisms of cancer progression, and develop the novel therapy.

753

A Mix-And-Match Library Approach for Enriching Plasma Proteome Discovery

Zainab Noor[1], Abidali Mohamedali[1], Seong Beom Ahn[1], Mark S. Baker[1], Shoba Ranganathan[1]

1. Macquarie University, North Ryde, NEW SOUTH WALES, Australia

The advent of SWATH-MS analysis has led to technological advances as well as significantly improved data quality and quantity. Each SWATH experiment covers a large swathe of the proteomics discovery space. In practice, however, the only peptides (and by inference proteins) able to be identified are those that are present in the SWATH library from previous DDA experiments. Multiple studies have attempted to increase library depth and quality, including advanced peptide fractionation, abundant protein depletion and other methods using the same sample type. Low abundance plasma proteins suffer particularly due to the prevalence of high abundant proteins that mask these during DDA experiments required to construct SWATH libraries. Additionally, disparate results from different search engines make conclusions requiring extensive validation using orthogonal techniques time-consuming and not always accurate. Here, we overcame these obstacles by combining libraries using iSwathX [1] from two distinct samples (blood plasma and cells) that were run on the same machine and platform to generate a large extended SWATH library. We used two commonly-available search engine pipelines to design a high coverage library dataset against which SWATH patient data from 4 stages of colorectal cancer were searched. Our extended cross-sample SWATH library was used to; (i) identify and quantify a large number of proteins from neat “undepleted” plasma (~2000 proteins reliably quantified using ~4500 unique peptides of at least 8 amino acids long) and (ii) detect ~70 differentially-expressed proteins. SWATH library concatenation can lead to the discovery of new disease biomarkers by enhancement of proteomics discovery space and by significantly increasing resolution.

1. Noor, Z., Wu, J. X., Pascovici, D., Mohamedali, A., Molloy, M. P., Baker, M. S., & Ranganathan, S. (2018). iSwathX: an interactive web-based application for extension of DIA peptide reference libraries. Bioinformatics, 35(3), 538-539.

754

Enhanced identification of protein ADP-ribosylation with an engineered Af1521 macro domain

Kathrin Nowak[2, 1], Florian Rosenthal[1], Deena Leslie Pedrioli[1], Mareike Bütepage[3], Birgit Dreier[4], Jens Sobek[5], Andreas Plueckthun[4], Herwig Schüler[6], Bernhard Lüscher[3], Michael Hottiger[1]

1. Department of Molecular Mechanisms of Disease, University of Zurich, Zurich, Switzerland

2. Molecular Life Science PhD Program of the Life Science Zurich Graduate School, University of Zurich, Zurich, Switzerland

3. Institute of Biochemistry and Molecular Biology, RWTH Aachen University, Aachen, Germany

4. Department of Biochemistry, University of Zurich, Zurich, Switzerland

5. Functional Genomics Center Zurich, ETH Zurich and University of Zurich, Zurich, Switzerland

6. Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden

Protein ADP-ribosylation is a covalent, reversible post-translational modification whereby ADP-ribose (ADPr) is transferred from nicotinamide adenine dinucleotide (NAD^+) onto different amino acid residues of target proteins to regulate protein structure, stability, activity and molecular interactions. To further elucidate the cellular function of ADP-ribosylation, it is of importance to detect and decode the target protein landscape. This is however challenging due to the low abundance of ADP-ribosylation. In order to decipher the ADP-ribosylome including the ADPr acceptor sites, we recently co-developed a mass spectrometry-based approach that enriches ADP-ribosylated peptides with the help of the Af1521 macrodomain, which binds the terminal ADP-ribose

units. To improve the enrichment and thus detection of ADP-ribosylated proteins in lysates, we aimed at improving the affinity of Af1521 to ADP-ribosylated peptides.

We thus evolved by random mutagenesis and *in vitro* selection an engineered Af1521 macrodomain with significantly increased affinity towards ADP-ribosylated peptides as verified by ELISA and Surface Plasmon Resonance (SPR). The crystal structure of the evolved macro domain revealed that two amino acid substitutions form an additional salt bridge forcing a rotation of the ribose and subsequently leading to less conformational flexibility and trapping of the ADP-ribose unit. The comparison of the conventional and the engineered Af1521 macro domain in the peptide enrichment of our proteomic ADP-ribosylome approach revealed that the identification rate of the ADP-ribosylated targets increased 2-fold with the mutated binder when assaying genotoxic stress condition that leads to mainly Serine-ADP-ribosylation. Further characterization of the evolved Af1521 domain as fusion protein with a Fc fragment confirmed the improved detection of cellular ADP-ribosylation by immuno blot and immunofluorescence, suggesting that our new engineered Af1521 macro domain can also serve as a valuable tool for the identification and detection of ADP-ribosylation.

755

MALDI imaging identifies transketolase to be up-regulated in serous ovarian cancer patients following chemotherapy-resistant disease relapse

Tannith M Noye[1], Parul Mittal[2], Noor A Lokman[1], Oliver Schilling[3], Peter Hoffmann[4], Martin K Oehler[5, 1], Carmela Ricciardelli[1]

1. Adelaide Medical School, Obstetrics and Gynaecology, University of Adelaide, Adelaide, South Australia, Australia

2. Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

3. Institute for Clinical Pathology, Albert-Ludwigs Universtität, Freiburg, Baden-Württemberg, Germany

4. Future Industries Institute, University of South Australia, Adelaide, South Australia, Australia

5. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

Ovarian cancer is the most lethal gynaecological disease, with poor 5-year survival rates and limited treatment options for patients who develop resistant disease. The majority of ovarian malignancies, up to 70% of cases, are high-grade serous carcinomas that have high chemosensitivity to first line platinum-based therapies. However, 75% of patients will become chemoresistant, following relapse. The underlying mechanism for developing resistance to chemotherapy in ovarian cancer is poorly understood. In this study, we employed peptide matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) in formalin fixed paraffin embedded sections of ovarian cancer tissues at the time of diagnosis and following relapse from 4 patients with serous cancer. Using, MALDI-MSI we have identified *m/z* features that were present in relapsed tissues but absent in ovarian cancer tissues at diagnosis. One of these was identified as transketolase using LC-MS/MS and data dependent analysis on paraffin sections. Transketolase was validated by immunohistochemistry and was elevated by ~3 fold in relapse tissues compared to matching ovarian cancer tissues at diagnosis (P=0.035, paired t-test). In addition, transketolase (*TKT*) expression was significantly increased (P=0.0075, unpaired t-test) in carboplatin resistant CAOV3 cells compared to parental cells measured by qRT-PCR. In summary, MALDI-MSI has the potential to identify proteins associated with chemotherapy resistance that can be evaluated as a novel therapeutic target.

756

Altered glycosylation of genetically engineered pigs for successful xenotransfusion using LC-MS/MS approach

Myung Jin Oh[1, 2], Nari Seo[1, 2], Jinyoung Park[1, 2], Jaekyoung Ko[1, 2], Jjin-dan Kang[3], Xi-jun Yin[3], Hyun Joo An[1, 2]

1. Chungnam National University, Yuseong-gu,, DAEJEON, South Korea

2. Asia Glycomics Reference Site, Chungnam National University, Daejeon, Korea

3. Department of Animal Sciences College of Agriculture, Yanbian University, Yanji City Jilin Pro, China

Xenotransfusion of pig blood has been considered as one of potential solutions to the frequent shortage of human blood. One key bottleneck to xenograft is allogenic immune rejection by human antibodies caused by species-specific glycosylation features. Thus, glycosylation directly related to the immune system has been recognized as a gatekeeper for successful xenotransplantation. The major target antigens of human natural antibodies and complements are non-human glycan moieties containing terminal galactose-a-1-3- galactose (alpha-Gal) and N-glycolyl neuraminic acid (NeuGc), respectively. They have been genetically modified to overcome rejection mechanism and incompatibility in xenotransfusion. In order to investigate whether glycosylation can be altered by destruction of the CMAH and GGTA genes, N-glycans released from RBC and plasma of WT and double KO pigs (n=8, respectively) were examined by PGC LC/MS and MS/MS approach. Species-specific glycans which are linked to target genotype containing specific glycan moieties such as sialic acids, linkage isomers, and alpha-Gal epitope were preferentially characterized. In particular, NeuGc vs NeuAc-sialylated glycan ratio of plasma and RBC of wild pigs are 1:10 and 1:40, respectively. NeuGc glycan was not detected in KO plasma, but less than 1% in KO RBC. The unique structure of alpha-Gal epitope was confirmed by the simultaneous loss of galactose and HexNAc from the terminal as well as the absence of diagnostic ion indicating the hybrid type glycan in MS/MS spectrum. Alpha-Gal epitopes were discovered with significant quantities (> 10% of total glycan) in WT pigs. Although GGTA gene was destructed in KO pigs, the relative amount of alpha-Gal still remained about 3%. Here, for the first time, we characterized structure- and linkage-specific glycans found in RBC and plasma of WT and genetically modified pigs, respectively.

757

The Human Protein Atlas, version 19: accessing data for your research

Per Oksvold[1], Kalle von Feilitzen[1], Martin Zwahlen[1], Mattias Forsberg[1], Fredric Johansson[1], Mathias Uhlen[1]

1. KTH Royal Institute of Technology, Solna, STOCKHOLM, Sweden

Programmatic access to open database resources are as important as it can be a tough exercise to get right. In the latest version of the Human Protein Atlas portal (www.proteinatlas.org, v19) we have addressed this by adding new ways to export the data following the FAIR principles as well as providing more comprehensive datasets, both for the new Blood and Brain atlases, but also for external sources. In terms of functionality, we have extended the XML format, added custom data selection and JSON format export for user generated searches as well as added access to remapped transcriptome data from external sources. These new features will empower researchers in using the Human Protein Atlas in more advanced ways.

758

Progress on identifying and characterizing the human proteome: 2018-2019 metrics from the HUPO Human Proteome Project

Gil Omenn[1], Lydie Lane[2], Chris Overall[3], Fernando J Corrales[4], Jochen M Schwenk[5], Young-Ki Paik[6], Jennifer E. Van Eyk[7], Siqi Liu[8], Stephen Pennington[9], Michael P Snyder[10], Mark S. Baker[11], Eric W Deutsch[12]

1. University of Michigan,, Ann Arbor, MI, United States

2. CALIPHO Group, SIB Swiss Institute of Bioinformatics and Department of Microbiology and Molecular Medicine, Faculty of Medicine, University of Geneva, CMU,, Geneva, Switzerland

3. Life Sciences Institute, Faculty of Dentistry, University of British Columbia, Vancouver, BC, Canada

4. Centro Nacional de Biotecnologia (CSIC), Madrid, Spain

5. Science for Life Laboratory, KTH Royal Institute of Technology, Solna, Sweden

6. Yonsei Proteome Research Center, Yonsei University, Seoul, South Korea

7. Advanced Clinical BioSystems Research Institute, Cedars Sinai Precision Biomarker Laboratories, Barbra Streisand Women's Heart Center, Cedars-Sinai Medical Center, Los Angeles, CA, United States

8. BGI Group-Shenzhen, Shenzhen, China

9. School of Medicine, University College Dublin, Conway Institute, Dublin, Ireland

10. Department of Genetics, Stanford University, Palo Alto, CA, United States

11. Department of Biomedical Sciences, Faculty of Medicine & Health Sciences, Macquarie University, Sydney, Macquarie Park NSW 2109, Australia

12. Institute for Systems Biology, Seattle, WA, United States

The Human Proteome Project (HPP) annually reports on progress made throughout the field in credibly identifying and characterizing the complete human protein parts list and making proteomics an integral part of multi-omics studies in medicine and the life sciences. NeXtProt release 2019-01-11 contains 17,694 PE1 proteins, which represent 89% of all 19,823 neXtProt predicted coding genes (all PE1, 2, 3, 4 proteins), up from 17,470 in release 2018-01. Conversely, the number of neXtProt PE2, 3, 4 proteins, called the "missing proteins" (MPs), has been reduced from 2,949 to 2,129 over the past three years. Since the inception of the Human Proteome Project, PeptideAtlas has been the source of uniformly re-analyzed raw mass spectrometry data for neXtProt. PeptideAtlas gained 495 canonical proteins between 2018 and 2019. Multiple strategies have been employed to detect hard-to-identify proteins. Meanwhile, the Human Protein Atlas has released version 18.1 with immunohistochemical evidence of expression of 17,000 proteins, survival plots as part of the Pathology Atlas, and its Cell Atlas, and is moving toward completion of a harmonized resource on tissue-specific RNA expression data. Many investigators apply multiplexed SRM-targeted proteomics for quantitation of organ-specific popular proteins in studies of various human diseases. The 19 teams of the Biology and Disease-driven B/D-HPP published a total of 382 publications in 2018, bringing proteomics to a broad array of biomedical research.

759

Quantitative, comprehensive multi-pathway signaling analysis using an optimized phosphopeptide enrichment method combined with an internal standard triggered targeted MS assay

Bhavin Patel[1], Penny Jensen[1], Aaron S Gajadhar[2], Sebastien Gallien[3], Jae Choi[1], Romain Huguet[2], Derek Bailey[2], Graeme McAlister[2], Shannon Eliuk[2], Markus Kellmann[4], Tabiwang N. N Arrey[4], Alexander Harder[4], Andreas Huhmer[2], Kay Opperman[1], John Rogers[1]

1. Thermo Fisher Scientific, Rockford, IL, United States

2. Thermo Fisher Scientific, San Jose, CA, United States

3. Precision Medicine Science Center, Thermo Fisher Scientific, Cambridge, MA, United States

4. Thermo Fisher Scientific, Bremen, , Germany

Introduction

There is broad interest in quantifying dynamic protein phosphorylation states in cellular signaling pathways under different conditions. Enrichment is necessary for better detection of the low abundant phosphorylated proteins, and multiplexed quantitation reagents parallelize processing across a multitude of experimental conditions. We have combined SMOAC (Sequential enrichment of Metal Oxide Affinity Chromatography), 146 AQUATM heavy-labeled phosphopeptide standards, and

SureQuantTM targeted MS to evaluate changes in phosphorylated protein abundance under different stimulation conditions. The specific phosphopeptides have been chosen to cover biologically interesting phosphosites from several different signaling pathways.

Method

HeLa/A549 cells were grown with different stimulation conditions (hIGF-1/hEFG) before in-solution digestion. One milligram of each digest spiked with phosphopeptides standard was subjected to SMOAC analysis using the Thermo Scientific Pierce Hi-SelectTM TiO2 and Fe_NTA phosphopeptide enrichment kits. Both eluents were combined before LC-MS analysis using Thermo Scientific Dionex nanoLC™ system coupled to Thermo Scientific™ Orbitrap Exploris™ 480 or Orbitrap Eclipse™ Tribrid™ Mass Spectrometers. To ensure optimal measurement of each target, a novel SureQuant method was performed where real-time heavy peptide detection triggered high-sensitivity measurement of endogenous targets. Data analysis was performed with Proteome Discoverer and Skyline software.

Results

We have previously described our optimized SMOAC phosphopeptide enrichment method and we have shown with that method significant improvement in the number of phosphopeptides identified. In this study, we developed a targeted assay based upon 146 AQUA heavy-isotope phosphopeptide standards. More than 90% of heavy peptides were quantified with high sensitivity and reproducibility across different MS acquisition methods. The phosphopeptide standards spiked into stimulated HeLa/A549 cell digest, followed by enrichment using the SMOAC method, allowed quantitation of about 60 endogenous phosphopeptides and 134 heavy phosphopeptides by PRM or SureQuant method.

Conclusion

This phosphopeptide standard with novel targeted MS analysis allowed quantitation of phosphorylation changes from >80 signaling pathway proteins.

76

Development of an ESI-MS based serotyping assay for S*almonella*

Hwee Tong Tan[1], Dave Ow[1], Wei Ling Tan[2], Calvin Yeo[2], Xuezhi Bi[1, 3]

*1. Bioprocessing Technology Institute, Agency for Science, Technology and Research (A*STAR), Singapore, SINGAPORE*

2. National Centre for Food Safety, Singapore Food Agency, Singapore

3. Duke-NUS Medical School, Singapore

Publish consent withheld

760

Tools to search, explore and interrogate your Proteomics data with UniProtKB.

Sandra Orchard[1], Emily Bowler[1], Emanuele Alpi[1]

1. EMBL-EBI, Cambridge, CAMBS, United Kingdom

The identification and analysis of proteomics data is inherently reliant on databases, requiring them to be of a consistent high-quality. UniProt is a comprehensive, expert-led, publicly available database of protein sequence, function and variation information. UniProt currently holds over 13,000 reference proteomes, that are constantly updated and reviewed based on collaborations with a variety of sources such as Ensembl, RefSeq, ENA and proteome-centric repositories such as ProteomicsDB, Peptide Atlas, MaxQB etc.

To facilitate searching of proteomics data, reference proteomes can be downloaded in FASTA format or queried programmatically using the UniProt API, allowing researchers direct access to data from large scale studies, variation annotation, and proteomics data amongst others mapped to UniProt from cross-referenced databases. Data is available for download and querying in a range of formats; including XML, FASTA and the recently published HUPO-PSI PEFF.

To facilitate further investigation of target proteins, mapped proteomics data pertaining to unique peptides is available graphically alongside variant, domain, and post-translational modification sites for each canonical protein sequence, allowing researchers to reference their data within the context of a proteins' peptide sequence and understand the current functional biological information available for their proteins of interest.
All data are freely accessible from www.uniprot.org

761

The prognostic impact of post translational modification in childhood acute myeloid leukemia patients

Aya Osama[1], Shahd ezzeldin[1], Eman Ahmed[1], Sameh Magdeldin[1]

1. Children Cancer Hospital in Egypt 57357-CCHE, Cairo, CAIRO, Egypt

Acute myeloid leukemia (AML) is a frequent hematological malignancy. Despite enormous therapeutic efforts that range from various cytotoxic agents to allogeneic stem cell transplantation, overall survival of patients with AML remains unsatisfying. The poor survival rates are mainly due to therapy-related mortality, failure of induction chemotherapy and early relapse. In this study, we investigated the possible role of protein post translational modifications on 3 AML subtypes; KMT2A t(9-11), normal karyotype and complex karyotype in an attempt to better understand the mechanistic and cancer pathophysiology. Our finding revealed that a characteristic pattern in acetylation and ubiquitination of c terminal zing finger protein might be a good prognostic indicator.

762

Comparative Proteome Profiling of Acute Myelogenous Leukemia Cell Lines under acute and chronic starvation condition

Aya Osama[1], Eman Ahmed[1], shahd ezzeldin[1], Sameh Magdeldin[1]

1. Children Cancer Hospital 57357, Cairo, CAIRO, Egypt

Myeloid leukemias are a heterogeneous group of diseases originating from bone marrow myeloid progenitor cells. Patients with myeloid leukemias can achieve long-term survival through targeted therapy, cure after intensive chemotherapy or short-term survival because of highly chemoresistant disease. To elucidate the proteomic characteristics AML, we applied a comparative analysis between different leukemic cell lines under acute and chronic starvation conditions. Here we tested acute promyelocytic leukemia (HL60), acute myeloid leukemia (Kasumi), acute myeloid leukemia (Nomo-I) and acute myelomonocytic leukemia (ML-2). Our finding revealed significant and characteristic metabolic cell dependent pathways reflecting different resistance capabilities between AML cell lines. This finding might be helpful to better understand chemoresistance patterns between different AMLs and associated recovery of patients.

763

A complementary peptide enrichment and mass spectrometric strategy for in-depth mining of immunopeptidome

Kirti Pandey[1], Sri H Ramarathinam[1], Nicole A Mifsud[1], Anthony W Purcell[1]

1. Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Background: The field of immunopeptidomics has provided critical insights into crucial aspects of antigen presentation and human leukocyte antigen (HLA) peptide interactions there by helping to deconvolute the complexity associated with immunopeptidome. HLA polymorphism also generates tremendous diversity in the processing and presentation of different peptide antigens and adds to the challenge of analysing these diverse immunopeptidomes. The use of high-resolution mass spectrometry (MS) and different peptide enrichment techniques can help in further exploring and delineating the immunopeptidome. These techniques can be exploited to unravel the immunopeptidome of cancer including leukaemia, thereby making inroads towards identifying tumour specific peptides which can be used for T cell-based vaccines or targets for CAR T cell therapies.

Methods: The eluate of 1 X 10^9 cells of human acute myeloid leukaemia (AML) cell line, THP-1 (HLA-A*02:01, -B*15:11 and -C*03:03) were lysed and HLA- class I bound peptides purified by immunoaffinity chromatography. The eluate from this immunoaffinity step was then further separated by two different techniques; namely off-line reverse phase HPLC (RP-HPLC) or using a molecular weight cut off filter (MWCO) followed by liquid chromatography and tandem mass spectrometry (LC-MS/MS) on a Tribrid Orbitrap Fusion (Thermo Scientific).

Results: A total of 37675 HLA class I bound peptides were identified across the triplicate dataset with 17434 HLA A*02:01 restricted and 20241 restricted by the other two HLA class I alleles expressed on THP-1 cells. The use of the two peptide enrichment steps was shown to expand the peptide repertoire identified. Several cancer testis antigen and oncogenic peptides were identified along with post translationally modified (PTM) peptides.

Conclusion: This study identified naturally presented peptide antigens derived from an AML cell line at unprecedented depth. This approach will aid in development of improved diagnostic and patient stratification tools along with peptide based immunotherapeutic approaches.

764

Chemical modification of proteins to mimic LysC proteolysis: Application of 1,2-dicarbonyl compounds for arginine modification

Boomathi P Pandi[1], Nagarjuna R Chary[2], A.S. Kamalanathan[1], Sripadi Prabhakar[2], Varatharajan Sabareesh[1]

1. Advanced Centre for Bio Separation Technology (CBST), Vellore Institute of Technology (VIT), Vellore - 632014, Tamil Nadu, India

2. Centre for Mass Spectrometry, Analytical Department, CSIR-Indian Institute of Chemical Technology (CSIR-IICT), Hyderabad, Telangana - 500007, India

Introduction:

We are interested to develop a new method for an emerging area, middle-down proteomics (MDP). In MD approach, proteases such as LysC, AspN, etc. are used to produce longer proteolytic peptides, which yields better sequence coverage than bottom-up approach and hence, post-translational modifications can be detected more reliably (e.g., in histones and antibodies). We apply the strategy of modifying the guanidine side chain of arginines of proteins by 1,2-cyclohexanedione (CHD) or phenylglyoxal (PG), prior to trypsin digestion, which would result in 'longer tryptic peptides' and such arginine-modified tryptic peptides mimic LysC derived peptides.

Methods:

Before applying this method for proteomics, we investigated five model proteins: β-lactoglobulin, β-casein, RNase A, ovalbumin and human transferrin. Carbamidomethylation of proteins was done before arginine modification and ~100 molar excess of CHD or PG was used (16 hrs, pH 8.4 (borate), ~25^0 C) for arginine modification reactions. Subsequently, each sample was digested with trypsin (37^0C) for different incubation times, which was monitored by liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS: 1290 Infinity LC attached to 6540 or 6545 Q-TOF (Agilent) and Acquity UPLC attached with Quattro Premier XE (Waters)).

Findings:

Three arginine-modified tryptic peptides of lengths in the range: ~26 - 50 amino acid residues (a.a.r), were detected from each of β-lactoglobulin and β-casein. In all these modified tryptic peptides, one molecule of CHD or PG was covalently added to the sidechain of the arginine residue. Tryptic peptides of very short lengths (< 5 a.a.r) and not longer than 25 a.a.r. were not observed from arginine-modified RNase A. Similarly, longer arginine-modified tryptic peptides were observed in other model proteins as well.

Conclusion:

Thus, in cases, where LysC is useful for MD approach based proteomics and protein sequencing, our strategy of arginine-modification-cum-trypsin digestion can be a new approach, which can be an alternative method and cost-effective too.

1. Takahashi, K., (1968). The Reaction of phenylglyoxal with arginine residues in proteins. J. Biol. Chem. 243, 6171-6179
2. Patthy, L., and Smith, E.L. (1975). Reversible modification of arginine residues. J. Biol. Chem. 250, 557-564
3. Brock, J.W.C., Cotham, W.E., Thorpe, S.R., Baynes,J.W., and Ames,J.M. (2007). Detection and Identification of arginine modifications on methylglyoxal-modified ribonuclease by mass spectrometric analysis. J.Mass Spectrom. 42, 89-100

765

Analytical guidelines for co-fractionation mass spectrometry obtained through global profiling of gold standard *Saccharomyces cerevisiae* protein complexes

Chi Nam Ignatius Pang[1], Sara Ballouz[2], Daniel Weissberger[1], Loïc M. Thibaut[3], Joshua J. Hamey[1], Jesse Gillis[2], Marc R. Wilkins[1], Gene Hart-Smith[1]

1. School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia

2. Stanley Center for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, The United States of America

3. School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW, Australia

Protein Correlation Profiling (PCP) enables many protein complexes to be identified in single experiments. A typical PCP experiment involves fractionation of endogenous and untagged protein complexes by size or other physiochemical parameters, followed by LC-MS/MS and label-free quantification of each fraction. Proteins in the same intact complex are co-eluted and have highly correlated abundances across fractions. Although PCP can be used to identify intact complexes, the best approaches for the collection and analysis of PCP data remain undefined. This study aims to gain insight into the collection and analysis of PCP data by benchmarking PCP datasets against gold standard complexes in a well-characterized model organism (*Saccharomyces cerevisiae*). Our analysis of experimental and modelled PCP datasets suggests that using a combination of fractionation methods and combining these results, for example using Fisher's combined probability test, is more beneficial than using a stand-alone fractionation method to collect the same number of fractions. From benchmarking the effects of 17 correlation metrics on the identification of known complexes, we showed that some metrics (e.g. Spearman correlation) were more effective than others (e.g. mutual information). While PCP identified many complexes observed in traditional experiments (e.g. AP-MS and Y2H), PCP also identified putative novel complexes. To measure the overlap of the PCP datasets with orthogonal gene expression data, we ran EGAD (Ballouz *et al.* 2017) on an aggregate PCP and gene co-expression network. We find the addition of gene co-expression to PCP data contributed mainly to confident identification of known complexes (e.g. EGAD scores of 0.63 for PCP alone, 0.72 for co-expression alone, and 0.71 for PCP and co-expression). The similarity in performance of EGAD scores suggests that novel complexes within PCP data are rare, and that confirmation of these novel complexes may require orthogonal experimental validation, for example using cross-linking mass spectrometry.

1. Ballouz S, Weber M, Pavlidis P, Gillis J. (2017) EGAD: ultra-fast functional analysis of gene networks. Bioinformatics. 33(4):612-614.

766

High-throughput automated sample preparation workflow for extensive pan-species repository and SWATH-mass spectrometry

Selvam Paramasivan[1, 2], Rajesh Gupta[2], Nana Satake[1], Paul Mills[1], Pawel Sadowski[2]

1. University of Queensland, Brisbane, QLD, Australia

2. Central Analytical Research Facility, Queensland University of Technology, Brisbane, QUEENSLAND, Australia

Large scale label-free quantitative mass spectrometry that relies on the robust sample processing workflow. High resolution mass spectrometry instruments are so effective that the sample preparation has become the key limiting factor for consistent and reproducible large scale protein quantification. Here, we introduce a high-throughput and fully automated sample preparation pipeline that integrates the key steps in SWATH-based proteomic workflow including digestion, desalting and fractionation to build comprehensive pan-species spectral repositories and collect quantitative profiles for 96 samples.

The pipeline utilizes a liquid handling robot and filter plates for efficient FASP-based protein digestion. Desalting and fractionation are achieved within the same system using vacuum-aided STAGE tips filled with SCX and SDB-XC membrane without compromising reproducibility. We leveraged the improved throughput to create a compendium of cattle proteins from 30 clinically relevant tissues and body fluids and demonstrated uniform and effective digestion performance across all tissue types that allowed us to reproducibly quantify over 7000 proteins using SWATH-MS. This end-to-end automated proteomic sample preparation overcomes technical barriers, minimizes the cost and enables accurate quantitative proteomic analysis across any biological system.

767

Precise Identification of Site-Specific Glycoproteins using Multiple Fragmentation in Mass Spectrometry with GlycoProteome Analyzer (GPA)

GunWook Park[1], Jin Young Kim[1], Ji Won Lee[1, 2], Hyun Kyoung Lee[1, 2], Heeyoun Hwang[1], Jong Shin YOO[1, 2]

1. Biomedical Omics Center, Korea Basic Science Institute, Cheongju-si, Chungbuk, South Korea

2. Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon, Republic of Korea

We have developed Integrated GlycoProteomeAnalyzer (GPA) for high throughput analysis of N-and O-glycoproteome, which combines methods for tandem mass spectrometry with a database search and algorithmic suite. For precise identification, we created novel scoring algorithms such as M-score for glycopeptide selection from oxonium-ions, S-score for theoretical isotope pattern match of molecular ion, and Y-score for glycopeptide identification from MS/MS fragmentation with calculation of estimated false discovery rate (FDR). All glycoprotein samples were digested with trypsin. Then, those site-specific glycopeptides were enriched by hydrophobic interaction liquid chromatography and analyzed by nano-reversed-phase liquid chromatography coupled to tandem mass spectrometry with both HCD and CID-MS/MS fragmentation. The resultant LC/MS/MS data were then automatically analyzed using GPA: tryptic glycopeptides were identified against the database from Uniprot database. In GPA, all amino acid sequence and N-and O-glycosylation site information were obtained from the Uniprot database. From the Swiss-Prot accession number of human protein, GPA program automatically construct glycopeptide database for human sample. GPA has been designed to easily handle high-throughput glycoproteomic data with a graphical user interface. GPA software is demonstrated on website (https://www.igpa.kr/). Additionally, O-GPA, newly developed search engine, allows direct analysis of site-specific O-glycopeptides from glycoprotein mixtures using the O-GPA-DB from Uniprot with estimated FDR ≤ 1%. O-GPA is capable of automatic identification and quantification of O-glycopeptides using HCD and ETD-MS/MS spectra.

768

Temporal Proteomic Alterations in Mouse Brain during the Recovery Phase of Traumatic Brain Injury (TBI)

Joonho Park[1], Byung-Mo Oh[2], Youngsoo Kim[1]

1. Interdisciplinary Program for Bioengineering, Seoul National University College of Engineering, Seoul

2. Department of Rehabilitation Medicine, Seoul National University Hospital, Seoul

Traumatic brain injury (TBI) is accompanied by a disorder in cognitive, emotional, and behavioral functions without notable surface wounds after a physical impact to the head. Every year, 1.5 million people experience TBI from traffic accidents, falls, violence, and etc. currently, TBIs are classified into mild, moderate, and severe grades according to the consciousness status assessed by the Glasgow Coma Scale. In the case of mild TBI, injury in cranial nerves recovered after sufficient rest, however, it is difficult to expect a natural recovery in TBIs above the moderate level. Thus, molecular components involved in the recovery are deemed to be significantly different between mild TBI and moderate-to-severe TBI. Yet, studies that investigate what molecules or pathways are relevant to recovery phase of TBI are rare, especially for proteomic approaches. In this study, the prefrontal cortex tissues of mice undergone mild/severe TBI were collected over time periods of recovery, then their proteome were analyzed using LC-MS/MS. We performed quantitative analysis using 10-plex TMT, identifying about 300 proteins of which expression levels were temporally altered. We also discovered molecular pathways activated or inhibited under TBI condition through bioinformatics analysis. As a result, we found that the sirtuin signaling pathway was gradually inactivated over the phase of recovery and that the proteins involved in the lipid metabolism were stably up-regulated till two weeks from TBI. These findings will help to uncover molecular mechanisms leading recovery from TBI. Furthermore, the biomarkers that play important role in recovery of mild TBI can be used as a target of drug intervention intended to severe TBI patients.

769

Identification of misannotated COSMIC mutations based on the combination of AMP guideline and multiple public databases

Kyoung-Jin Park[1], Moon Jung Kim[1], Woori Jang[2]

1. Myongji Hospital, Goyang, Gyeonggi-Do, South Korea

2. Inha University School of Medicine, Incheon, Korea

Background: The presence of benign variants or passenger mutations in driver genes from public databases is one of the well-documented limitations when reporting clinical actionability of somatic mutations. However, our understanding on classification and interpretation of somatic mutations is far from being perfect. The aim of this study is to investigate the frequency of benign variants from the Catalogue of Somatic Mutations in Cancer (COSMIC) Cancer Gene Census (CGC) and the interpretation concordance of OncoKb variants in depth.

Methods: We used a total of 339,425 somatic mutations in CGC genes registered in the COSMIC. Multiple databases incluing CIViC and Gene Drug Knowledge database and *in silico* tools were used to interpretate the clinical actionability of the mutations. All mutations were reanalyzed by a four-tiered system based on the consensus recommendation of the Association for Molecular Pathology (AMP), American Society of Clinical Oncology, and College of American Pathologists, and compared to OncoKb classification.

Results: The frequency of benign or likely variants, which were categorized as Tier IV variants according to AMP guideline, in the COSMIC CGC was 1% (4,200/339,425) of the mutations. Although 75% (3,129/4,200) of the Tier IV variants were flagged as polymorphism, while there was no warning sign in remaining 25% of benign variants. In addition, 70% (2,952/4,200) of the Tier IV variants were recurrently observed in at least two samples. Overall concordance of actionable mutations between AMP classification (Tier I or Tier II) and OncoKb (Oncogenic or likely oncogenic) were 88% (3,118/3,565) of variants studied.

Conclusions: Interpretation and reporting of somatic mutations in cancer databases need to use with caution. We found that significant number of somatic mutations registered in COSMIC CGC may not be clinically-actionable mutations. This study could be a good starting point for in-depth review of somatic mutation information essential for precision oncology.

77

Leavening it for later: Proteomics of gluten-free bread

Mitchell G. Nye-Wood[1], Angela Juhasz[1], Michelle L. Colgrave[1, 2]

1. Food and Agricultural Proteomics, School of Science, Edith Cowan University, Perth, Western Australia, Australia

2. CSIRO Agriculture and Food, St Lucia, Queensland, Australia

A range of dietary proteins pose risks to human health. In susceptible individuals, consumption of cereals can cause coeliac disease by the autoimmune reaction to gluten proteins, wheat allergy through an IgE-mediated reaction to various types of prolamins, and non-coeliac wheat sensitivity by a reaction to non-gluten proteins such as alpha-amylase/trypsin inhibitors expressed by wheat. Importantly, the only effective treatment of these diseases is to eliminate sources of gluten from the diet.

Gluten-free (GF) bread can be made by substituting wheat flour for potato or corn starch. In some bread products, wheat flour is used wherein enzymes are proposed to degrade the gluten during fermentation. In this study, a number of GF bread varieties that are available commercially are explored with the aim to determine the GF status of these products and hence determine their safety for people with gluten intolerance.

Discovery proteomics is employed to characterise the protein components within the finished product, and to assess the degree of protein degradation during fermentation. The identities and abundance of gluten and homologous proteins are compared to a typical wheat bread. High sensitivity protein discovery was performed on a TripleTOF 6600 (SCIEX) mass spectrometer with microflow HPLC, and SWATH proteomics was used to track the global abundance of key peptides. These data informed the development of quantitative approaches targeting gluten proteins using multiple reaction monitoring on a QTRAP 6500+ (SCIEX) mass spectrometer coupled to an Exion HPLC system.

This study explores how extended microbial fermentation affects the proteome of commercial GF bread.

770

Multi-omics characterization of never-smoking lung cancer patients without known driver mutations

Seung-Jin Park[1], Sung-Ho Goh[2], Shinyeong Ju[3], Sang Jin Lee[2], Byung Ha Yoon[1], Jeong-Hwan Kim[1], Kyung Chan Park[1], Ji-Youn Han[2], Cheolju Lee[3], Seon-Young Kim[1]

1. Personalized Genomic Medicine Research Center, KRIBB, Daejeon, CN, South Korea

2. Research Institute, National Cancer Center, Goyang-si, Gyeonggi-do, South Korea

3. Center for Theragnosis, KIST, Seoul, South Korea

As an effort for international cancer proteogenomics consortia (ICPC) initiative, we are charactering the molecular changes among never-smoking lung cancer patients (up to 150) without known driver genes such as EGFR, KRAS, and ALK fusions. Until now, we have generated the whole genome, transcriptome, DNA methylome, proteome, and immune cell profiling of 116 Korean never-smoking lung cancer tissues selected from screening more than 1000 lung cancer tissues. Most never-smoking lung cancer patients showed lower mutation burden (2.6/Mb) than smoking lung cancer patients (11.8-14.4/Mb from TCGA data) except for a few patients (194/Mb) with POLE and POLD1 mutations or mismatch repair defects. We identified several significantly mutated genes among the 116 patients. Transcriptome analysis revealed three sub-groups which showed significant differences in the mutation burdens. We are currently performing integration of multi-genomics and proteomics data and continue to generate multi-omics data for more never-smoking ling cancer patients.

Grant support: NRF-2017M3A9F9030559 (to C Lee), NRF-2017M3A9F9030565 (to KC Park), and NRF-2017M3A9F9030648 (to JY Han)

771

NanoBlow: A simple device to limit contaminants during nanoLC-MS

Sean J Humphrey[1], Ben Crossett[1], Benjamin L Parker[2, 1]

1. University of Sydney, Sydney, NSW, Australia

2. University of Melbourne, Melbourne, VIC, Australia

Limiting contamination of LC-MS systems and reducing downtime associated with maintenance and cleaning is essential for productivity. We developed a simple device that creates a gas curtain barrier to prevent ions entering the MS inlet. The gas can be quickly and easily applied when certain contaminant ions are known to elute. We show the device can prevent the build up of contaminants on the heated transfer capillary following >100 injections of a crude tissue lysate and improves peptide identifications. The device may provide a promising approach towards improving instrument robustness.

772

Burn wound blister fluid proteomics to assist with early clinical decision making

Tuo Zang[2, 3, 1], Leila Cuttle[4, 2, 1], Daniel Broszczak[2, 1], James A Broadbent[2, 1, 5], Catherine Tanzer[4, 2, 1], Tony J Parker[2, 1]

1. Tissue Repair and Translational Physiology Program, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia

2. School of Biomedical Science, Faculty of Health, Queensland University of Technology, Brisbane, Qld, Australia

3. Wound Management Innovation Cooperative Research Centre, West End, Qld, Australia

4. The Centre for Children's Burns and Trauma Research, Queensland Children's Medical Research Institute, University of Queensland, Royal Children's Hospital, Herston, QLD, Australia

5. Agriculture and Food, CSIRO, Brisbane, Queensland, Australia

Burn wounds in children result in significant and long lasting physical and emotional impacts which require repeated surgeries and rehabilitation efforts well into adulthood. Burn wound treatment involves 'de-roofing' of blisters, however, little attention has been paid to the composition of the blister fluid and if the constituents have diagnostic or prognostic potential. Blister fluid analysis provides an opportunity to non-invasively investigate the biology of the initial response to burn injury and potentially uncover novel diagnostics or prognostics to assist in early clinical decision making that might assist healing and reduce scarring.

We performed a proteomic analysis of 87 paediatric burn blister fluid samples using liquid chromatography - tandem mass spectrometry with SWATH (data independent) acquisition. The blister fluid proteomes of all samples were compared to the key clinical features of burn depth classification (superficial-partial thickness; deep-partial thickness; and full-thickness) and time-to-reepithelialisation (time to healing). Both of these clinical parameters are critical for enabling accurate clinical decisions regarding early burn treatment options such as whether or not to graft.

More than 600 proteins were quantitatively compared between samples using partial least squares-discriminant analysis which revealed significant differences in the biochemistry associated with both burn depth and time-to-reepithelialisation. Interestingly, the protein profiles provided evidence of potential clinical misclassification of some burn wounds examined in this study.

While full-thickness burns are often grafted and superficial-partial thickness burns are often not grafted, it is more difficult to determine if deep-partial thickness burns should be grafted or not. Thus, with further validation, utilisation of the results of this study could translate to aid in clinical decision making. Overall, this study provides new insights into the early stages of burn wound biology in children and may help with the development of diagnostic or prognostic tools to assist with clinical decisions regarding burn treatment options.

773

Batch normalisation and mixed effects models in TMT or SWATH – two sides of the same coin

Dana Pascovici[1], Jemma Wu[1], Karthik Kamath[1], Yunqi Wu[1], Thiri Zaw[1], Mehdi Mirzaei[1]

1. Australian Proteome Analysis Facility, Macquarie University, NSW, Austria

With the advent of larger datasets being tackled in discovery proteomics, batch effects are inevitable, whether in labelled multiplexed formats such as TMT which come in batches of 6 or 10 or larger label-free experiments using SWATH/DIA. Various avenues exist for accounting for these effects, whether by normalisation to remove the variation between batches, or by using statistical models such as linear mixed effects models that allow for random batch effects. If using the normalisation approach, global methods that act at the sample level and normalise for the total or median sample amount often fail to completely remove the batch variability, while *IRS* or *ComBat* normalisation can more completely remove differences between batches. The *IRS* method as published can only be applied to batches of similar size and therefore is well suited for TMT cross-run normalisation, however we introduce a small variation of it which can also be applied to normalise when combining across SWATH batches of uneven size. We present the effects of these normalisation approaches in the context of a spike-in experiment containing three 10-plex TMT replicate runs with 2%, 5% and 10% of yeast peptides spiked into mouse cell lysate, which presents a known quantitation scenario. We also show that, from the point of view of determining differentially expressed proteins, similar results can be obtained by disregarding normalisation altogether, and applying mixed effects models with random batch effects, thus the two approaches can be used to provide additional computational checks and balances. However, normalisation remains crucial if the goal is to provide a complete dataset to be used for data mining or machine learning approaches.

774

Metabolomic/lipidomic DESI imaging of different cell cultures

Hadeer Mattar[1], Emrys Jones[2], Heather Patsiouras[3], Emmanuelle Claude[2], Clare Mills[1]

1. Division of Infection, Immunity & Respiratory Medicine, University of Manchester, Manchester, UK

2. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

3. Waters Corporation, Sydney, NSW, Australia

Lipidomics provides a comprehensive structural and functional characterization of various lipids in different biological samples. Desorption electrospray ion source (DESI) has been currently used as a mass spectrometry imaging (MSI) technique for the analysis as well as the identification of lipids in the cell culture.

In this study, we aimed to investigate the distribution and the localisation of lipids among three different cell lines: two of the gut epithelial cell lines (Caco2 and HT29-MTX cells) and one of cancerous basophil cells (RBL=Rat Basophilic Leukemia) . We also aimed to identify the master charged lipids by applying positive and negative modes.

Three Cell lines were investigated which are; Caco2, HT29-MTX as a gut epithelial cells, and RBL as basophil leukemia. Data were acquired using a DESI source attached to a Xevo G2-XS mass spectrometer, with data acquired in both positive and negative mode. DESI spray conditions were set at 2µl/min, 98:2 MeOH: water with nebulizing gas pressure of 5 bar.

Cell grown on cover slips had the media removed for the MSI experiment and were directly mounted onto microscope glass slides, using double sided tape and placed onto the DESI stage, with no sample preparation or pre-treatment.

Initial experiments were carried out in positive ion mode (50 um pixel) with mainly glycerophospholipids such as PC (34:1), K+ and triglyceraides being identified for all cell lines. However, differences in the lipid intensity profiles were observed between cell lines, including PC (36:4),K+ being over expressed in Caco2 whereas (PC(P-32:0), K+ was more pronounced in HT29-MTX. Furthermore, an experiment at 20um pixel size showed individual agglomerate of cells with a different distribution within the same cell line.

Negative ion mode showed a variety of molecular information for metabolite and lipid species, including oleic acid which was the most abundant in Caco2.

775

Unified access to cancer proteogenomics data

Caleb M Lindgren[1], David W Adams[1], Sadie Tayler[1], Sean J. I. Beecroft[1], Isaac D Taylor[1], Samuel H Payne[1]

1. Brigham Young University, Provo, UT, United States

Interdisciplinary consortia like the Clinical Proteomic Tumor Analysis Consortium (CPTAC) drive scientific discovery by creating coordinated and structured datasets on a large scale, amassing both a number of samples and diversity of measurements that requires a collaborative effort. Although the dataset is analyzed in a primary publication, these consortia encourage and promote reanalysis by the scientific community to explore new questions or apply novel methodologies. Therefore, expanding access to data is an important goal for publicly funded research.

To more seamlessly enable secondary analyses, data dissemination technologies need to meet their intended audiences in the most convenient and accessible way possible. Although storing data in supplemental tables or cloud-based archives is fine for historical records, it is not the optimal dissemination method. To facilitate easy reanalysis, data needs to be accessible within an analytical environment - it needs to be accessible to software via APIs. No such mechanism currently exists for the quantitative molecular data tables which form the foundation of data analysis tasks.

We present a unified API for accessing all CPTAC proteogenomic data for cancer cohorts. Each cancer type contains data for ~100 tumors, which are comprehensively characterized with genomics, transcriptomics, and proteomics, as well as relevant clinical information. This data is packaged within a python module, cptac, that is freely distributed through the Python Package Index (PyPI). This removes many of the common barriers to re-analysis. Most importantly, the data is accessible via code without parsing, formatting, web-browsing or passwords. The data tables of omics and clinical information come pre-formatted as dataframes ready for statistical and visual analytic packages, and the API handles any complex merging between data types. The module contains extensive tutorials and documentation to assist users in understanding the data. Additionally, the API wraps common algorithms used by CPTAC in their primary analyses.

776

Discovering metabolic alterations in glial tumors using a combination of ambient MS profiling and nanoLC/MS

Stanislav Pekov[1, 2], Anatoly Sorokin[1, 3], Vasiliy Eliferov[1], Vsevolod Shurkhay[4], Alexander Potapov[4], Eugene Nikolaev[5], Igor Popov[1]

1. Moscow Institute of Physics and Technology, Moscow, Russian Federation

2. V.L. Talrose Institute for Energy Problems of Chemical Physics, N.N. Semenov Federal Research Center of Chemical Physics RAS, Moscow, Russian Federation

3. Institute of Cell Biophysics RAS, Pushchino, Russian Federation

4. N. N. Burdenko Scientific Research Neurosurgery Institute, Moscow, Russian Federation

5. Skolkovo Institute of Science and Technology, Skolkovo, Russian Federation

Lipid and energy metabolism reprogramming is tightly connected with cancer and enables the possibility to differentiate malignant and intact tissues on the molecular level. Since cancer-related metabolic changes have multiple interconnections, it is required to implement the approaches of big data analysis to investigate metabolic alterations regardless of the natural variability. In this work, the ambient mass spectrometry molecular profiling was combined with conventional lipidomics pipeline to detect and identify key molecular alterations in human glial tumors.

Histologically annotated samples were provided by N.N. Burdenko National Scientific and Practical Center for Neurosurgery. More than 100 samples obtained from more than 30 patients were analyzed using Inline Cartridge Extraction approach developed for the mass spectrometry assisting in neurosurgery. For each sample, measurements were done twice - one part of a sample was investigated immediately after resection using Thermo LTQ XL. The other part of the sample was frozen and transferred into the scientific laboratory to perform high-resolution profiling with Thermo LTQ XL Orbitrap instrument. Since stereotaxic biopsy of brain tumors usually provide small amounts of sample for analysis, less than 10% of samples were analyzed by LC-MS/MS. Based on the results of histological examination, all data were classified, and sets of distinctive features were highlighted using non-tumor pathological tissues as a control. The analysis of the significant amount of samples analyzed by ambient mass spectrometry profiling allows refining lists of detected features while tandem mass spectrometry coupled with reversed-phase liquid chromatography was used for the identification of peaks, corresponding to these features.

The results of this study not only enhance the reliability of mass spectrometric approaches for intraoperative tumor tissue differentiation but shed light on the molecular mechanisms allowing such differentiation.

777

Glycoproteomic measurement of site-specific polysialylation

Ruby Pelingon[1], Cassandra Pegg[1], Lucia Zacchi[1], Toan Phung[1], Christopher Howard[1], Matthew Hardy[2], Catherine Owczarek[2], Ben Schulz[1]

1. Univeristy of Queensland, Brisbane, QUEENSLAND, Australia

2. CSL Limited, Melbourne, Victoria, Australia

Polysialylation is the enzymatic addition of a negatively charged sialic acid polymer (>8 residues) at the terminal end of glycans. Polysialylation plays important roles through embryonic development, and is involved in neurological diseases, neural tissue regeneration, and cancer. Polysialic acid (PSA) is also researched as a potential biodegradable and non-immunogenic biomolecule conjugated to therapeutic drugs to improve their pharmacokinetics. PSA chains can vary in length, composition, and linkages. In addition, the site of polysialylation and the abundance of this modification are important determinants of protein function. PSA is difficult to analyse by mass spectrometry due to the biochemical properties of the molecule (negative charge and size). Most analytical tools available focus on determining degree of polymerization and composition, but do not address the key question of site specificity and abundance. We developed a high-throughput LC-ESI-MS/MS glycoproteomic method to measure site-specific polysialylation of glycoproteins. This method combines enzymatic (endosialidase EndoNF) and chemical (mild acid hydrolysis) elimination of PSA and sialic acids, leaving the glycan backbone intact to provide site-specificity. As proof of principle, we tested this method in polysialylated recombinant human neural cell adhesion molecule (rHuNCAM) and non-polysialylated serum purified human IgG glycoproteins (as negative control). Control untreated glycopeptides (mono/polysialylated) and desialylated glycopeptides were analysed by LC-ESI-MS/MS, and data was processed using Proteome Discoverer (v2.0.0.802, Thermo Fisher Scientific), Byonic (v2.13.17, Protein Metrics), and in-house designed scripts to identify glycopeptides and calculate relative abundance of glycoforms. This new methodology allows for the detailed comparison between polysialylated and non-polysialylated versions of glycoproteins, and it is an effective analytical method to facilitate studies of biological polysialylation and for quality control of polysialylated therapeutic proteins.

778

Proteomics of triacylglycerol accumulation and DGAT inhibition in HepG2 cells

Lifeng Peng[1], Bhumika Bhatt-Wessel[1], Bill Jordan[1], John Miller[1]

1. Victoria University of Wellington, Kelburn, WELLINGTON, New Zealand

Excess accumulation of triacylglycerol (TAG) in liver occurs in a variety of diseases including non-alcoholic fatty liver disease (NAFLD) in which TAG-rich lipid droplets accumulate in the cytoplasm of hepatocytes leading to liver injury, including inflammation, and in some cases fibrosis and cirrhosis. This represents a huge public health problem worldwide. The final and committed step of TAG synthesis in the ER is catalysed by diacylglycerol O-acyltransferase (DGAT) enzymes DGAT1 and DGAT2 (Bhatt-Wessel *et al*., 2018). DGAT inhibition to eliminate TAG accumulation as a therapy for NAFLD has attracted interest. Changes in proteins, and the pathways they are a part of, represent the molecular mechanisms that lead to NAFLD and the way in which the cell copes with the reduced lipid accumulation caused by DGAT inhibition. Protein biomarkers and the underlying molecular mechanisms provide potential diagnostic and treatment measures of NAFLD; however, they have not been well-studied. Here we used HepG2 human liver carcinoma cells as a model for lipid accumulation, DGAT inhibition, and proteomics. We treated HepG2 cells with 1 mM of a mixture of fatty acids (FAs) for 18 hours and induced TAG accumulation by about 2-fold with a minor effect on cytotoxicity (82% cell viability). When DGAT was partially inhibited, the triacylglycerol accumulation was reduced by 50%. We compared the proteomics profiles of these cells. Among the 1,202 proteins identified, 111 proteins significantly changed in abundance in FA-treated cells compared to untreated control cells; 184 proteins significantly changed in inhibitor and FA-treated cells compared to the FA-treated cells, and 185 significantly changed in inhibitor and FA-treated cells compared to control cells. Bioinformatics analysis revealed that these changed proteins were organised into pathways responsible for lipid metabolism and dyslipidaemia disease. This study provides a resource for the proteomes of fatty liver cells.

779

Advancing high-throughput top-down analysis of proteoforms up to 60 kDa using an Orbitrap Eclipse Tribrid mass spectrometer

Scott Peterman[1], Romain Huguet[1], KRISTINA SRZENTIC[1], Christopher Mullen[1], John E. P. Syka[1], Jesse Cantebury[1], Michael Goodwin[1], Graeme McAllistar[1], Michael W Senko[1], Vlad Zabrouskov[1], Luca Fornelli[2]

1. Thermo Fisher Scientific, San Jose, CA, USA

2. University of Oklahoma, Norman, OK, United States

Top-down mass spectrometry (TDMS) offers the theoretical possibility of capturing the full molecular complexity of every gene product (proteoform) present in a cell. However, characterizing proteoforms in high-throughput proteomic experiments remains challenging. Electrospray ionization generates wide charge state distributions for each proteoform, leading to reduced spectral dynamic range. Furthermore, large cations are difficult to transfer to high resolution mass analyzers without undesired fragmentation. As a consequence, only the most abundant proteoforms are typically detected and selected for fragmentation, and the effect dramatically worsen for proteins > 30 kDa. We performed extensive TDMS characterization of the budding yeast proteome, to characterize subcellular compartment-specific proteoforms using an Orbitrap Eclipse Tribrid mass spectrometer. We first benchmarked the TDMS performance using LC-separated standard proteins from 8.5-29 kDa (Sigma Aldrich). Results indicate that reduced Orbitrap analyzer pressure and optimized ion transmission produce a 2-4 fold increase in spectral signal-to-noise ratio (SNR) in comparison to previous Orbitrap Tribrid MS. This SNR gain leads to more accurate on-the-fly charge state determination and facilitates *m/z* to mass deconvolution. Importantly, the improved performance in intact mass determination (i.e., MS^1) at high resolving power (120,000 at 200 *m/z* or higher) is reached without compromising the quality of data-dependent

(DDA) fragmentation spectra (i.e., MS^2) obtained via higher-energy collisional dissociation. To characterize large (30-60 kDa) proteoforms, we applied a targeted data acquisition strategy: a quadrupole-isolated small *m/z* region (1.5 *m/z* wide) was first expanded via Proton Transfer Charge Reduction to enable accurate determination of proteoforms' average mass (at medium resolution, 15'000 at 200 *m/z*) and subsequently fragmented via collisional dissociation. Results obtained indicate that the targeted PTR-based data acquisition method doubles the number of protein entry and proteoform identifications (at 1% FDR) compared to DDA experiments, allowing the characterization of large proteoforms fundamental for yeast central metabolism.

78

How are the scientific concepts evolved at the National Institutes of Health?

Sudhir Srivastava[1]

1. DHHS / NIH / NCI / Division of Cancer Prevention / Cancer Biomarkers Research Group, ROCKVILLE, MARYLAND, United States

-

780

Software for automation of MS/MS glycoproteomic analysis

Toan K Phung[1], Lucia F Zacchi[2], Cassandra Pegg[1], Ruby Pelingon[2], Nazmi Nazir[1], Aaron Li[1], Benjamin L. Schulz[1, 2]

1. School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia

2. ARC Training Centre for Biopharmaceutical Innovation, The University of Queensland, Brisbane, Queensland, Australia

While mass spectrometry (MS) is at the core of many glycoproteomic workflows, limitations of the typical data dependent acquisition (DDA) MS experiment mean that often information of lower abundance glycoforms is lost and tools are limited for quantitative analyses. The introduction of sequential windows acquisition of all theoretical fragment ion spectra (SWATH), a data independent acquisition (DIA) MS technique, into the workflow has greatly expanded the scope of quantifiable glycoforms. However due to the technique being relatively young, the toolsets for processing of the huge amount of data produced are still sub-optimal. We would like to introduce two prototype programs targeting limitations in MS/MS programs for glycoproteomic data.

DIALib is a program to be placed at the top of the SWATH data extraction workflow, which aims to complement or replace the usage of the typical DDA based library. DIALib can generate a theoretical ion library currently targeted for PeakView from a simple protein sequence for interrogation of SWATH raw data. Due to its theoretical nature, the DIALib library could potentially cover all glycoforms selected within the chromatographic run of a sample without any information beyond protein/peptide amino acid sequence.

Glypnir is a tool for calculating the relative abundance from AUC values of different *N*-glycoforms across different complex samples from manually validated data. Glypnir's input requires output from Byonic software (Protein Metrics) as a node in Proteome Discoverer and user supplied sequence information of the proteins of interest. Using this information, the tool can map any identified precursor back to the protein sequence as well as any accompanying modifications, and group the identified *N*-glycopeptides by site and monosaccharide composition.

Together DIALib and Glypnir offer an automated solution for different parts of the current glycoproteomic pipelines.

781

Hands-off: fully automated & TMT-compatible sample preparation on the PreON liquid handling platform employing the iST-NHS workflow

Garwin Pichler[1], Fabian Hosp[1], Nils A. Kulak [1]

1. PreOmics, Martinsried, BAVARIA, Germany

Sample preparation is an important component of the overall mass spectrometry-based proteomics workflow and remains to be a limiting factor for high-throughput analyses. Here, we present a fully automated end-to-end solution for standardized sample processing, including cell lysis, digestion, TMT labeling and peptide cleanup within 4 hours. In order to minimize sample loss and improve reproducibility, we sought to completely automate TMT sample processing from cell lysis to ready-to-measure labeled peptides. To facilitate this, we aimed to combine the NHS adaption of the iST technology (Kulak et al., 2014) with a newly developed automation platform called PreON.

The workflow described here integrates robotic handling with chemical labeling of peptides in a one-pot reaction device, including cell lysis, protein denaturation, reduction, alkylation, digestion as well as the peptide cleanup. The PreOn platform is a newly developed benchtop liquid handling system capable of processing up to 12 samples in parallel. It is integrated with a built-in centrifuge with evaporation feature, ultrasound liquid detection, deck load check and a drag-and-drop method selection for easy, fast and convenient operation.

Here, we demonstrate full automation of sample processing without any manual interference from cell lysis to ready-to-measure peptides for various biological sample types such as multiple cancer cell lines, bacteria, yeast or human plasma, scaling from 1-100 µg of protein input material. Furthermore, we integrate chemical labeling such as iTRAQ or TMT in the whole workflow and thus permit multiplexing in a fully automated fashion. To this end, we have employed 11-plex TMT to successfully label several cell lines achieving >98% labeling efficiency and a reproducibility of $R^2 = 0.98$ for biological replicates (CVs <10%). The combination of the one-pot iST workflow and the fully automated liquid handling enables a simplified, efficient and highly reproducible preparation of biological input materials and ranges.

782

In-situ exploration of lipid composition remodeling using sperm maturation in the rat epididymis by MALDI mass spectrometry imaging

Melanie Lagarrigue[1], Regis Lavigne[1], Blandine Guevel[1], Andrew Palmer[2], Karine Rondel[1], Jan H Kobarg[3], Dennis Trede[3], Charles Pineau[1]

1. PROTIM, Inserm U1085 - Irset, Rennes, France

2. European Molecular Biology Laboratory , EMBL, Heidelberg, Germany

3. Bruker Daltonik, SCiLS GmbH, Bremen , Germany

Spermatozoa acquire their fertilizing capacity during a complex maturation process that occurs in the epididymis. This process involves a substantial remodeling of the proteins and lipids located at the surface of the gamete. Epididymis can be divided into three regions (the caput, corpus, and cauda) or into 19 intra-regional segments based on histology. Most studies carried out on the evolution of the lipid composition during epididymal maturation has been performed on sperm samples or tissue extracts.

Here, we used MSI to study the precise localization of lipids directly in the rat epididymis tissue. The spatial segmentation of molecular information provided by MSI revealed that the rat epididymis could be divided into 21 molecular clusters different from the 19 intra-regional segments. The discriminative *m/z*values that contributed the most to each molecular cluster were then annotated and corresponded mainly to phosphatidylcholines, phosphatidylethanolamines, triacylglycerols, plasmalogens, phosphosphingomyelins, glycerophosphates, lysophosphatidylethanolamines and lysophosphatidylcholines.

Although phosphatidylcholines, triacylglycerols, phosphatidylethanolamines were predominantly detected in the epididymis caput, an important increase in the number of plasmalogens, phosphosphingomyelins and lysophosphatidylcholines annotated in the cauda was observed. MALDI images reveal that molecules belonging to the same family can have different localizations along the epididymis. For some of them, annotation was confirmed by on-tissue MS/MS experiments. A 3D-model of the epididymis head was reconstructed from 61 sections analyzed with a lateral resolution of 50 µm and can be used to obtain information on the localization of a given analyte in the whole volume of the tissue.

Work is currently ongoing to integrate our high quality in-situ lipid-distribution dataset with quantitative proteomes obtained for the three main epididymal regions, with the aim to identify the key enzymes involved in lipid metabolism within the organ. This pioneering work opens new perspectives for elucidating the role of lipid metabolism in sperm maturation during its epididymal transit.

783

Precision, personalized and digital medicine: research trends over the past 10 years

Elena A. Ponomarenko[1], Ekaterina I. Ilgisonis[1], Katerina V. Poverennaya[1], Mikhail A. Pyatnitskii[1], Andrey V. Lisitsa[1]

1. Institute of Biomedical Chemistry, Moscow, RUSSIA, Russia

The number of papers in all areas of science increased over the last ten years, leading it to be considered Big Data. Thus, according to the PubMed database, the number of papers in the field of medicine and biology over the last year amounted to more than a million. MeSH indexing terms (the main headings and subheadings) are served as a rich resource for extracting a broad range of domain knowledge. Automated analysis and comparison of MeSH terms of the papers allowed the formation of groups of relevant papers in accordance with user-defined criteria and highlighting key concepts within the groups of papers, expressed in the form of relationship between MeSH or other concepts.

Using text-mining algorithm in automatic mode, publications available in the PubMed retrieved by the keywords "*Precision medicine*", "*Personalized medicine*" and "*Digital medicine*" were analyzed for decipher the trends of scientific articles in the field of personalized medicine over the past. The simple idea of using text-mining tools allows users to form concept-centered semantic networks (maps) based on real-time Pubmed-available knowledge. Networks represent relationships between various objects: genes (proteins), MeSH, chemical compounds, diseases, etc. Semantic networks were visualized using Cytoscape according to the matrix of similarity, and the distance between the nodes (concepts) was correlated with the normalized number of scientific articles. Analysis of changes in time the concept-centered semantic networks allows us to find the main trends and key concepts in a given subject area.

Acknowledgement. This work supported by RFBR grant (#19-29-01138).

1. Ilgisonis E, Lisitsa A, Kudryavtseva V, Ponomarenko E. Creation of Individual Scientific Concept-Centered Semantic Maps Based on Automated Text-Mining Analysis of PubMed. Adv Bioinformatics. 2018 Jul 26;2018:4625394. doi: 10.1155/2018/4625394. eCollection 2018.

784

CPPD – a rapid proteolytic method enables complete protein sequence coverage in bottom-up proteomics

Irene Ling ANG[1], Melody Man Ting LAM[1], Kate Man Kei LEI[1], Rui WEI[1], Pengwei Zhang ZHANG[1], Terence Chuen Wai POON[1]

1. Faculty of Health Sciences,, University of Macau, Macau, China, Taipa, Macau, China

In bottom-up proteomics, incomplete protein sequence coverage leading to uncertainty about sequence information and post-translational modifications (PTMs) within the coverage gaps. Complete sequence coverage is particular important for differentiating proteoforms. Unfortunately, a simple and robust method for obtaining 100% sequence coverage has not yet been available, no matter whether the approach is bottom-up, middle-down or top-down. Here we report a rapid (≤1 hour) proteolytic method (termed CPPD) for generation of overlapping peptides for LC-MS/MS analysis. This method capacitates 100% protein sequence coverage, sequence variant identification and comprehensive modification characterization for analyses of single

proteins in solutions or in gel bands, simple protein mixtures and high abundance proteins in complex protein mixtures. As observation of overlapping peptides could reduce ambiguity in mapping a covalent modification to a specific amino acid residue, we applied the CPPD-based one-shot LC-MS/MS to examine the PTMs of bovine asialofetuin and a theranostic antibody 8H9. In addition to the known phosphorylation sites, glycosylation sites, *N*-linked glycans and *O*-glycans, a novel *O*-glycosylation site and several previously unreported *N*- and *O*-linked glycans were discovered in asialofetuin. For the 8H9 antibody, N-terminal microheterogeneity, N-terminal glutamine to pyroglutamine conversion and C-terminal lysine deletion were detected in the heavy chain. Besides, known *N*-linked glycans and a rarely reported glycan were identified. As the 8H9 antibody is a theranostic antibody, the value of the CPPD approach in quality control of biotherapeutic/theranostic protein products was demonstrated. This research was funded by the Science and Technology Development Fund (FDCT) of the Macau SAR Government (Reference Number: 0102/2018/A3) and by the Faculty of Health Sciences, University of Macau.

785

AD and MCI associated changes in blood plasma proteome studied by high resolution mass spectrometry

Maria Indeykina[1], Igor Popov[2, 1], Anna Bugrova[1], Natalia Zakharova[1], Alexey Kononikhin[1], Svetlana Gavrilova[3], Eugene Nikolaev[4]

1. Emanuel Institute of Biochemical Physics RAS, Moscow, Russia

2. Mowcow Institute of Physics and Technology, Dolgoprudny, MOSCOW REGION, Russia

3. Mental Health Research Center (MHRC), Moscow, Russia

4. Skolkovo Institute of Science and Technology, Moscow, Russia

Introduction

Alzheimer's disease(AD) is the most common socially significant neurodegenerative pathology worldwide. Since existing therapy methods are low-effective, development of reliable early diagnostics approaches is one of the global tasks. Increasing attention is being paid to AD-associated blood plasma proteome changes, including post-translational modifications (PTMs) studies. Mass spectrometry (MS) approaches seem to be the most promising tool due to high sensitivity, specificity, and multiplexing capacity.

Methods

Plasma samples were collected from three groups of patients – AD (16), mild cognitive impairment (MCI, 20) mentally healthy elders (14). Plasma was subjected to abundant protein depletion, denaturation, filtration, alkylation and trypsinolysis (16 hours, 37C). Samples were analyzed by LC-ESI-MS/MS using nano-LC Agilent 1100 system coupled to 7T LTQFT Ultra tandem high-resolution mass-spectrometer.

Results and Conclusions

Significant changes in plasma levels of 22 proteins were observed, including 14 whose dysregulation was previously shown to be associated with AD/MCI in studies on CSF and postmortem brains. Obtaining consistent results on different cohorts using different methods is an extremely important stage for validation of pathology markers and one of the central tasks for determination of biomarkers specific for AD and MCI. Analysis PTMs distribution showed significant changes correlating with pathologies in modification levels even for those proteins whose expression levels remain unchanged. Since early AD diagnostics methods remain a hot issue, obtained results may be important for creation of new MS-based diagnostics methods.

786

The study of human plasma proteome changes under the influence of spaceflight factors

Alexander G Brzhozovskiy[1], Alexey S Kononikhin[1, 2], Maria I Indeykina[3], Daria N Kashirina[1], Igor A Popov[4], Ludmila Ch Pastushkova[1], Marc-Antoine Custaud[5], Irina M Larina[1], Evgeny N Nikolaev[3, 2]

1. Russian Academy of Sciences, Institute of Biomedical Problems – Russian Federation State Scientific Research Center , Moscow, Russia

2. Skolkovo Institute of Science and Technology, Moscow, Russia

3. Russian Academy of Sciences, Emanuel Institute for Biochemical Physics, Moscow, Russia

4. Moscow Institute of Physics and Technology, Moscow, Russia

5. Angers University, Angers, France

The aim of the study was to compare proteomic data on the effects of spaceflight factors on the human body, including both real space missions and ground-based experiments.

LC–MS/MS-based proteomic analysis of blood plasma samples obtained from 13 cosmonauts before and after long-duration (169–199 days) missions on the International Space Station (ISS) and for five healthy men included in 21-day-long head-down bed rest (HDBR) and dry immersion experiments were performed. The samples were prepared via the filter-aided sample preparation (FASP) protocol using 10 kDa filters. The tryptic peptide fraction was analyzed in triplicate on a nano-HPLC Dionex Ultimate3000 system coupled to a MaXis 4G using a nanospray ion source (positive ion mode, 1600 V). HPLC separation was performed on a C18 capillary column (75 µm x 50 cm, C18, 3 µm, 100 A) at a flow rate of 0.3 µL/min by gradient elution. The mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in acetonitrile. The separation was carried out by a 120 min gradient from 3% to 90% of phase B.

Semi-quantitative label-free analysis revealed significantly changed proteins: 19 proteins were significantly different on the first (+1) day after landing with respect to background levels; 44 proteins significantly changed during HDBR and 31 changed in the dry immersion experiment. Comparative analysis revealed nine common proteins (A1BG, A2M, SERPINA1, SERPINA3, SERPING1, SERPINC1, HP, CFB, TF), which changed their levels after landing, as well as in both ground-based experiments. Common processes, such as platelet degranulation, hemostasis, post-translational protein phosphorylation and processes of protein metabolism, indicate common pathogenesis in ground experiments and during spaceflight. Dissimilarity in the lists of

significantly changed proteins could be explained by the differences in the dynamics of effect development in the ground-based experiments.

787

MetaNovo: a probabilistic pipeline for peptide and polymorphism discovery in complex mass spectrometry datasets

Matthys MG Potgieter[1, 2], Andrew AJM Nel[2], David DL Tabb[2, 3, 4, 5], Suereta S Fortuin[2], Shaun S Garnett[2], Jerome JM Wendoh[6], Katie K Lennard[1], Jonathan JM Blackburn[2], Nicola NJ Mulder[1]

1. Computational Biology Division, Department of Integrative Biomedical Sciences, IDM, University of Cape Town, Cape Town, South Africa

2. Division of Chemical and Systems Biology, Department of Integrative Biomedical Sciences, IDM, University of Cape Town, Cape Town, South Africa

3. Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Stellenbosch University, Cape Town, South Africa

4. Bioinformatics Unit, South African Tuberculosis Bioinformatics Initiative, Stellenbosch University, Cape Town, South Africa

5. The African Microbiome Institute, The African Microbiome Institute, Stellenbosch University, Stellenbosch, South Africa

6. Division of Immunology, Department of Pathology, IDM, University of Cape Town , Cape Town, South Africa

The characterization of complex mass spectrometry data obtained from metaproteomics or clinical studies presents unique challenges and potential insights in the pathogenesis of human disease. Previous approaches essentially rely on prior expectation or knowledge of likely sample composition in order to construct focussed search libraries, but this is potentially limiting in many cases. Here we present a novel software pipeline to directly estimate the proteins and species present in complex mass spectrometry samples at the level of expressed proteomes, using *de novo* sequence tag matching and probabilistic optimization of very large sequence databases prior to target-decoy search. We validated our pipeline against the results obtained from the recently published **MetaPro-IQ** (Zhang et al., 2016) pipeline on 8 human mucosal-luminal interface samples, with comparable numbers of peptide and protein identifications being found. We then showed that using an unbiased search of the entire release of **UniProt** (ca. 90 million protein sequences) MetaNovo was able to identify a similar bacterial taxonomic distribution compared to that found using a small, focused matched metagenome database, but now also simultaneously identified proteins present in the samples that are derived from other organisms that are missed by 16S or shotgun sequencing and by previous metaproteomic methods**.** Using **MetaNovo** to analyze a set of single-organism human neuroblastoma cell-line samples (*SH-SY5Y*) against **UniProt** we achieved a comparable MS/MS identification rate during target-decoy search to using the **UniProt** human Reference proteome, with 22583 (85.99 %) of the total set of identified peptides shared in common. Taxonomic analysis of 612 peptides not found in the canonical set of human proteins yielded 158 peptides unique to the *Chordata* phylum as potential human variant identifications. Of these, 40 had previously been predicted and 9 identified using whole genome sequencing in a proteogenomic study of the same cell line.

1. Zhang, X., Ning, Z., Mayne, J., Moore, J. I., Li, J., Butcher, J., … Figeys, D. (2016). MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota. Microbiome, 4(1), 31. https://doi.org/10.1186/s40168-016-0176-z

788

A multi-instrument, longitudinal assessment of high-throughput proteomics using 1,560 SWATH-MS profiles from standardised cancer tissue samples

Rebecca C Poulos[1], Peter G Hains[1], Rohan Shah[1], Natasha Lucas[1], Dylan Xavier[1], Sadia Mahboob[1], Max Wittman[1], Jennifer MS Koh[1], Steven G Williams[1], Srikanth Manda[1], Michael Hecker[1], Asim Anees[1], Michael Dausmann[1], Rosemary Balleine[1], Jean Yang[2], Terence P Speed[3, 4], Brett Tully[1], Yansheng Liu[5, 6], Roger Reddel[1], Phillip J Robinson[1], Qing Zhong[1]

1. ProCan®, Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

2. School of Mathematics and Statistics, The University of Sydney, Sydney, NSW, Australia

3. Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

4. Department of Mathematics and Statistics, University of Melbourne, Melbourne, VIC, Australia

5. Department of Pharmacology, Yale University School of Medicine, New Haven, CT, USA

6. Yale Cancer Biology Institute, Yale University, West Haven, CT, USA

SWATH mass spectrometry (MS)-based proteomics is a valuable tool for biomedical research. A selection of landmark studies have examined high-throughput SWATH-MS. However, industrial-scale reproducible proteomic measurements across multiple instruments, over time in a single laboratory, have not been investigated. To this end, with a cautious study design, we acquired 1,560 SWATH-MS runs using six SCIEX 6600 TripleTOFs. These instruments operated in harmony at ProCan over a four-month period, collecting approximately 5,500 additional MS runs in the interim to reflect a real-world scenario. Our experimental samples were a dilution series containing prostate cancer in fixed proportion (50%), with a variable fraction of ovarian cancer (3.125% - 50%) offset by yeast. We identified 6,865 proteins in a combined spectral library generated from pooled DIA runs searched with Mascot, X!Tandem and MSGF+. Our proteomic data were processed using OpenSWATH, with PyProphet for FDR-control. The median unnormalised coefficient of variation measured within a single instrument during the first experimental week approximated 10%. We applied RUV-III (Remove Unwanted Variation) for experiment-wide normalisation to correct for machine, temporal and unknown batch effects. We then evaluated and demonstrated the benefits of imputing missing values with non-missing measurements from technical replicates spanning multiple instruments. After normalisation, we observed a strong linear relationship between the ovarian tissue proportion in a sample and intensity of ovarian-specific peptides. Moreover, we assessed

the sample sizes required to overcome technical variation in order to identify significantly different peptide intensities across dilutions. Finally, we applied machine learning to predict ovarian tissue concentration in each sample with high accuracy. We have produced large-scale SWATH-MS data, obtained over time across multiple instruments with varying maintenance schedules. We establish for the first time, that such measurements can be effectively integrated and analysed using appropriate statistical methods, to enable the reproducible and high-throughput proteomics required for precision medicine.

789

Focus on Gene Editing for human protein annotation

Ekaterina Poverennaya[1], Leonid Kurbatov[1], Petr Sergiev[2], Andrey Lisitsa[1]

1. Institute of biomedical chemistry, Moscow, MOSCOW, Russia

2. Lomonosov Moscow State University, Moscow, Russia

Gene editing based on CRISPR-Cas9 system opens up broad possibilities for the study of the genome and its' transfer to phenotype. The changes of gene will be reflected on the expression and translation processes. Therefore, in case of gene editing we really work with three levels of realization of genomic information - genome, transcriptome and proteome - which allows us not only to describe the function of target gene but also to approach understanding of all cell processes' details.

The scheme of CRISPR-Cas9 system is described as delivering to cell the guide RNA-Cas9 complex via plasmid where guide RNA is complementary to target DNA. The DNA target must be adjacent to a short stretch of sequence termed PAM that is met in DNA approximately every 20-30 nucleotides. And Cas9 endonuclease causes a double strand break at the target site. And in the moment of put the native system of DNA reparation we have an opportunity to add something new to target gene.

Our idea is to add a tag to the gene to further extract the corresponding protein. Using tagged protein is desirable to fish out interacting partners, and it gives hints on protein function. Such hints could be followed by focused functional screens to decipher biological role and activity of proteins under study.

We developed an AP-MS protocol modified with gene editing for HA-tag addition as applied to several protein with uncertain function encoded by genes of human Chr18. For this proteins protein-protein interaction data were obtained in HEK293 cell line under different conditions. As result we showed that this approach is promising for real human interactome construction and protein's annotation due to its' "ecological compatibility" for cell and repeatability.

This work was supported with RSF grant #18-74-00144

79

How Does Japan’s Agency for Medical Research and Development fund proteomic research?

Tesshi Yamada[1]

1. National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, TOKYO, Japan

-

790

Cellular Thermal Shift Assay (CETSA): From drug targets to horizontal cell biology

Nayana Prabhu[1], Lingyun Dai[1], Tianyun Zhao[1], Ying Yu Liang[1], Wendi Sun[1], Pär Nordlund[1]

1. Nanyang Technological University, Sinagpore, SINGAPORE, Singapore

The Cellular Thermal Shift Assay i.e. CETSA is based on the biophysical principle of ligand-induced thermal stabilization of proteins. At inception about 5 years ago, CETSA was considered a tool for drug-target deconvolution and mainly used to study direct target engagement *in situ*. Using quantitative mass spectrometry, proteome-wide CETSA (MS-CETSA) was established, which allows off-target binding as well as downstream effects to be discovered, in addition to direct drug-target interactions. MS-CETSA has since then been used widely in the pharmaceutical industry for such studies and is currently a valuable tool for biomedical research and drug development.

Today, MS-CETSA has evolved to encompass several other applications including novel ways to study cell biology. With the advent of a highly sensitive multidimensional implementation of MS-CETSA, comprehensive studies of protein level changes as well as stability changes are plausible. In addition, MS-CETSA can also be used to access binding of physiological ligands to proteins, such as metabolites, nucleic acids and other proteins. Such studies will aid in understanding “Horizontal Cell Biology” where modulations of protein interaction states (PRINTs) are studied (as opposed to the “vertical information” focusing on the protein and RNA levels) to provide new insights on cellular functions and processes.

791

Nuclear signaling and CHO proteome altered after retinoic acid and sodium butyrate treatment

Antonia Pries[1], Louise Schelletter[1], Thomas Noll[1], Raimund Hoffrogge[1]

1. Institute of Cell Culture Technology, Bielefeld University, Bielefeld, Germany

Chinese hamster ovary (CHO) cells are a preferred expression system for biopharmaceuticals like monoclonal antibodies (mAb). Sodium butyrate (NaBu) is known for increasing the cell specific productivity by altering product gene expression via epigenetic changes. Similar effects were recently found for retinoic acid (RA) [1]. Therefore, unraveling specific underlying signaling events and variations on proteome level will provide a new basis for future engineering of production processes.

MAb producing CHO cells were cultivated in SILAC medium (light, medium, heavy lysine and arginine). 1 h and 24 h after agent addition (150 nM RA or 2 mM NaBu) cells were harvested for phosphoproteome as well as proteome analysis. A nuclear extraction was followed by a tryptic digest and SIMAC phosphopeptide enrichment. The obtained samples were analysed by nLC-ESI-orbitrap MS measurements. Finally, data evaluation was performed by Proteome Discoverer and in-house Fusion software.

Addition of NaBu resulted in increased productivity linked to decreased growth as observed with other CHO lines before [2,3]. Similar effects for RA were found accompanied by a 23 % higher product concentration. Statistical evaluation of MS data resulted in 256 significantly regulated phosphopeptides after NaBu and 407 by RA treatment. The data reveal a changed phosphorylation pattern of a variety of histone modifying enzymes like methyl- or acetyltransferases such as Suz12 and KAT8. These findings implicate also putative differences in transcriptional regulation of e.g. product genes. On proteome level almost 2000 nuclear proteins were quantified and 355 (NaBu) or 243 (RA) differentially expressed. KEGG-mapping revealed a broad coverage of RNA transport and spliceosome pathways.

The impact of both agents on CHO cells resulted in a temporary increased cell specific productivity. Retinoic acid induced changes in nuclear signaling and phosphorylation pattern of histone modifying enzymes similar to sodium butyrate next to effects on other interesting pathways like RNA transport.

1. Rahimi-Zarchi, M., Shojaosadati, S.A., Amiri, M.M., Jeddi-Tehrani, M., Shokri, F. (2018). All-trans retinoic acid in combination with sodium butyrate enhances specific monoclonal antibody productivity in recombinant CHO cell line. Bioprocess and Biosystems Engineering 41 (7), 961–971.
2. Wippermann, A., Rupp, O., Brinkrolf, K., Hoffrogge, R., Noll, T. (2017). Integrative analysis of DNA methylation and gene expression in butyrate-treated CHO cells. Journal of biotechnology 257, 150 161.
3. Müller, B., Heinrich, C., Jabs, W., Kaspar-Schonefeld, S., Schmidt, A., Rodrigues de Carvalho, N., Albaum, S.P., Baessmann, C., Noll, T., Hoffrogge, R. (2017). Label-free protein quantification of sodium butyrate treated CHO cells by ESI-UHR-TOF-MS. Journal of biotechnology 257, 87 98.

792

Systematic analysis of nuclear deubiquitylases – insight into BRCC3-regulated signaling

Gabriela Prus[1], Shanka Sapathy[1], Magdalena Budzowska[1], Thomas Wild[1], Chuna Ram Choudhary[1]

1. Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen N, DK, Denmark

BACKGROUND: Ubiquitination is intimately involved in key nuclear processes such as gene transcription, DNA replication, and DNA damage repair. Therefore, many nuclear deubiquitylating enzymes (DUBs) play important roles in regulating these processes. For example, the deubiquitylase BRCC36 (BRCC3) is part of the BRCA1 (breast cancer type 1 susceptibility protein) complex, which includes BRCA1, Rap80, MERIT40, BRE, and the scaffolding protein Abraxas. The BRCA1 complex plays an important role in DNA double strand-break repair and facilitates error-free repair through homologous recombination. It binds to ubiquitylated histones H2A and H2AX at DNA lesion, promoting DNA end resection. Intriguing, BRCC36 is thought to prevent exceeded DNA resection via counteracting ubiquitylation of H2A. However, whether BRCC36 plays another role in BRCA1 complex as well as its site-specificity remains unclear, leaving a considerable gap in understanding the relationship between the DUB and its substrates.

METHODS: Here, we apply a combination of CRISPR-based genome engineering and quantitative proteomics to obtain a systematic and comprehensive picture of DUB-regulated signaling in the nucleus on the endogenous level. Our multi-step workflow provides a new depth to ubiquitylome studies and allows for comprehensive analysis of DUB site-specificity.

RESULTS AND CONCLUSIONS: In our ongoing work we have generated KO cell lines for 8 nuclear DUBs, including BRCC36-KO cell line. Using SILAC-based proteomics, we mapped ubiquitylation sites regulated upon BRCC36 deletion. Our preliminary data shows the significant decrease in several ubiquitylation sites, including the components of BRCA1 complex: MERIT40 and BRE. These preliminary results indicate a role of BRCC36 in regulating ubiquitylation of the BRCA1 complex members.

793

The Impact of Asialylated Glycoproteins in Coronary Artery Disease.

Benjamin J Pullen[1, 2, 3], Achini K Vidanapathirana[1, 2, 3], Jarrad M Goyne[1, 2, 3], Stephen J Nicholls[4, 2], Nicolle H Packer[2, 5], Christina A Bursill[1, 2, 3]

1. School of Medicine, University of Adelaide, Adelaide, South Australia, Australia

2. Australian Research Council (ARC) Centre of Excellence for Nanoscale BioPhotonics (CNBP) , Adelaide, South Australia, Australia

3. Vascular Research Centre, South Australian Health and Medical Research Institute (SAHMRI), Adelaide, South Australia, Australia

4. MonashHeart, Monash University, Clayton, Victoria, Australia

5. Department of Biomolecular Design and Discovery Centre, Macquarie University, Sydney, New South Wales, Australia

Coronary artery disease (CAD) is the build-up of fatty plaque within the walls of the blood vessels within the heart, resulting in reduced blood flow to the heart tissue. Macrophages are cells that are key to the build of fatty atherosclerotic plaque. A recent landmark study has revealed that the presence of a rare loss-of-function mutation in a glycoprotein receptor, ASGR1, significantly correlates with a reduction (34%, p=4.0×10−6) in the risk of CAD. The dramatic reduction in CAD risk could not be explained purely by the reduction in plasma cholesterol levels, suggesting that other mechanisms are involved. This project aims to determine whether: 1) ASGR1 is expressed in macrophages, 2) expression changes with macrophage polarity and 3) ASGR1 plays a role in important macrophage functions.

Bone marrow cells weredifferentiated from monocyte-like cells to bone marrow-derived macrophages (BMDMs) over seven days. BMDMs were also polarised to either an ‘M1-like’ inflammatory or ‘M2-like’ anti-inflammatory macrophage phenotype. Samples

were collected throughout the differentiation and polarisation process for analysis of mRNA using quantitative PCR and protein using Western blotting. We found that ASGR1 is expressed in macrophages at both the mRNA and protein level. ASGR1 expression increases with differentiation from a monocyte to a macrophage. There were no changes in ASGR1 levels between the M1 and M2-like phenotype.

We have demonstrated, for the first time, the presence of ASGR-1 in BMDMs, highlighting a possible new role for ASGR1 in atherosclerosis. Future studies will compare changes in the function and glycoprotein profile of BMDMs from ASGR1$^{-/-}$and wildtype mice *ex vivo*. These studies will determine if ASGR1 represents a new therapeutic target for the prevention of CAD.

794

The age-dependent proteome landscape of platelet in healthy population

Yueqin Zhu[1], Roy Bhaskar[1], Yanbin He[1], Da Qi[1], Siqi Liu[1]

1. BGI-Shenzhen, Shenzhen, GUANGDONG, China

It is generally accepted that blood platelet plays a critical function during aging, while some bio-markers were reported indicative to age. Herein, we present a global analysis of platelet proteins towards 45 normal samples, 18 males and 27 females, whose ages span from 22 to 64 years old. The platelets were well prepared, and the corresponding proteins were extracted followed by a completely tryptic digestion. The digested peptides were delivered to Q Exactive™ HF hybrid quadrupole-orbitrap™ mass spectrometer, and the acquired MS/MS signals were treated with Spectronaut for protein identification and quantification. The total of 3703 platelet proteins identified and quantified were divided into three groups on basis of principal component analysis. Intriguingly, the groups contained the samples with clear cut-off of age, 22-34 (young), 35-54 (middle) and 55-64 (old), respectively. Of these proteins, 417 proteins were found their abundance as age-dependent, and were further hierarchically divided to four clusters, in which the samples in cluster 1 and 4 belonged to young group, while that in cluster 2 and 3 fit in old group. The cluster analysis demonstrated some platelet proteins closely associated with aging. Pathway enrichment to all the age-dependent proteins revealed a wide functional categories, while their functions or pathways were reported as age-involvement, such as vesicle-mediated transport, oxidative phosphorylation, Parkinson disease, and Huntington disease. Besides, with multivariate analysis several proteins could be defined a proteins panel that clearly indicated age.

795

Impact of proteomic analyses for understanding cellular responses to nanoparticles: toward mechanistic data and evidence for cross-toxic effects

Thierry Rabilloud[1], Bastien Dalzon[1], Hélène Diemer[2], Sarah Cianférani[2]

1. Chemistry and Biology of Metals, UMR CNRS-CEA-UGA 5249, Grenoble, RA, France

2. IPHC/LSMBP, Strasbourg, Alsace, France

Nanoparticles are recently-introduced industrial products which impact on the environment and on human health is intensely debated, due in part to their long persistence in vivo. Most of the toxicological research carried to date on nanoparticles follows very classical tracks, so that little is known yet on their mechanisms of toxicity. Furthermore, classical toxicology on laboratory animals cannot mimick the high diversity of situations encountered in human populations, so that mechanistic data are needed to make predictions that could be applied in human toxicology of nanoparticles. Because of their major scavenging function and their role in inflammation, macrophages are one of the major cell types to study when working on nanoparticles. In order to gain new insights on the interaction of nanoparticles with macrophages, we carried out a proteomic screen on cells treated with different nanoparticles (e.g. copper oxide, zinc oxide, metallic silver or silica). Very different responses have been observed : Copper oxide induces mainly a mitochondrial response and a strong response at the glutathione level. Zinc oxide rather induces a proteasomal and a metabolic response (glucolysis and pentose phosphate pathways). Proteomics was followed by targeted studies aiming at validating the proteomic results and at investigating major functions of macrophages (e.g. phagocytosis and cytokine production). In this way, we could demonstrate a role of DNA repair pathways for zinc oxide and silica nanoparticles, as well as a critical role of glutathione and heme oxygenase for survival to a copper oxide challenge. Mitochondrial dysfunction is also prominent following exposure to zinc oxide or copper oxide, and much less important following exposure to silver or silica. Different changes in the actin cytoskeleton were also observed in response to the nanoparticles tested. In conclusion, this combined approach has provided new ideas on how nanoparticles can exert their toxic effect.

796

About the noisiness and predictive value of proteomic data : lessons from enzymes

Thierry Rabilloud[1]

1. Chemistry and Biology of Metals, UMR CNRS-CEA-UGA 5249, Grenoble, RA, France

Since the beginning of proteomics, the quality of the data has been an important question. Furthermore, the predictive value of omics data is also an important point to be taken into account for advancing biology. To investigate these questions, a model study on macrophages submitted to treatment by copper ion complexed or not with polyacrylate has been used. In this study, the same protein extracts were used in two different proteomic setups, i.e. 2D-gel based and label-free shotgun proteomics.

The quality of the data was investigated by using several metrics, including clustering techniques. While the 2D gel data produced a clustering tree reflecting the expected biological situation, the shotgun data did not produce a consistent grouping, showing that the noise of the data was not negligible compared to the biological signal.

To address the even more important question of the positive predictive value, a "gold standard" is required. The fact that several enzymes appeared as modulated in the various proteomic setups allowed testing the predictive value against the enzyme activities, used as the gold standard. In this dataset, at a cutoff value of $p<0.05$, the predictive value of the 2D gel data was slightly

higher than 40%, while the one of the shotgun data was <20%. It increased at 25% with a cutoff value of p<0.01, and further filtering with the fold change did not induce any improvement, but only a loss in the true positives detected. I the same trend, the calculated FDRs for individual enzymes were not different between the true and false positives.

Although these figures are expected to be variable from one dataset to another, they show that the predictive value of the proteomic data is a figure to be investigated, and that simple enzyme assays are an interesting tool to investigate this parameter.

797

Breast implant capsular contracture reveals changes in the regulation of low-abundance plasma proteins using TMT-based quantitative MS

Md Arifur Rahman[1], Ardeshir Amirkhani[2], Maria Mempin[1], Honghua Hu[1], Alamgir Khan[2], Dana Pascovici[2], Charlie Ahn[1], Anand Kumar Deva[1], Mark S. Baker[1], Karen Vickery[1]

1. Department of Biomedical Sciences, Macquarie University, Sydney, NSW, Australia

2. Australian Proteome Analysis Facility, Sydney, NSW, Australia

Background

Capsular contracture (CC) is one of the most common post-operative complications among breast-implant associated infections. However, the mechanisms that lead to CC are still poorly understood. Therefore, we employed depletion of high-abundance plasma proteins followed by TMT-based quantitative mass spectrometry (MS) to construct the plasma proteome in healthy controls and patients with biofilm-related breast-implant CC to explore disease-associated alterations of plasma proteins.

Methods

Plasma from each patient (10 healthy control, 10 CC) was processed for the depletion of high-abundant proteins using immune affinity-based depletion method (MARS-14) to improve the depth of detection in plasma samples. Protein extraction, fractionation, reduction, alkylation, digestion, and 10-plex Tandem Mass Tag (TMT) labeling steps were performed, respectively. TMT-based MS was performed, and protein identification and relative quantitation of protein levels were analysed using Proteome Discoverer (version 2.1). Statistical analysis was done using the TMTPrePro R package.

Results

A total of 450 proteins were identified from these samples. Among the significant differentially expressed proteins tropomyosin alpha-4, talin-1, transferrin receptor protein 1, lipopolysaccharide-binding and fibrinogen alpha chain were exclusively upregulated whereas cartilage intermediate layer protein 2, bone marrow proteoglycan, complement component C7, neural cell adhesion molecule 2, etc were downregulated in the breast implant CC patients as compared to healthy controls. Interestingly, we found a correlation with the significant upregulation of tropomyosin alpha-4, increased myofibroblast activity, and the regulatory mechanism by TGF-β which demonstrated a correlation with bacterial biofilms. Further pathway analysis revealed an inflammatory response, focal adhesion, platelet activation, complement and coagulation cascades, as enriched pathways identified in this study.

Conclusions

This is the first report using high throughput TMT-based MS in the plasma of patients with breast implant CC. The identified differentially expressed proteins has the potential to provide important information for future mechanism studies and in the development of breast implant CC biomarker/s.

798

Exosomes derived from human primary and metastatic colorectal cancer cells contribute to functional heterogeneity of activated fibroblasts by reprogramming their proteome

Alin Rai[1], David W Greening[1, 2], Richard J Simpson[2]

1. Baker heart and diabetis instutitbe, Melbourne, VIC, Australia

2. La Trobe University, Melbourne, VIC, Australia

Cancer-associated fibroblasts (CAFs) are a heterogeneous population of activated fibroblasts that constitute a dominant cellular component of the tumor microenvironment (TME) performing distinct functions. Here, the role of tumor-derived exosomes (Exos) in activating quiescent fibroblasts into distinct functional subtypes is investigated. Proteomic profiling and functional dissection reveal that early- (SW480) and late-stage (SW620) colorectal cancer (CRC) cell-derived Exos both activated normal quiescent fibroblasts (α-SMA , CAV , FAP , VIM) into CAF-like fibroblasts (α-SMA , CAV , FAP , VIM). Fibroblasts activated by early-stage cancer-exosomes (SW480-Exos) are highly pro-proliferative and pro-angiogenic and display elevated expression of pro-angiogenic (IL8, RAB10, NDRG1) and pro-proliferative (SA1008, FFPS) proteins. In contrast, fibroblasts activated by late-stage cancer-exosomes (SW620-Exos) display a striking ability to invade through extracellular matrix through upregulation of pro-invasive regulators of membrane protrusion (PDLIM1, MYO1B) and matrix-remodeling proteins (MMP11, EMMPRIN, ADAM10). Conserved features of Exos-mediated fibroblast activation include enhanced ECM secretion (COL1A1, Tenascin-C/X), oncogenic transformation, and metabolic reprogramming (downregulation of CAV-1, upregulation of glycogen metabolism (GAA), amino acid biosynthesis (SHMT2, IDH2) and membrane transporters of glucose (GLUT1), lactate (MCT4), and amino acids (SLC1A5/3A5)). This study highlights the role of primary and metastatic CRC tumor-derived Exos in generating phenotypically and functionally distinct subsets of CAFs that may facilitate tumor progression.

799

Turnover optimized short nanoLC gradients on a tims equipped QTOF for high throughput and deep proteome measurements

Thomas Kosinski[1], Markus Lubeck[1], Scarlet Koch[1], Pierre-Olivier Schmit[2], Adam Rainczuk[3], Thorsten Lederteil[1], Christian Meier-Credo[1], Christoph Gebhard[1], Heiner Koch[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Bruker France, Wissembourg, France

3. BRUKER Pty. LTD, Preston, Australia

Introduction

The analysis of large sample cohorts requires regular QC injections to monitor the instrument's performance over the entirety of the experiment. As throughput concerns grow with the size of the sample cohort, we have developed a QC method that relies on the additional separation power and high sequencing speed of Parallel Accumulation Serial Fragmentation (PASEF) to deliver reproducible and quantitative information from a high-complexity QC sample despite a short injection to injection time.

Methods

HeLa peptide digest was dissolved to 250 ng/ µL, where 1 µL injection volume was delivered to a 100 mm fritted column of ID 75 µm (Bruker). The nano LC business logic was modified in a way that turnover times were reduced to achieve a 21.4 min gradient time and 28.8 min inject to inject time. Separation was performed at 500 nl/ min and loading was performed at a max pressure of 500 bar. This method allows up to 50 samples per day to be measured, maximizing throughput. PEAKS studio (Bioinformatics Solutions Inc.) was used for data processing and results were corrected to 1% PSM FDR.

Results

Using the optimized 28.8 min method, more than 4000 protein groups could be identified from 250 ng of a proteolytic digest of a human cancer cell line (HeLa). In parallel, offline LC-fractionation was performed on a HeLa digest and digests of murine cerebellum, and the fractions analyzed using the same column and method. Samples were fractionated on high pH reversed-phase columns into 24 concatenated samples. Subsequent measurement allowed the identification of more than 100000 unique peptides and 9052 protein groups in less than 12 h of measurement time.

Conclusions

This method allows to perform rapid QC from complex samples and expedite the assembly of large spectral libraries from fractionated samples.

8

Proteome-wide systems genetics to interrogate metabolism

Benjamin L Parker[1, 2], Marcus M Seldin[3], Anna C Calkin[4], Michael F Keating[4], Elizabeth J Tarling[3], Pengyi Yang[2], Sarah C Moody[4], Yingying Liu[4], Eser J Zerenturk[4], Elise J Needham[2], Matthew L Miller[3], Bethan L Clifford[3], Pauline Morand[3], Matthew J Watt[1], Ruth C Meex[1], Kang-Yu Peng[4], Richard Lee[5], Kaushala Jayawardana[4], Calvan Pan[3], Natalie A Mellett[4], Jacquelyn M Weir[4], Ross Lazarus[4], Peter J Meikle[4], Thomas W Vallim[3], Brian G Drew[4], Aldons J Lusis[3], David E James[2]

1. University of Melbourne, Melbourne, VIC, Australia

2. University of Sydney, Sydney, NSW, Australia

3. University of California Los Angeles, Los Angeles, CA, USA

4. Baker Heart & Diabetes Institute, Melbourne, VIC, Australia

5. Ionis Therapeutics, Carlsbad, CA, USA

Genetic reference panels (GRPs) using model organisms have become a more tractable way of studying the influence of genetics and environment on complex metabolic traits. Unlike studies in humans, GRPs allow for accurate control of environment as well as access to critical metabolic tissues. Importantly, systems genetic integration with intermediate phenotypes such as proteomic and lipidomic analysis of such tissues facilitates the discovery of previously unknown linkages between several layers of molecular information. Here, we present our latest systems genetic analysis of the Hybrid Mouse Diversity Panel involving the integration of genomics with proteomics and lipidomics analysis in liver from >100 strains of mice. This revealed functional protein and genetic variants that modulate pathological lipid abundance including the validation of PSMD9 as a previously unknown lipid regulatory protein. To further understand how genetic variants influence potential adaptations to the environment, we have also integrated the quantification of lysine acetylation modifications across >70 strains of mice. A *trans*-quantitative trait loci analysis has identified precise SNPs that are associated with modification sites. Integration of these data with established mouse GWAS has revealed novel causative mutations effecting histone modifications and whole-body metabolic traits. Associating natural variations in the abundance of PTMs across a GRP to phenotypic measurements is a powerful approach to pinpoint functional modifications influencing gene expression and metabolism.

800

Exploiting 4D in Omics: Mass Offset Mobility Aligned (MOMA) and Mass Aligned Mobility Offset (MAMO) of Biological Analytes

Christopher M Adams[1], Michael Krawitzky[1], Guillaume Trementin[1], Adam Rainczuk[2], Pierre-Olivier Schmit[3], Gary Kruppa[4]

1. Bruker Daltonics Inc, San Jose, CA, USA

2. BRUKER Pty. LTD, Preston, Australia

3. Bruker France, Wissembourg, France

4. Bruker Daltonics Inc, Billerica, MA, USA

Small molecules, lipids, glycans and peptides are just a few of the classes of compounds, represented by vast diversity, that are measured in “OMXs” studies by mass spectrometers. By definition, the mass spectrometer measures molecular mass , and as such is incognizant to physiochemical properties that differentiate such things, as for example peptide isobars. Trapped Ion Mobility Spectrometry (TIMS) has the unique capability of separating gaseous ions by mobility (CCS) prior to injection into the mass spectrometer . This additional dimension (CCS) coupled to a sensitive and fast scanning TOF affords deconvolution of analytes whose molecular mass is different but CCS (mobility) is the same, as is the case for SILAC, mTRAQ and di-methyl labeled peptides. We describe this principle with the acronym MOMA- Mass Offset Mobility Aligned where the precise delta mass of peptide labeled pairs co-eluting in time give rise to the same CCS term. Utilizing the delta mass shift in combination with the same retention time is a methodology previously describe as TOMAHAQ[a] (Triggered by Offset, Multiplexed, Accurate-mass, High-resolution, and Absolute Quantification), where TMT^0 at high levels conjugated to a desired target analyte is run in conjunction with a traditional TMT experiment and multiplex targeted quantitation can occur simultaneously with discovery multiplex proteomics . One of the pitfalls with this approach is, without the *a priori* knowledge of the target analyte retention time, targeted quantitation is highly prone to being falsely triggered. We and others have previously shown that CCS values for lipids and peptides can be predicted with high fidelity. Within this work we intend to describe how using MOMA and predictive CCS we can perform TOMAHAQ-like experiments with higher throughput, sensitivity and target analyte fidelity on the timsTOF Pro.

1. Erickson et al., 2017, Molecular Cell 65, 361–370 January 19, 2017 ª 2016 Elsevier Inc. http://dx.doi.org/10.1016/j.molcel.2016.12.005

801

Orthogonal evidence for Olfactory Receptors can be used for agonist prediction

Amara Jabeen[1], Shoba Ranganathan[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

Proteins are important biological macromolecules critical for structure, function and regulation of human cells and tissues. The human genome draft is available since 2003 but until now not all the coding genes have known protein products. Human Proteome Project (HPP) was launched in 2010 with the aim of mapping entire human proteome. The HPP community has identified 88.62% of the coding genes as protein products. The remaining 11.38% of the proteins are missing. Since most of the missing proteins are membrane proteins which might have clinical implications, therefore it is important to identify these proteins for utilizing their therapeutical potential. There are several technical challenges that make the missing protein (MP) identification through mass spectrometry (MS), a difficult task.

The largest family among the missing proteins is the olfactory receptors (ORs) which are the superfamily of G-protein coupled receptors (GPCRs). There is no convincing MS evidence is found, even for the single OR. Four of the ORs are given the protein status based on orthogonal evidence. Therefore, we collated the available orthogonal evidence for ORs from published literature. In particular, available ligand evidence can be used for novel agonist prediction.

We have applied different classical ML and deep learning methods to an ectopic OR, with a broad ligand spectrum. OR1G1 (UniProt ID: P47890) is the member of family 1 of the OR superfamily, located on Chromosome 17. OR1G1 is ectopically expressed in gut enterochromaffin cells (normal and tumours) where it is known to be responsible for serotonin release. On the basis of classifier performance, we applied the naive Bayes classifier to a large test dataset, resulting in high probability predictions [1]. Such an approach will assist in collecting experimental evidence for the missing olfactory proteome.

1. Jabeen A, Ranganathan S (2019) Applications of machine learning in GPCR bioactive ligand discovery. Curr Opin Struct Biol, 55:66–76.

802

Multifactorial omics platforms for studying Alzheimer's disease olfactory cells

Laura Rantanen[1, 2], Anthony White[1], Larisa Haupt[2], Katja Kanninen[3]

1. QIMR Berghofer Medical Research Institute, Brisbane City, QDL, Australia

2. Queensland University of Technology, Brisbane City, QDL, Australia

3. University of Eastern Finland, Kuopio, Finland

Neurodegenerative diseases, such as Alzheimer’s disease (AD) cause a significant burden to health and socio-economic sectors worldwide. Previous attempts to generate amyloid-targeting strategies have failed to translate to clinical outcomes. Hence, new approaches targeting the underlying molecular pathways of AD are urgently required. An impaired sense of smell is one of the earliest symptoms and a significant predictor of conversion to AD from mild cognitive impairment (MCI). The olfactory mucosa, the organ of smell in the nose, contains a unique niche of stem cells that are capable of replenishing nasal receptor neurons and supporting cells in the nose throughout adult life. These cells can be collected from patients with relative ease and can easily be expanded in the laboratory to generate human olfactory neurosphere-derived cells (hONS). Patient-derived olfactory cells have demonstrated disease-associated differences in several neuronal diseases, such as Parkinson’s disease, schizophrenia, familial dysautonomia and ataxia telangiectasia. The main objective of this study is to identify significant differences in gene, protein, and

lipid metabolism between control and AD patient-derived olfactory cells. This will also be extended to people with MCI, who may have early changes of AD. Analyses consist of global RNA sequencing, and proteomic screening of cells using subcellular fractionation and liquid chromatography in combination with tandem mass spectrometry (LC-MS/MS). In addition, we will also carry out assays to determine changes in neuroinflammation and mitochondrial function between control, MCI and AD-patient-derived olfactory cells. This study will help elucidate the involvement of the olfactory system in pathogenesis of AD and provide new insights into Alzheimer's disease pathways, biomarkers and therapeutic targets.

803

Dehydration-responsive multi-omics landscape of grasspea: novel findings and unique insights into abiotic stress tolerance

Divya Ms. Rathi[1], Akanksha Ms. Pareek[1], Subhra Prof. Chakraborty[1], Niranjan Prof. Chakraborty[1]

1. National Institute of Plant Genome Research, New Delhi, DELHI, India

Grasspea is a legume known as an excellent source of protein and antioxidants, besides other nutritional traits. It is notable for its hardy nature, water-use efficiency and efficacy as a stress-tolerant pulse as compared to major legume species. Despite superior morpho-physiological attributes, this pulse has largely remained outside the realm of systematic molecular profiling.

Three-week-old grasspea (cv. LP-24, Prateek) seedlings were subjected to dehydration over a period of 144 h. This was followed by evaluation of the temporal effects of dehydration at the proteomic, transcriptome and metabolome levels using 2-DE, RNA-seq and GC-MS. Moreover, the 10-day-old subcultured grasspea suspension cells were subjected to dehydration using 10% PEG and metabolites were analysed using MRM-MS. The RNA-seq data was used to generate an *in-house* database for grasspea (LSDB), which aided in proteomic identifications. Conclusively, we compared the differentially expressed mRNA with of the mRNA products so as to obtain insights into post-transcriptional regulation.

The physiological responses of grasspea were construed by an increase in ROS, disruption in membrane integrity and osmotic imbalance during 72-96 h. The dehydration-induced differential proteomic, transcriptomic and metabolomic analyses, revealed 120 proteins, 5201 genes and 230 metabolites, of various functional classes. The proteogenomic analyses provide crucial insights into the dehydration response, presumably orchestrated by proteins belonging to an array of functional classes including photosynthesis, protein and RNA metabolism, protein folding, antioxidant enzymes and defense.

The dehydration tolerance and/or avoidance mechanism of grasspea appeared to be enforced through remodulation of DRPs belonging to protein biosynthesis, protein folding, photosynthesis and stress response. We report, for the first time, the dehydration-induced proteogenomic-cum-metabolomic landscape of grasspea, whose genome is yet to be sequenced. The cross-species comparison of the proteomes, transcriptomes and metabolomes provides evidence for marked molecular diversity of grasspea.

1. Rathi D, Chakraborty S and Chakraborty N (2015) Proteomics of an orphan legume, grasspea: current status and future strategy. Plant Tissue Cult Biotechnol 25:117-141.
2. Rathi D, Gayen D, Gayali S, Chakraborty S and Chakraborty N (2016) Legume proteomics: Progress, prospects, and challenges. Proteomics 16:310-327.
3. Rathi D, Pareek A, Gayali S, Chakraborty S and Chakraborty N (2018) Variety-specific nutrient acquisition and dehydration-induced proteomic landscape of grasspea (Lathyrus sativus L.). J Proteomics 183:45-57.
4. Rathi D, Gayali S, Pareek A, Chakraborty S and Chakraborty N (2019) Transcriptome profiling illustrates expression signatures of dehydration tolerance in developing grasspea seedlings. Planta, https://doi.org/10.1007/s00425-018-03082-2.
5. Rathi D, Pareek A, Zhang T, Pang Q, Chen S, Chakraborty S, Chakraborty N (2019) Metabolite signatures of grasspea (Lathyrus sativus L.) suspension-cultured cells illustrate dehydration response and acclimation. Planta, https://doi.org/10.1007/s00425-019-03211-5.

804

Comprehensive dog plasma spectral libraries for SWATH-MS Data Acquisition

Halley Ravuri[1], Nana Satake[2], Paul C Mills[1], Pawel Sadowski[3]

1. Veterinary Pharmacology, School of Veterinary Sciences, The university of Queensland, Gatton, Queensland, Australia

2. Analytical Services Unit, School of Agriculture & Food Sciences, The University of Queensland, Brisbane, QLD, Australia

3. Proteomics and Small Molecule Mass Spectrometry Laboratory, Central Analytical Research Facility (CARF), Queensland University of Technology (QUT), Brisbane, Queensland, Australia

Data independent acquisition mass spectrometry, and in particular, SWATH-MS strategy, has revolutionized human biomedical research and is slowly finding its way in clinical veterinary applications. SWATH-MS relies on comprehensive peptide spectral libraries to ease data interpretation and minimise the number of false positives. Plasma is an easily accessed source of protein biomarkers in clinical studies and yet, generating peptide spectral libraries for plasma is not trivial due to a large dynamic range of protein concentrations. In our study in which we employ SWATH-MS based protein quantitation to study a canine model of pain and inflammation, we are therefore employing various fractionation strategies to increase the depth of plasma spectral library.

Fractionation of the dog plasma has been carried out using SDS-PAGE, acetonitrile precipitation and ProteoMiner Protein Enrichment kit. The fractionated samples were digested either in gel or using FASP protocol desalted using StageTip and analysed on TripleTOF 5600 in data dependent acquisition mode. The use of a combination of ProteoMiner and SDS-PAGE resulted in the highest number of proteins followed by a combination of acetonitrile precipitation and SDS-PAGE.

To our knowledge this is the most comprehensive dog spectral library collected to date. Canine models are now increasingly used in translational biomedical research, and hence the generation and availability of a canine plasma proteome would aid advancements in both human and veterinary biomedicine.

805

An alternative strategy to explore missing proteins with low molecular weight

Zhilong Lin[1], Yuanliang Zhang[1], Piliang Hao[2], Yanbin He[1], Siqi Liu[1], yan ren[1]

1. BGI-Shenzhen, Shenzhen, GUANGDONG, China

2. ShanghaiTech University, Shanghai, China

Digging more missing proteins (MPs) is an important mission of C-HPP. With the MP number is being attenuated year by year, we have realized that the difficulty to explore the remaining MPs is a challenge in technique. Herein, we propose a comprehensive strategy to effectively enrich, separate and identify the proteins with low molecular weight, aiming at discovery of MPs. Basically, the protein extract from human placenta were passed through a C18 SPE column, and the bound proteins eluted were further separated by SDS-PAGE gel or 50kD cutoff filter. The separated proteins were undergone trypsin digestion and the MS/MS signals were searched against datasets with two different digestion modes. The strategy was adopted, resulting in 4 MPs (≥2 non-nested unique peptides with ≥9 amino acids) identification containing 8 unique peptides, one of which was only composed of 80 a.a.. Importantly, seven out of eight unique peptides derived from the MPs were verified by parallel reaction monitoring, demonstrating that the placenta tissues indeed possessed the MPs. The study does not only establish a feasible strategy for analysis of the proteins with low molecular weight, but also fill a small gap to the MP list.

806

An optimization towards post-search process to improve the identification of HLA-associated peptides

Zhe Ren[1], Jie Chen[1], Da Qi[1], yan ren[1], Siqi Liu[1]

1. BGI-Shenzhen, Shenzhen, GUANGDONG, China

In cancer immunotherapy, the precise and efficient identification to HLA-associated peptides (HAP) is a critical step, while liquid chromatography-tandem mass spectrometry (LC-MS/MS) is a premising means to discover these HAPs that are partially not derived from the traditional annotation of the human genome. The HAPs are generally considered as the endogenous peptides cleaved by uncertain enzyme(s). In contrast to a conventional search upon tryptic peptides, the non-specific cleavage peptides are expected with their unique features and are identified through a search specially optimized. By inspecting several publicly available datasets and home-made ones, we found the search results for HAPs quite unstable with lower identification rate and high false positive rate by using several common search engines, such as Mascot, Comet, X! Tandem and MSGF+ and some post-processor, like PeptideProphet and Percolator. Herein, we have focused on the characteristics of non-specific cleavage peptides and optimized the post-search process to improve the estimation of posterior error rate. In the primary test using Mascot, with only one additional feature and percolator training, the identified PSMs in two datasets increased from 1,401 to 1,986 and from 9,133 to 9,351, respectively. With more features added, the results were further improved. In future, we are looking forward to finding a proper combination of features and training strategy to further optimize the post-process with a robust increase of identification rate and consistency using different search engines. And the optimization is expected not only limited in the identification of HAPs but also beneficial to the identification of all the peptidome with non-specific cleavage search involved.

807

Development of Integrative Protein and Lipid Organelle Profiling (iPLOP) method for high throughput organelle analysis

Harley R Robinson[2, 1], Jeffrey Molendijk[1], Ahmed Mohamed[1], Leonard Foster[3], Michelle M Hill[4, 1]

1. Precision and Systems Biomedicine, QIMR Berghofer Medical Research Institute , Brisbane, QLD, Australia

2. Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia

3. Michael Smith Laboratories, University of British Colombia, Vancouver, BC, Canada

4. UQ Diamantina Institute , Brisbane, QLD, Australia

Mammalian cells are compartmentalised into membrane bound organelles, which serve distinct functions. Dysregulation of cellular lipid distribution leads to loss of organelle homeostasis and underlie numerous chronic diseases. While methods of spatial proteomics are relatively well established, such as Protein Correlation Profiling (PCP) and hyperLOPIT methods, high throughput organellar lipidomics and integration with organellar proteomics are yet to be established. This study aims to develop a high-throughput method for proteo-lipidomic profiling of organelles, termed integrative Protein and Lipid Organelle Profiling (iPLOP). iPLOP uses a continuous sucrose gradient to roughly separate various membrane bound organelles into different profiles. Lipids and proteins are extracted from the same sucrose gradient fractions and subjected to shotgun proteomics and targeted lipidomics using mass spectrometry. As most lipid species localize to several organelles at varying abundances, lipidomics will aim to quantify the lipid species at each organelle. Several computational analysis approaches will be explored to assign proteome and lipidome to appropriate organelles. Downstream pathway analyses and computational functional prediction will be integrated to determine functional effects of altering organelle composition. Here, we present the development of the iPLOP method and application in studying membrane remodelling abnormalities facilitated by cancer mediators, caveolin-1 and cavin-1, in advanced prostate cancer. We anticipate that applications of iPLOP method will contribute to an improved understanding of the spatial relationship between cellular lipids and proteins in health and disease.

808

Post-translational modifications on recombinant human factor IX from fed-batch and perfusion chinese hamster ovary cell culture

Dinora Roche Recinos[1], Lucía F. Zacchi[1, 2], Casandra Pegg[2], Toan K. Phung[2], Campbell Aitken[3], Ellen Otte[3], Tony Hunt[3], Mark Napoli[3], Yih Yean Lee[3], Ben Schulz[1, 2], Christopher B. Howard[1]

1. Centre for Biopharmaceutical Innovation (CBI), Australian Institute of Bioengineering and Nanotechnology (AIBN) at The University of Queensland (UQ), Brisbane, QLD, Australia

2. School of Chemistry and Molecular Biology, University of Queensland, Brisbane, QLD, Australia

3. CSL Limited, Melbourne, Victoria, Australia

Human coagulation factor IX (FIX) is a protein that relies on an extensive spectrum of posttranslational modifications (PTMs) for its correct and efficient function in the coagulation pathway [1, 2]. These PTMs include seven disulfide bridges, two *N*-glycans, six *O*-linked glycans, one sulfation site, one phosphorylation site, 12 g-carboxylation (GLA) sites, and one b-hydroxylation site [3-9]. Here, we investigated the differences in the PTMs of human recombinant factor IX (rFIX) produced in CHO fed-batch and perfusion cultures, compared with native plasma-derived factor IX (PD-FIX). The cell line used was a CHO-K1SV expressing rFIX. Two fed-batch bioreactors were conducted using commercial CD-CHO media and EfficientFeed A and B respectively. Perfusion cultures were conducted in the same base medium using an Applisens Biosep acoustic perfusion unit at a dilution rate of one reactor volume per day. The bioreactors were sampled daily for off-line measurements to track cell growth, metabolism and productivity. These samples were also used for Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS) [10] analysis of the culture supernatant proteome during the time-course of fermentation. We also performed detailed LC-MS/MS characterization of purified rFIX from each culture.The fed-batch cultures responded differently to each of the feeds despite achieving similar peak cell densities of ~15x10^6 cells/mL. Pseudo steady-states were established in the perfusion cultures at 15x10^6 cells/mL via bleeding of the cultures under the control of an online turbidity probe. Almost all the PTMs present on PD-FIX were also observed in rFIX of fed-batch and perfusion cultures, although they showed partial occupancy and higher heterogeneity in rFIX. The extent of gamma-carboxylation in the rFIX GLA domain, and of diverse other PTMs across the protein, was affected by fermentation conditions, emphasizing the utility of LC-MS/MS techniques in monitoring the quality of recombinant biopharmaceuticals, especially those heavily modified by diverse PTMs.

809

Welfare Biomarkers For Farmed Gilthead Sea Bream (*Sparus Aurata*): Integrating Multi-Omics Data

Cláudia R Magalhães[1], Marco Cerqueira[1], Denise Schrama[1], Paula F Resende[1], Annette Kuehn[2], Dominique Revets[2], Surintorn Boonanuntanasarn[3], Chatsirin Nakharuthai[3], Pedro M Rodrigues[1]

1. CCMAR, Universidade do Algarve, Campus de Gambelas, Ed. 7, Faro, Portugal

2. Department of Infection and Immunity, Luxembourg Institute of Health, 29, rue Henri Koch, L-4354 , Esch-sur-Alzette, Luxembourg

3. School of Animal Production Technology, Institute of Agricultural Technology, Suranaree University of Technology, 111 University Avenue, Muang, Nakhon Ratchasima, Thailand

Managing the welfare of fish in captivity is crucial to ensure a sustainable aquaculture production. Physiological stress responses, such as the plasma levels of cortisol, glucose and lactate, are the most common indicators of assessing farmed fish welfare, but their reliability has been questioned due to high biological variability and fish adaptation processes. An integrated multi-omics approach can be a promisor strategy to discover more robust fish welfare biomarkers, since it can offer the possibility of understanding the complete flow of information in the fish biological system. The aim of this work is to use proteomics to identify a restricted protein map as putative fish welfare biomarkers and integrate these results with transcriptomics and metabolomics data to achieve a global picture of the fish response to stress. *Sparus aurata* was reared under different stressful conditions: overcrowding, repetitive net handling (air exposure), and hypoxia, using fish reared under optimal conditions as control. Fish were sampled after 45 days of trial and proteins extracts were prepared from liver samples. Proteins were separated by 2D-DIGE and identified by MALDI-TOF/TOF MS. Putative welfare biomarkers were then chosen based on their stress-related function, fold-change and score, and used for primer design. Total RNA was extracted from liver samples using Trizol reagent with DNase treatment and used for cDNA synthesis. The mRNA levels of the target genes were assessed by real-time PCR. Comparative proteomics showed, in the liver, a total of 147 proteins statistically different in their abundance among conditions, from which 24 were indicated as putative welfare biomarkers and chosen for their transcription level analysis. Quantitative gene expression analysis reveals that the levels of transcripts of 7 of the target genes were modulated. This joint analysis provides a starter point for the development of more reliable fish welfare assessment measures to improve aquaculture sustainability.

810

Proteomic profiling identifies novel pathways modulated by repurposed drug candidate Nitroxoline in AsPC-1 pancreatic cancer cells

Maurizio Ronci[1], Serena Veschi[2], Laura De Lellis[2], Luca Scotti[1], Paola Lanuti[3], Antonio Aceto[1], Alessandro Cama[2]

1. Department of Medical, Oral and Biotechnological Sciences, University G. d'Annunzio of Chieti-Pescara, Chieti, Italy

2. Department of Pharmacy, University G. d'Annunzio of Chieti-Pescara, Chieti, Italy

3. University G. D'Annunzio of Chieti Pescara, Chieti, ABRUZZO, Italy

Background Nitroxoline is an antibiotic widely used in several countries from 1960s known to act by chelating divalent metal ions such as Mg2+ and Mn2+. In a recent study, we analysed the effects of nitroxoline as a candidate for repurposing in pancreatic cancer cells (AsPC-1), showing that the drug has antitumor effects comparable to erlotinib, a drug approved for PC treatment.

Methods. In the present study, by exploiting a label-free shotgun proteomics approach, we analysed the effect of nitroxoline on the protein repertoire of AsPC-1 cells after 24 and 48 hours of treatment.

Results The label-free shotgun proteomics analysis resulted in the detection of 81 proteins consistently deregulated at both timepoints. Those protein were subjected to a few bioinformatic analyses including STRING and IPA. STRING analysis showed that many of the proteins are connected in a single network, with a highly significant protein-protein interaction enrichment (p-value < 10-16). The highest enriched pathways, as showed by IPA analysis, are related to some of the pathways known to be modulated by nitroxoline such as mTOR and sirtuins. Among the additional highest ranked pathways several were related to metabolism including Krebs cycle and aminoacid, lipid and carbohydrate metabolism. The level of expression of ATP1B3, β-catenin and TFR1 were confirmed by western blot. Our results suggest that Na/K-ATPase downregulation could to be linked to several antitumor actions of nitroxoline, including inhibition of the PI3K/AKT/GSK3β/β-catenin pathway, induction of oxidative stress and triggering of DNA damage response.

Conclusions With the present study we identified target pathways previously unknown to be affected by Nitroxoline.Our results indicate that Nitroxoline treatment impairs cell migration and invasion, increases intracellular ROS and activates oxidative DNA damage response, reduces cotranslational protein targeting to membrane and increases lactate production.

811

Cellular Imprinting Proteomics applied to ocular disorders elicited by Congenital Zika virus Syndrome

Livia Rosa Fernandes[1], Raquel H Barbosa[2], Maria Luiza B dos Santos[2], Claudia B Angeli[1], Thiago P Silva[3], Rossana CN Melo[3], Gilberto SO Oliveira[1], Martin R Larsen[4], Claudete A Cardoso[2], Giuseppe Palmisano[1]

1. GlycoProteomics Laboratory, Institute of Biomedical Sciences, University of Sao Paulo, Sao Paulo, Brazil

2. UFF, Niteroi, Rio de Janeiro, Brazil

3. FUJF, Juiz de Fora, Brazil

4. University of Southern Denmark, Odense, Denmark

Congenital Zika Syndrome (CZS) was identified due to the increased incidence of congenital defects associated with Zika virus (ZIKV) infection, which include morphological, behavioral, neurological and ocular impairments. The eye has been described as a specialized CNS compartment being able to display symptoms of neurodegenerative diseases. Ocular lesions may result from defects that occurred during embryogenesis and are apparent in newborns exposed to ZIKV. However, late pregnancy infection or the absence of microcephaly cannot preclude the occurrence of ocular lesions and other CNS manifestations. Considering the need for surveillance of babies with possible ZIKV congenital exposure, we developed a method termed Cellular Imprinting Proteomics, CImP, for the identification and quantification of the ocular surface proteome of infants exposed to ZIKV during gestation.

The cohort was divided into CTRL (no infectious diseases), ZIKV (infants exposed to ZIKV, without microcephaly) and $Zikv^{CZS}$ (infants exposed to ZIKV, with microcephaly). The CImP is based on an improved impression cytology method and an optimized pipeline to extract and analyze the ocular proteome using mass spectrometry-based technology. A total of 2209 proteins were identified with modulation of neutrophil degranulation, cell death, ocular and neurodevelopment pathways in $Zikv^{CZS}$ compared to CTRL. Moreover, the molecular pattern of ocular surface cells retrieved from infants infected during the gestation but with no CZS was different from matched controls.

Molecular alterations in the ocular surface associated to ZIKV infection with and without CZS complications are reported for the first time. We predict that this method will be introduced successfully in the study of several neurological diseases with the aim to identify novel diagnostic and therapeutic biomarkers.

812

Fast and accurate bacterial identification in clinical samples by MS-DIA crude signal mining using Machine Learning: application to Urinary Tract Infections

Florence Roux-Dalvai[1], Mickael Leclercq[1], Marion Narbeburu[1], Clarisse Gotti[1], David Bouyssie[2], Boissinot Maurice[3], Odile Burlet-Schiltz[2], Michel Bergeron[3], Arnaud Droit[1]

1. Proteomics Platform and Computational Biology laboratory, CHU de Québec - Université Laval, Quebec, QUEBEC, Canada

2. Institut de Pharmacologie et Biologie Structurale - Université de Toulouse, Toulouse, France

3. Centre de Recherche en Infectiologie, CHU de Québec - Université Laval Research Center, Quebec, QUEBEC, Canada

Bacterial identification is essential to many applications in biology, health, and environment. Low cost and fast MALDI-TOF technology has become an approach of choice for this purpose but has several drawbacks: a long step of bacterial culture prior to analysis (>24h), low specificity and is not quantitative.

We developed a new strategy using Machine Learning algorithms to explore into LC-MSMS-DIA (Data Independent Acquisition) data in order to detect specific bacterial signals without the need of a bacterial culture nor peptide/protein identification.

As a proof of concept, we used the 15 bacteria most commonly found in urinary tract infections (UTI). To do so, 200 bacterial inoculated urine specimens were analyzed on an Orbitrap Fusion instrument in DIA mode. Raw data were converted into LC-MS maps corresponding to each precursor window, then a systematic binning in both m/z and time dimensions was performed and compared to a peptide feature detection strategy. Data tables resulting from both methods were tested with various machine learning classifiers associated to dimensionality reduction techniques to determine the best conditions for species discrimination. Moreover, mass recalibration and retention time alignment tools were used to improve the prediction accuracy and make it transferable to other laboratories. Standard LC-MS gradient (90min) and short gradient (15min), more suitable for routine analyses, were also compared. With this strategy, we were able to obtain 90 to 95% accuracy in bacterial prediction for bacterial concentration < $1x10^5$ CFU/mL.

Our new approach, using cutting-edge technologies in proteomics and computational biology, is able to identify bacteria responsible for 85% of UTI in few hours without the need of bacterial culture or peptide/protein identification. This work paves the way to development of new generation diagnostic methods and could be extended in the future to other biological specimens and to bacteria having specific resistances.

813

Quantitative proteomics for studying the effect of lysine succinylation on aflatoxin production in *Aspergillus flavus*

Li Tao[1], Li ying[1], Hu Rui[1], Yang yunhuang[1]

1. Wuhan Institute of Physics and Mathematics, the Chinese Academy of Sciences, Wuhan City, HUBEI PROVINCE, China

Publish consent withheld

814

Serum profiling using NAPPA reveals an autoantibody signature for the early diagnosis of knee osteoarthritis: Data from the Osteoarthritis Initiative

María Camacho-Encina[1], Vanesa Balboa[1], Ignacio Rego-Pérez[1], Florencia Picchi[1], Jennifer VanDuin[2], Ji Qiu[2], Manuel Fuentes[3], Natividad Oreiro[1], Joshua LaBaer[2], Cristina Ruiz-Romero[1, 4], Francisco J Blanco[1, 5]

1. Instituto de Investigación Biomédica de A Coruña (INIBIC-CHUAC), A Coruña, Spain

2. Virginia G. Piper Center for Personalized Diagnostics, Biodesing Institute-Arizona State University, Tempe, Arizona, USA

3. Department of Medicine and General Cytometry Service-Nucleus. Proteomics Unit. CIBER-ONC CB16/12/00400. Cancer Research Center (IBMCC/CSIS/USAL/IBSAL), Salamanca, Spain

4. CIBER-BBN Instituto de Salud Carlos III, INIBIC-CHUAC, A Coruña, Spain

5. RIER-RED de Inflamación y Enfermedades Reumáticas, INIBIC-CHUAC, A Coruña, Spain

Osteoarthritis (OA) is a disease commonly diagnosed at advanced stages, when the damage to the joint is irreversible. In OA pathogenesis, changes involving the joint could lead to the production of autoantibodies (AAbs) through a humoral response, even at asymptomatic stages. Therefore, the discovery of an AAb signature might facilitate the identification of those patients who will develop the disorder.

In this study we analyzed sera obtained at baseline from subjects belonging to the Osteoarthritis Initiative (OAI) cohort that were subsequently followed for 96-months. Ten pools of sera per study group (incident and not-incident) were blindly analyzed by Nucleic Acid-Programmable Protein Arrays (NAPPA) to screen immunoreactivity against 2,125 human proteins. Quantitative data were normalized following the Biodesign Institute criteria and a cut-off level was calculated by the median intensity absolute deviation rule from all the spots through all the pools to determine AAb reactivities. A differential spot analysis was carried out with the antigens over the cut-off by Wilcoxson Rank-Sum test. Among the final panel of candidate AAbs, the association of the reactivity levels against methionine adenosyltransferase 2 subunit beta (MAT2ß) with OA onset was verified in 327 sera from the OAI cohort by NAPPA-ELISA assays.

Among the 2125 proteins in the array, a panel of 6 AAbs showed significant ($p<0.05$) different baseline reactivity levels between the incident and not-incident subjects. The NAPPA-ELISA analysis verified the presence of significant higher baseline reactivity levels of MAT2ß–AAb in those patients who did develop OA at some point of the 96-months follow-up period (0.58 ± 0.22 vs 0.49 ± 0.23 a.u., p= 3.140E-04). The odds ratio of incident OA was 5.99 (2.16–16.63) times higher per one-unit increase of MAT2ß-AAb.

In conclusion, we have identified an autoantibody signature in serum that could be useful for the diagnosis of osteoarthritis at early, asymptomatic stages.

815

Development of sandwich antibody microarrays to validate a panel of potential protein biomarkers of osteoarthritis

Rocío Paz-González[1], María Camacho-Encina[1], Valentina Calamia[1], Lucía Lourido[1], Patricia Fernández-Puente[1], Beatriz Rocha[1], Lucía González-Rodríguez[1], Cristina Ruiz-Romero[1, 2], Francisco J Blanco[1]

1. Instituto de Investigación Biomédica de A Coruña (INIBIC-CHUAC), A Coruña, Spain

2. CIBER-BBN Instituto de Salud Carlos III, INIBIC-CHUAC, A Coruña, Spain

In complex diseases such as osteoarthritis (OA), the measurement of multiple biomarkers is expected to provide valuable information to improve diagnosis, prognosis and therapy monitoring. Bead-based xMAP technology offers a flexible and open platform to simultaneously quantify proteins. The aim of this study has been to develop sandwich antibody microarrays to validate a panel of 6 potential OA biomarkers in clinical samples.

Antibody pairs and recombinant proteins tested to work in an ELISA were purchased for each of the 6 analytes of interest. First, 1.6 µg of each monoclonal antibody were coupled to $5x10^5$ of an activated fluorescent bead region and the efficiency of the coupling was confirmed by Phycoerythrin-labelled anti-species antibodies. Assay conditions, including the standard curve range and sample dilution, were optimized for each protein individually. Cross-reactivity of antibodies with the non-target analyte was assessed between those proteins requiring the same sample dilution to generate multiplex sandwich immunoassays. Then, analytical characteristics in terms of accuracy, precision and limit of detection (LLOD) and quantification (LLOQ) were evaluated and finally their utility to quantify the potential biomarkers was assessed.

Individual capture immunoassays were successfully converted for all biomarkers to a protein microarray platform. According to the serum sample dilution, the six proteins were finally grouped in three different immunoassays: a duplex, a triplex and a singleplex sandwich immunoassays. No cross-reactivity was observed for the antibody pairs against the non-targeted proteins. Median Fluorescence Intensity of the negative controls in 8 replicates was used to evaluate the LLOD and LLOQ of each microarray. Precision and accuracy were calculated using 3 different known concentrations of the standards. In all cases precision was below 10% and accuracy ranged between 70–130%.

In conclusion, three different sandwich immunoassays have been successfully developed in a suspension beads array format to absolutely quantify six potential OA biomarkers.

816

Ion chromatography coupled with mass spectrometry (IC-MS): an innovative tool for the characterization of the sulfated O-linked glycans

Zoltan Szabo[1], Neil G Rumachik[2], Jim Thayer[2], Gina Tan[3], Rosa Viner[3], Andreas Huhmer[3], Jeff Rohrer[2], Yan Liu[2], Dietmar Reusch[4], Christopher A Pohl[2]

1. Thermo Fisher Scientific, South San Francisco, CA

2. Thermo Fisher Scientific, Sunnyvale, CA, United States

3. Thermo Fisher Scientific, San Jose, CA

4. Roche Diagnostics GmbH, Penzberg, Germany

High Performance Anion Exchange Chromatography (HPAE) has long been used to separate ions and polar molecules. Here, we demonstrate that HPAE can be coupled to an Orbitrap™ mass spectrometer (IC-MS) to facilitate the separation and sequencing of sulfated O-glycan alditols. Porcine gastric mucin (type III) harbors high levels of sulfated but fewer sialylated glycans, and after release and rapid purification, these O-linked glycans were directly injected and analyzed. However, to analyze bovine submaxillary mucin, which harbors fewer sulfated but more sialylated O-glycans, a new workflow based on weak anion exchange (WAX) fractionation was necessary to minimize ion-suppression by the sialylated glycans. Using a post column splitter, 60% of the flow was desalted with an electrochemically regenerated desalter prior to introduction into the mass spectrometer. Due to the ability of HPAE to resolve isomers, 27 sulfated O-linked glycans were identified from Porcine gastric mucin (type III), and 9 from bovine submaxillary mucin. Using the data-dependent mode of the Thermo Scientific™ Q Exactive™ hybrid quadrupole-Orbitrap mass spectrometer, highly informative MS^2 spectra enabled confident identification of the separated species.

817

Structural Basis For DNA Recognition And Transcription Activation By The Response Regulator OmpR

Sushant Sadotra[1, 2], Yuan-Chao Lou[1], Chun-Hua Hsu[3], Chinpan Chen[1]

1. Institute of Biomedical Sciences, Academia Sinica, Taipei City, Taiwan

2. Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan

3. Department of Agricultural Chemistry, National Taiwan University, Taipei, Taiwan

Publish consent withheld

818

ProTargetMiner: a proteome signature library of anticancer molecules for functional discovery

Amir Ata Saei[1], Christian Beusch[1], Alexey Chernobrovkin[1], Pierre Sabatier[1], Bo Zhang[1], Ülkü Güler Tokat[1], Eleni Stergiou[1], Massimiliano Gaetani[1], Ákos Végvári[1], Roman A. Zubarev[1]

1. Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

ProTargetMiner is the first publicly available expandable proteome signature library of anticancer molecules in cancer cell lines. Based on 287 A549 adenocarcinoma proteomes affected by 56 compounds, the main dataset contains data on 7,328 proteins and 1,307,859 refined protein-drug pairs. The proteomic signatures cluster into groups by the compound targets and action mechanisms. The targets and mechanistic proteins can be deconvoluted by partial least square modeling provided online through an R Shiny package and a website (http://protargetminer.genexplain.com). For 9 molecules representing most diverse action

mechanisms and common cancer cell lines MCF-7, RKO and A549, deep proteome datasets were obtained. Combining the data from the three cell lines identified common drug targets and mechanisms, while also highlighting important cell-specific differences. ProTargetMiner can also provide information on resistance factors, affected protein complexes and drug metabolizing enzymes. The database can be easily extended and merged with new compound signatures. We expect ProTargetMiner to serve as a valuable chemical proteomics resource for the cancer research community, and become a popular tool in drug discovery.

819

Chondroitin sulfate and heparan sulfate determine axonal regeneration and its inhibition through PTPRσ-Cortactin-Autophagy axis

Kazuma Sakamoto[1], Kenji Kadomatsu[1]

1. Nagoya University Graduate School of Medicine, Nagoya, AICHI, Japan

The human neural circuit is about 50 thousand kilometers in length which is connected by axons. During development of the neuronal network, neurons connect with one another by extending their axons. The longest axon in human can stretch up to approximately 1 meter.

The delicate neural circuit can be easily disrupted by physical forces. For example, spinal cord injury can cause a series of damage to the neuronal axons. Distal part of neurons undergo Wallerian degeneration and removed from tissue, while proximal part of neurons are still alive and try to extend their axons again to reconnect neural circuits. However, their efforts end up in vein. As a result, neural network permanently remains disconnected and patients suffer from paralysis for life. One of the reasons is the accumulation of chondroitin sulfate (CS) at the injury site. In contrast, heparan sulfate (HS) promotes axon growth. Interestingly, HS and CS are similar in molecular structure and bind to the same receptor, PTPRσ (a receptor type tyrosine protein phosphatase). However, it has been elusive why or how these similar glycans cause opposite effects on the axon regeneration through the same receptor.

Based on chemically synthesized a series of HS and CS, which are different both in length and sulfation patterns. We found that HS could polymerize PTPRσ and promote axonal extension, while CS monomerized it and disrupted the extension. Upon binding to PTPRσ, CS activated this receptor's enzymatic activity, and consequently dephosphorylated cortactin. As cortactin is critical for autophagy, CS-induced cortactin de-phosphorylation stopped autophagy flux and transformed axon tips to ball-like structures, so-called dystrophic endballs, the hallmark of injured axon. Indeed, an artificial disruption of autophagy induced dystrophic endballs.

Taken together, our results clearly revealed novel CS-PTPRσ-Cortactin-autophagy pathway which was involved in axonal regeneration inhibition.

820

Detection of ultra-low abundant epitopes by targeted mass spectrometry

Mojjib Salek[1, 2], Jonas Föster[1, 2, 3], Rebecca Köhler[1, 2], Angelika B. Riemer[1, 2]

1. Immunotherapy & Immunoprevention, German Cancer Research Center (DKFZ), Heidelberg, Germany

2. Molecular Vaccine Design, German Center for Infection Research (DZIF), partner site Heidelberg, Heidelberg, Germany

3. Faculty of Biosciences, Heidelberg University, Heidelberg, Germany

Rational design of therapeutic cancer vaccines or TCR-based adoptive transfer approaches requires knowledge of the epitope repertoire presented on the surface of tumor cells. However, direct detection of viral or mutation-derived epitopes remains an analytical challenge, as currently several hundred million cells are required for their detection, which so far prevents systematic application in clinical settings.

Previously we reported a multiple reaction monitoring with multi-stage fragmentation (MRM3) targeted strategy for the detection of HLA-presented human papillomavirus (HPV)16 epitopes, undetectable by untargeted approaches. Here we present the transfer and adoption of the strategy to a high-resolution mass spectrometer that ensures confident detection based on high mass accuracy in the parallel reaction monitoring (PRM) mode.

To achieve ultra-high sensitivity for the detection of low abundant target HLA-presented peptides, extremely long MS2 injection times are required. However, due to the presence of interfering signals, including background proteins, detergent and other contaminants present in the immune-isolates, careful selection of acquisition parameters is mandatory to prevent excessive overfilling of the orbitrap, which may compromise the sensitivity of detection. Our approach includes the generation of a high quality spectral library, and dominant charge state determination of the predicted target epitopes by prior analysis of their stable isotope labeled (SIL) surrogates. Although the default collision energy is normalized to the mass-to-charge ratio (NCE) of the precursor ion, we found that additional tuning of the NCE greatly improves detection. Moreover, spiking an unrelated peptide-rich background proteome into the sample allows detection of very low abundance peptides that otherwise would remain undetected.

Our preliminary results show better consistency of detection of HPV-derived epitopes than with the previously used MRM3 approach. However, further optimization is mandatory to achieve the analysis of tumor biopsies, for instance as part of larger personalized clinical diagnostic and therapeutic pipelines.

821

High-Throughput Discovery and Validation of Post-Translational Modifications on HLA-Presented Peptides from Patient Tumor Samples

Robert Salzler[1], Darshit Shah[1], Lisa Chen[1], David DiLillo[1], Olav Olsen[1], William Olson[1]

1. Regeneron, Tarrytown, NEW YORK, United States

Introduction: Tumor specific antigens can be processed and presented by HLA class I alleles just as normal 'self" proteins, adding to the immunopeptidome. High Resolution Accurate Mass Mass Spectrometry (HRAM-MS) is the tool of choice to interrogate the immunopeptidome in an accurate, systematic and unbiased manner. Post-translational modifications (PTMs) are known to play a critical role in the etiology of cancer and other disease states. Cancer-associated post-translational modifications provides a potential source of tumor-specific epitopes (1,2,3,4).

Goal: The objective of this study was to use HRAM-MS to detect aberrant PTMs (methylation, phosphorylation, deamidation, citrullination, acetylation etc.) from HLA Class I presented tumor-specific peptides.

Method: We used an optimized workflow consisting of anti-HLA affinity chromatography, C18 solid phase capture and selective elution follow by LC-HRAM-MS for immunopeptidomics experiments. Data analyisis was performed using PEAKS Studio X (BSI) database searching workflow incorporating PEAKS PTM which is specifically designed to discover hidden modifications by integrating the powerful *de novo* sequencing algorithm and database searching.

Results: Typically, ~7000 unique HLA-associated peptides were identified in a single 120 minutes LC-MS analysis using 1 × 10^8 cells or 0.5 g of fresh frozen tumor tissue. In this study, we identified a robust pattern of preferential binding of peptides with PTMs to specific HLA molecules such as methylated peptides to HLA-B*07 (Class I) from both cell lines and a variety of human tumors. There were also several HLA peptides with PTMs detected consistently in our study which were predicted binders to a wide array of HLA Class I alleles.

Conclusion: Our current workflow routinely detects HLA-peptides with specific aberrant PTMs potentially providing a unique source of disease related HLA class I presented peptides that can trigger specific immune response and hence provides additional opportunity for targeted immunotherapy.

822

Proteomics and metabolomics analysis of PBMCs from first-episode psychosis patients

Cátia Santa[1, 2], Margarida Coelho[1, 3], Manuel Coroa[4, 5], Sofia Morais[4, 5], Miguel Bajouco[4, 5], Vera M. Mendes[1], Sandra I. Anjo[1], Inês Baldeiras[1, 4], Nuno Madeira[4, 5], António Macedo[4, 5], Bruno Manadas[1]

1. Center for Neuroscience and Cell Biology of the University of Coimbra, Coimbra, Baixo Mondego, Portugal

2. Institute for Interdisciplinary Research, University of Coimbra, Coimbra, Baixo Mondego, Portugal

3. Chemistry Department, Faculty of Sciences and Technology, University of Coimbra, Coimbra, Baixo Mondego, Portugal

4. Faculty of Medicine, University of Coimbra, Coimbra, Baixo Mondego, Portugal

5. Psychiatry Department, CHUC, Coimbra, Baixo Mondego, Portugal

Schizophrenia is a chronic and debilitating psychiatric disorder for which there is still no biomarker. The diagnosis of the disease is based primarily on clinical interview with no biomolecular support which can be used to increase diagnosis confidence or to guide prognosis. Moreover, medication-resistant patients need to be subjected to long and unhelpful therapy trials before initiating clozapine, which is far from the goal of personalized and preventive medicine.

In the present project, we aim at analyzing the proteic and metabolic content of the peripheral blood mononuclear cells (PBMCs) of minimally medicated first-episode psychosis (FEP) patients, comparing with healthy controls, in order to depict which pathways may play a role at the onset of the disease.

To achieve this goal a state-of-the-art quantitative approach, SWATH-MS, was used for both proteomics and metabolomics analysis. In total, thousands of metabolites and proteins were quantified. After the application of quality and statistical filters, several proteins and metabolites were defined as differentially expressed between FEP and control samples. Multivariate analysis using the differentially expressed proteins and metabolites demonstrated that these were enough to distinguish between the study groups, and after functional analysis of these same proteins some pathways were highlighted as potentially interesting for further studies.

In the future, the results here discussed will be further pursued and the analysis of the proteomic and metabolomic signatures throughout the course of disease and treatment, as well the comparison with other major psychiatric disorders, will be addressed.

This work was financed by ERDF through COMPETE 2020 and Portuguese funds via FCT, under projects: POCI-01-0145-FEDER-007440, POCI-01-0145-FEDER-016428, POCI-01-0145-FEDER-30943 and POCI-01-0145-FEDER-016795; and by The National Mass Spectrometry Network (LISBOA-01-0145-FEDER-402-022125). CS was supported by PhD fellowship SFRH/BD/88419/2012, co-financed by the European Social Funds through the POCH and FCT.

823

SWATH-MS based proteomics enabled studying the semen quality in Brahman bulls

Nana Satake[1], Gry Boe-Hansen[2], Rajesh Gupta[3], Pawel Sadowski[3]

1. School of Agriculture and Food Science, The University of Queensland, St Lucia, QLD, Australia

2. School of Veterinary Sciences, University of Queensland, Gatton, QLD, Australia

3. Central Analytical Research Facility (CARF), Queensland University of Technology, Brisbane, QLD, Australia

Over the last two decades, mass spectrometry (MS)-based proteomics techniques have emerged as powerful tools for enabling complex biological systems to be pieced together. However, the adoption of proteomic approaches into studies of animal physiology lags behind the human biomedical research. Traditionally, veterinary proteomics investigations are often conducted using laborious 2D-gel techniques, with difficulty in resolving proteins within "gel spots". Recently, SWATH-MS approach has emerged as reliable and high throughput alternative. We have applied SWATH-MS based quantitation workflow on a study assessing the changes in bull semen characteristics after an acute temperature increase by scrotal insulation (SI), electroejaculated *Bos indicus* Brahman bulls. Brahmans bulls (n=6) were split into two groups of three bulls with and without SI, and semen was collected by electroejaculation at three days intervals from before (-10 d), until 74 d after initiation of SI. A subset of samples identified by peripheral semen characteristics, were selected for SWATH-MS assessment, at 21 and 24d after SI, where there were significant differences in the "semen quality" between the two experimental groups. DDA acquisition identified a total of 418 proteins from a pool of seminal plasma samples, in order to obtain the base spectral library. Subsequent SWATH-MS analysis reproducibly quantified 158 proteins and a significant difference in magnitude of change in 29 proteins (adjusted $P<0.05$) was found between the two treatment groups. SWATH-MS result was compared to previously acquired 2D-gel data on the same set of samples revealing vast agreement between these two techniques, although SWATH-MS showed a greater resolution of changes observed in the proteome. The change in seminal plasma proteome profile, stimulated by an acute heat insult, has direct ramifications on maturing spermatocytes during spermatocytogenesis, subsequently semen quality. The verification of these biomarkers assists management of bull fertility from testicular insult to recovery during acute environmental heat stress.

824

MS-based single-cell proteomic screening for oocyte quality during *in vitro* production a reality?

Gry Boe Hansen[1], Rajesh Gupta[2], Pawel Sadowski[2], Nana Satake[3]

1. School of Veterinary Sciences, University of Queensland, Gatton, QLD, Australia

2. Central Analytical Research Facility (CARF), Queensland University of Technology, Brisbane, QLD, Australia

3. School of Agriculture and Food Science, The University of Queensland, St Lucia, QLD, Australia

Remarkable developments in resolution and sensitivity of MS-based proteomics techniques are continuously revolutionising the way we study cellular communication. Single cell proteome analysis is fast becoming a reality with a variety of sample input methods, such as imaging or flow cytometry interfaces, in addition to traditional chromatography giving a new dimension to study physiological mechanisms at the individual cellular level. The mammalian oocyte is one of the largest single cell with the ability to ultimately produce a multicellular being. Thus, the implications of application and integration of advanced MS-based single cell proteomics is vast. This is particularly true as a powerful tool in *in vitro* production (IVP) systems where individual oocyte qualities may be verified or embryo health screened prior to implantation without destruction of the main oocyte/embryo. One of the most important cellular collaborations for the successful maturation of an oocyte is the cumulous-oocyte complex (COC). This complex is absolute paramount in oocyte maturational health, with the presumptive changing combined proteome. Within the past few years, our group was given the opportunity to trial a technique developed by Bruker, Parallel Accumulation Serial Fragmentation (PASEF) coupled with Trapped Ion Mobility Spectrometry (timsTOF) on samples consisting of a single digested *in vitro* matured cattle COCs. The initial preliminary results of this trial showed the identification of approximately 4000 peptides and 2000 proteins, a ten-fold improvement compared to DDA technique trialled within our group on a pooled sample of 10 COCs. To expand and explore the limitations and applications of new MS techniques, the oogenesis to embryogenesis may be an ideal platform to assess changeable proteomes dependant on maturational stage. Furthermore, development of practical screening methods to assess these changes in the live individual oocyte and embryo now appears to be a genuine prospect that would improve production efficiency.

825

Proteomic analysis of murine alveolar-like macrophages infected with wild-type *Mycobacterium bovis* BCG

Darien T Schell[1], Naadir Ganief[1], Brandon D Murugan[1], Bridget Calder[1], Tariq Ganief[1], Nelson C Soares[1], Jonathan M Blackburn[1, 2]

1. Division of Chemical and Systems Biology, Department of Integrated Biomedical Science, University of Cape Town, Cape Town, Western Cape, South Africa

2. Institute of Infectious Diseases & Molecular Medicine, University of Cape Town, Cape Town, Western Cape, South Africa

Background

Protein tyrosine phosphatase B (PtpB) is a mycobacterial phosphatase secreted during early infection of macrophages[1]. It has been demonstrated that PtpB is imperative to the intracellular survival of *Mycobacterium tuberculosis*, and infections with PtpB-deficient *M. bovis* BCG results in a reduced capacity to maintain infection[2,3,4]. Previous studies do not necessarily reflect the true primary site of infection, as it has been demonstrated that alveolar macrophages are the site in animal models[5]. Although PtpB

inhibitors exist they don't effectively inhibit PtpB activity. Better inhibitors can be generated through a better understanding of the role PtpB[6]. We, therefore, hypothesise that PtpB assists in the maintenance of infection.

Methods

Confocal microscopy analysis of infected macrophages was used to determine when to sample early infection the images were analysed using ImageJ. Label-free shotgun LC-MS/MS analysis was conducted on alveolar-like macrophage cell-line infected with *M. bovis* BCG wild-type for 0, 3, 11 and 24 hours. The mass spectrometry-based analyses were conducted using MaxQuant, R and Cytoscape.

Results

Analysis of the confocal data indicates that an optimal infection time was 1 hour. Preliminary proteomic results showed 57 differentially expressed proteins between all time points. Functional analysis of the infection shows enrichment of proteins related to signal transduction.

Conclusion

Preliminary evidence suggests that knocking out PtpB alters host signalling, preventing maintenance of infection by the bacterium. These results would be complemented by phosphoproteomic data to identify significant pathways in the infection. Future analysis of the difference in the phosphoproteome of the cell-line infected with *M. bovis* BCG wild-type and ΔPtpB will be done.

1. Wong, D., Chao, J. D. & Av-Gay, Y. Mycobacterium tuberculosis-secreted phosphatases: From pathogenesis to targets for TB drug development. Trends in Microbiology 21, 100–109 (2013).
2. Singh, R. et al. Disruption of mptpB impairs the ability of Mycobacterium tuberculosis to survive in guinea pigs. 50, 751–762 (2003).
3. Fan, L. et al. MptpB Promotes Mycobacteria Survival by Inhibiting the Expression of Inflammatory Mediators and Cell Apoptosis in Macrophages. Frontiers in Cellular and Infection Microbiology 8, (2018).
4. Chauhan, P. et al. Secretory Phosphatases Deficient Mutant of Mycobacterium tuberculosis Imparts Protection at the Primary Site of Infection in Guinea Pigs. PLoS ONE 8, 1–15 (2013).
5. Cohen, S. B. et al. Alveolar Macrophages Provide an Early Mycobacterium tuberculosis Niche and Initiate Dissemination. Cell Host & Microbe 24, 439-446.e4 (2018).
6. Soellner, M. B., Rawls, K. A., Grundner, C., Alber, T. & Ellman, J. A. Fragment-Based Substrate Activity Screening Method for the Identification of Potent Inhibitors of the Mycobacterium tuberculosis Phosphatase PtpB. Journal of the American Chemical Society 129, 9613–9615 (2007).

826

Media dependent proteomic composition of CHO cells affecting IGF-signaling

Louise Schelletter[1], Thomas Noll[1], Raimund Hoffrogge[1]

1. Bielefeld University, Bielefeld, NRW, Germany

In chinese hamster ovary cells specific signaling pathways are largely unknown. Often, optimization approaches of monoclonal antibody production process with CHO cells are untargeted and therefore partially ineffective. To overcome this, signaling effects of cell growth enhancing agents are analysed on molecular level via quantitative MS approaches.

CHO cultivation was performed in chemical defined medium. For SILAC experiments cells were cultivated with l/m/h lysine and arginine. Bottom-up proteomic MS sample preparation, measurement and data analysis was performed. Phospho-targets were evaluated by Western blots of label-free cultivated cells.

Insulin-like growth factor stimulation induced rapid changes in signaling directly linked to cellular metabolism and in parallel, long-term differences in growth and viability was observed [1]. Next to this, media supplementation with the non-essential amino acid L-glutamine resulted in increased growth of mAb-producing CHO cells, but in also in a remarkebly reduced IGF-induced upregulation of phosphorylation sites.

To elucidate this, the proteomic composition of CHO cells growing on two different media is analyzed via triple SILAC-MS (including label-swap conditions). Significantly differential expressed proteins calculated for both label conditions were compared and validated via two quantification workflows, Proteome Discoverer and MaxQuant.

As expected, in glutamine containing media the glutamine synthesis and also parts of glucose metabolism pathways were found to be downregulated. However, also MAPK interacting proteins, like GSTP1, were downregulated and protein phosphatases as well as cytoskeletal proteins were upregulated. Based on this, the cellular reportoire contributing to decreased signaling intensity is presented on a new designed signaling network by expression data mapping.

The combination of validated Western blot-, SILAC- and LFQ-MS-based data led to a first characterization of a specific proteome, which contributes to a shift in IGF-signaling in CHO in media with or without glutamine. With this knowledge, a more targeted process design can be realized and will improve mAb-production efficiency.

1. 1. Schelletter, L., Albaum, S., Walter, S., Noll, T., Hoffrogge, R. (2019). Clonal variations in CHO IGF signaling investigated by SILAC-based phosphoproteomics and LFQ-MS. Applied microbiology and biotechnology (accepted).

827

Quantitative elution of biotinylated peptides and proteins from streptavidin complexes

Martina Schnölzer[1], Uwe Warnken[1]

1. German Cancer Research Center (DKFZ), Heidelberg, Germany

Introduction

Biotin and streptavidin form one of the strongest non-covalent biological bonds. Due to the extraordinary stability of the complex, biotin labeling of molecules has emerged as a widely used technique for numerous applications. However, the stability of the biotin-streptavidin complex poses a major problem when biotinylated molecules need to be recovered. Harsh elution conditions are required to dissociate the complex using heat or chemical reagents. Here we describe a simple and mild method for quantitative dissociation of biotin-streptavidin complexes which is compatible with downstream applications such as mass spectrometry.

Materials and Methods

Biotinylated BSA and transferrin were used alone or in spike-in experiments with protein lysates from HeLa cells. Different elution conditions were applied to release biotinylated proteins and peptides from streptavidin-coated dynabeads (Life Technologies) and from MSIA D.A.R.T.'s streptavidin (Thermo Fisher). Protein eluates were separated by 1D PAGE and band intensities were compared using ImageJ. For peptide elution proteins were digested with trypsin and peptides were incubated with the streptavidin surfaces. Mass spectrometry was used as read-out.

Results

The high affinity of streptavidin to biotin makes complex formation essentially irreversible unless the biotin-streptavidin complex is exposed to harsh conditions such as heat or detergents. We developed a one-step elution strategy using an organic solvent and compared our workflow with different elution protocols. Biotinylated BSA and transferrin or biotinylated tryptic peptides were bound to either streptavidin coated magnetic beads or streptavidin coated tips. Elution of biotinylated molecules with the organic solvent was almost quantitative as shown by 1D PAGE and mass spectrometry. In contrast to other elution methods which require further purification steps, removal of the organic solvent from the eluates is simply achieved by lyophilisation in a speed-vac concentrator.

Conclusion

Our quantitative elution protocol is directly applicable to mass spectrometry without the need for any further workup procedures.

828

Unconventional epitope discovery: a novel bioinformatics workflow for immunopeptidogenomics

Katherine E Scull[1], Kirti Pandey[1], Sri H Ramarathinam[1], Anthony W Purcell[1]

1. Department of Biochemistry and Molecular Biology and Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Background: Human leukocyte antigen (HLA) molecules are cell-surface glycoproteins that present peptide antigens for surveillance by T lymphocytes seeking signs of disease. Mass spectrometric analysis allows us to identify large numbers of these peptides (the immunopeptidome) following affinity purification from cell lysate. However, in recent years there has been a growing awareness of the 'dark side' of the immunopeptidome: unconventional peptide epitopes that elude detection by conventional search methods because their sequences are not present in reference protein databases.
Methodologies: Here we establish a bioinformatic workflow to aid identification of peptides generated by non-canonical translation of mRNA. The workflow incorporates both standard transcriptomics software and novel computer programs to produce cell line-specific protein databases based on 3-frame translation of the transcriptome. Optionally, known mutations can be included to produce an 'alternate' transcriptome and corresponding imputed protein database. We then search our experimental data against both transcriptome-based and standard databases using PEAKS Studio. Finally, further novel software helps to compare the various result sets arising for each sample, and pinpoint putative genomic origins for the identified unconventional sequences.
Results: We have trialled the workflow to study the immunopeptidome of the acute myeloid leukaemia cell line THP-1, starting with THP-1 RNA-Seq data downloaded from the Sequence Read Archive. We confidently identified over 8000 peptides from 3 replicates of purified THP-1 HLA peptides using Swissprot. Using the transcriptome-based databases, we recapitulated >70% of these, and also identified over 250 unconventional peptides, many of which might be generated by translating UTRs or the 'wrong' frame.
Conclusions: Our workflow, which we term 'immunopeptidogenomics', can provide databases which include pertinent unconventional sequences, and can also be tailored for neoepitope discovery in cancer, without becoming unsearchably large. Immunopeptidogenomics is a step towards the unbiased search approaches needed to illuminate the dark side of the immunopeptidome.

829

Generating high-quality chromatogram libraries for DIA-MS with empirically corrected peptide predictions

Brian C Searle[1, 2], Kristian E Swearingen[1], Christopher A Barnes[3], Tobias Schmidt[4], Siegfried Gessulat[4], Bernhard Kuster[4], Mathias Wilhelm[4]

1. Institute for Systems Biology, Seattle, WA, United States

2. Proteome Software, Portland, OR, USA

3. Novo Nordisk, Seattle, WA, USA

4. Technical University of Munich, Munich, Germany

Data-independent acquisition (DIA) is a powerful technique for deep, proteome-wide profiling. DIA methodologies often rely on sample-specific spectrum libraries from data-dependent acquisition (DDA) experiments. This approach produces high-performance DDA libraries with instrument-specific fragmentation and retention times at the expense of time, sample, and significant offline fractionation effort. Previously we demonstrated a DIA-only workflow that built chromatogram libraries by searching gas-phase fractionated (GPF) DIA runs with PECAN, a FASTA search engine. However, we found success varied due to the lack of fragmentation and elution information, as well as a substantially increased search space. Here we leverage fragmentation prediction by Prosit to generate chromatogram libraries by replacing the retention times and fragmentation in the predicted library with sample- and instrument-specific empirical values found in the six runs.

We benchmarked our workflow in yeast with ten Prosit-predicted libraries at various normalized collision energy (NCE) settings. While spectrum prediction accuracy is highly dependent on NCE and has significant effect when matching to single-injection DIA, GPF-DIA is less sensitive to spectrum library quality and can produce filtered libraries of equal size to a sample-specific 10-fraction DDA library with high-pH reverse-phase at a wide range of NCE settings. We find that the filtered chromatogram libraries have both more accurate fragmentation and retention times than the DDA library because GPF-DIA fragmentation patterns match to wide-window DIA closer than DDA, and GPF does not affect chromatographic interactions with matrix.

We applied our workflow to analyze gametocyte cultures of Plasmodium falciparum, the parasite responsible for 50% of all malaria cases. We were able to detect parasite peptides in up to 1:100 dilution with uninfected red blood cells in wide-window DIA experiments, while maintaining higher quantification accuracy than comparable DDA experiments. In conclusion, our approach to library generation produces high-quality, sample-specific libraries without offline fractionation using only six GPF-DIA runs.

830

Isomer-Specific Quantitation of Serum Acidic N-Glycans for Behcet's Disease Screening

Nari Seo[1, 2], Jaehan Kim[3], Hyun Joo An[1, 2]

1. GRAST, Chungnam National University, Daejeon, South Korea

2. AGRS, Daejeon, South Korea

3. Department of Food and Nutrition, Chungnam National University, Daejeon, South Korea

Aberrant glycosylation is a well-known event in autoimmune diseases and cancers. Indeed, several glycans are widely recognized as the powerful biomarker for screening disease states. Bechet disease (BD) is a type of inflammatory disorder include painful mouth sores, genital sores, and arthritis. Although many diagnostic criteria of BD have been proposed, it is still challenging to establish reliable diagnosis method. Here, we created glycan isomeric-based novel approach for inflammation disease monitoring. Unlike the conventional glycomic approach with global mass spectrometric profiling of serum released glycan, we have developed new platform to quantify isomer-specific glycan using PGC/MRM-MS. Especially, we have targeted sialylated glycans which are known to highly present in immunoglobulin proteins closely related with inflammation diseases. Serum N-glycans released from BD patients (n=47) and heathy control (n=47), respectively were selectively fractionated by PGC-SPE prior to MS analysis. Targeted acidic glycans were chromatographically separated to obtain isomer-specific information and relatively quantified by PGC-MRM MS. Total eleven N-glycan isomers were monitored for the diagnosis of BD. We found that BD patients contained significant quantity of mono/di-sialylated bi-antennary N-glycans consisting of Hex_5 $HexNAc_4$ $NeuAc_{1-2}$ compared with healthy control group. Expression level in patient cohort were 5 to 10-folds higher than those in control group showing complete separation. Four bi-antennary acid glycan isomers showed high diagnostic efficacy providing AUC of ROC curve over 0.98. Interestingly, these marker glycans exhibits unique correlations with biological characteristics of cohorts. Specially, two isomers of Hex_5 $HexNAc_4$ $NeuAc_1$ and two isomers of Hex_5 $HexNAc_4$ $NeuAc_2$ could completely distinguish male and female groups in healthy control cohort. Glycan expression level in female groups were 6 to 10 times higher than that of male groups resulting in 1.0 AUC in ROC curve. Our novel approach that targeting isomer-specific glycan will provide significant insight into the glycobiological aspects of the BD process.

831

The role of upstream phosphorylation in the regulation of histone methylation

Ryan J Separovich[1], Joshua J Hamey[1], Marc R Wilkins[1]

1. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia

Histone methylation is a central means by which gene expression is controlled. In the lower eukaryote, *Saccharomyces cerevisiae,* histone methylation is regulated by a reduced, but evolutionarily conserved set of methyltransferases (Set1, Set2, Set5, Dot1) and demethylases (Jhd1, Jhd2, Rph1, Gis1). While the catalytic activity and specificity of these enzymes have been established, knowledge of how they themselves are regulated by post-translational modification is surprisingly limited. Consequently, the regulatory network of histone methylation in yeast remains unknown and is also unknown in all other eukaryotes. To this end, we aimed to comprehensively characterise the modifications occurring on the eight yeast histone

methyltransferases and demethylases *in vivo.* This was achieved by purification of these proteins, and their analysis by targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS). With respect to phosphorylation, to date, we have identified modification sites on the histone methyltransferases Set5 (nine sites) and Dot1 (three sites), and the demethylases Gis1 (seven sites) and Jhd2 (two sites). Fourteen of these sites validate those observed previously in high-throughput phosphorylation screens, while seven sites are novel. To determine the upstream kinases responsible for the phosphorylation, and potential regulation of these enzymes, mass spectrometric analysis was employed to monitor levels of histone methylation in kinase knockout yeast strains. As a proof of concept, quantification of H3K79 methylation in the knockout cells established twenty-five kinases that are not responsible for the regulation of Dot1 methyltransferase activity. The screening of all other non-essential kinases is in progress. We plan to extend this methodology to the other yeast histone methyltransferases and demethylases in order to comprehensively integrate these enzymes into intracellular signalling pathways, and ultimately facilitate the assembly of the first regulatory network of histone methylation in eukaryotes.

832

LFQ-Analyst: An easy-to-use interactive web-platform to analyze and visualize quantitative proteomics data

Anup D. Shah[1, 2], Robert J.A Goode[1], Cheng Huang[1], David R. Powell[2], Ralf B. Schittenhelm[1]

1. Monash Proteomics and Metabolomics Facility, Monash University, Clayton, VIC, Australia

2. Monash Bioinformatics Platform, Monash University, Clayton, VIC, Australia

Relative label-free quantification (LFQ) of shotgun proteomics data using precursor (MS1) signal intensities is one of the most commonly used applications to comprehensively and globally quantify proteins across biological samples and conditions. Due to the popularity of the technique, several software suites – such as MaxQuant – have been developed to extract, analyze and compare spectral features, and to report quantitative information of peptides, proteins and even post-translationally modified (PTM) sites. However, there is still a lack of accessible tools for the interpretation and downstream statistical analysis of these complex datasets, in particular for researchers and biologists with no or only limited experience in proteomics, bioinformatics and statistics.

We have therefore created *LFQ-Analyst*, which is a web application developed to perform differential expression analysis with "one click" and to visualize label-free quantitative proteomic datasets preprocessed with the popular software suite MaxQuant. *LFQ-Analyst* provides a wealth of user-analytic features including differential expression analysis, dimensionality reduction, clustering and various quality control checks in tabular and graphical format to facilitate exploratory and statistical analysis of label-free quantitative datasets. *LFQ-Analyst* is freely available at https://bioinformatics.erc.monash.edu/apps/LFQ-Analyst

833

Developing a mass spectrometry-based immunopeptidomics platform to analyze thermostability profiles of peptides bound to HLA class I molecules

Mohammad Shahbazy[1], Emma C. Jappe[2, 3], Sri Ramarathinam[1], Pouya Faridi[1], Christian Garde[2], Jens Kringelum[2], Anthony Purcell[1], Nathan P. Croft[1]

1. Infection and Immunity Program, Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Clayton, Victoria, Australia

2. Evaxion Biotech, Bredgade 34E, DK-1260 Copenhagen, Denmark

3. Department of Bio and Health Informatics, Technical University of Denmark, DK-2800 Lyngby, Denmark

Publish consent withheld

834

Identifying neoantigens with LC-MS by personalized de novo peptide sequencing

Hieu Tran[1], Rui Qiao[1], Lei Xin[2], Xin Chen[2], Baozhen (Paul) Shan[2], Ming Li[1]

1. Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada

2. Bioinformatics Solutions, Waterloo, ONTARIO, Canada

1. Introduction

Currently, candidate neoantigens could be identified with a proteogenomics method that involves two major steps: whole exome sequencing and RNA sequencing to find somatic mutations, and mass spectrometry to find neoantigens by database search. In this study, we propose a workflow to identify neoantigens solely from mass spectrometry data by integrating de novo sequencing and database search. In addition, deep learning model for de novo sequencing is tailed to each individual patient based on their own MS data (personalized proteome). Such a personalized approach enables faster and more accurate identification of neoantigens for personalized vaccines.

1. Methods

Our personalized de novo sequencing based neoantigen finding workflow involves five steps:

(1). LC-MS/MS data from a patient was searched against canonical database. The identified normal HLA peptides represent the patient's immunopeptidome. Spectra generated by mutated peptides remain unmatched.

(2). DeepNovo [1], a neural network model, was used for de novo peptide sequencing. The model was trained with the identified spectrum acquired from the first step. This unique advantage allows DeepNovo to adapt to a specific immunopeptidome of an individual patient.

(3). Perform de novo sequencing on those unidentified spectrums with the trained model, and only keep high-confidence de novo peptides with an expected mass tag accuracy of 95%.

(4). Quality control of de novo peptides.

(5). Select candidate neoantigens.

1. Results

We test our workflow on a public available dataset from a melanoma patient [2]. MS/MS data was searched against UniProt database with PEAKS X at 1% of false discovery rate (FDR). After the aforementioned 5 steps, our workflow reported 158 HLA class I candidate neoantigens, which include 4 out of the 5 neoantigens identified by the current proteogenomics method at 1% FDR. More importantly, our workflow identified an extra neoantigen that matched a nucleotide mutation reported by RNA-seq.

1. Tran, N.H, Zhang, X., Xin, L., Shan, B. & Li, M. De novo peptide sequencing by deep learning. Proc. Natl. Acad. Sci. U. S. A. 114, 8247-8252 (2017). 2. Van Allen, E.M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science 350, 207-211 (2015).

835

Novel Method for Proteome-scale Analysis of Protein Substrates and Modification Sites of Small Ubiquitin-like Modifier

Yichu Shan[1], Yang Li[1], Lihua Zhang[1]

1. Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, LIAONING, China

Small Ubiquitin-like Modifier(SUMO) plays important roles in regulating many physiological activities of cells. Accurate identification and quantification of the SUMO substrates and modification sites at proteome scale is helpful for studying the molecular mechanisms of SUMO modification and its' influence on the functions of protein substrates. However, there is no method for the selective enrichment and efficient and accurate identification of wild-type SUMO modified peptides in various kinds of protein samples. In this work, a method for enrichment of endogenous SUMOlyated peptides based on anion chromatography and antibody was developed. Standard SUMO modified peptides library was synthesized and high efficient and accurate algorithms for the identification of SUMO modified peptides was developed by investigating their fragmentation patterns in mass spectrometer. The developed method was applied to the proteome-scale analysis of SUMO modified proteins and modification sites to the analysis of SUMO substrates and their modification sites in HeLa cells, which provides important technical support for disclosing the regulation networks of SUMO modification and studying the functions of SUMO modification in biological processes.

836

Comparative analysis of mRNA and protein degradation in prostate tissues indicates high stability of proteins

Wenguang Shao[1], Tiannan Guo[2], Nora Toussaint[1], Peng Xue[1], Ulrich Wagner[3], Yi Zhu[2], Marija Buljan[1], Jianmin Wu[4], Andreas Beyer[5], Ben Collins[1], Yansheng Liu[6], Gunnar Rätsch[1], Peter Wild[7], Ruedi Aebersold[1]

1. ETH Zurich, Zurich, ZURICH, Switzerland

2. Westlake University, Hangzhou, China

3. Institute of Surgical Pathology, University Hospital Zurich, Zurich, Switzerland

4. Center for Cancer Bioinformatics, Peking University Cancer Hospital & Institute, Beijing, China

5. University of Cologne, Cologne

6. Cancer Biology Institute, Yale University School of Medicine, West Haven, USA

7. University Hospital Frankfurt, Frankfurt am Main, Germany

Deterioration of biomolecules in clinical tissues is an inevitable pre-analytical process, which affects molecular measurements and thus potentially confounds conclusions from cohort analyses. While mRNA degradation has been studied extensively, the extent of protein degradation in human tissues and its implication for molecular classification remain largely unexplored. Here, we comprehensively compare the degradation patterns of mRNA and protein in a prostate cancer tissue cohort. We subject 68 pairs of adjacent prostate tissue samples to RNA-Sequencing (RNA-Seq) or proteomic analysis by pressure cycling technology (PCT) coupled with SWATH mass spectrometry. To objectively quantify the extent of protein degradation, we develop a numerical score, the Proteome Integrity Number (PIN), that faithfully measures the degree of protein degradation. We benchmark the PIN algorithm using a set of ground-truth samples in which the levels of proteome degradation were artificially controlled and independently validated, and assess the relative degree of mRNA and protein degradation in adjacent samples. Our results indicate that the PIN score faithfully indicates the degree of protein degradation of a sample and is robust across different types of clinical samples and mass spectrometric measurement methods. We show that protein degradation only affects 5.9% of the samples tested and shows negligible correlation with mRNA degradation in the adjacent samples. These findings are confirmed by independent analyses on additional clinical sample cohorts and across different mass spectrometric methods.

Concluding statement

Overall, the data show that the majority of samples tested are not compromised by protein degradation, even if the corresponding samples show substantial transcript degradation and establish the PIN score as a generic and accurate indicator of sample quality for proteomic analyses. Our results thus provide important information and resources for proteomic measurements in clinical cohort studies.

1. Shao W, et al. Nat Commun. 2019 Jun 7;10(1):2524. doi: 10.1038/s41467-019-10513-5

837

CRCprot: a novel and practical protein-based classification system for CRC prognosis using FFPE biopsy tissues

Yingkuan Shao[2, 1], Kailun Xu[2, 1], Xi Zheng[2, 1], Fangfei Zhang[1], Qi Yuan[1], Biting Zhou[2], Ziqi Li[1], Xue Cai[1], Yi Zhu[1], Shu Zhen[2], Tiannan Guo[1]

1. Westlake University, Hangzhou

2. Cancer institution, Zhejiang University, Hangzhou, ZHEJIANG, China

Background Colorectal cancer (CRC) is the third most common cancer worldwide and the fourth leading cause of cancer-related deaths. Genomic and transcriptomic classification system has been developed for heterogeneous CRC tumors, but their applications in clinical tissues are limited due to the degradability of mRNAs. No protein-based classification system has been reported.

Methods We analyzed the proteome of FFPE biopsy samples from 217 CRC patients with up to ~9 years survival using pressure cycling technology (PCT) and data-independent acquisition (DIA) mass spectrometry. Then we trained a model using deep neural network. An independent CRC cohort of 117 patients was further procured to validate the protein-based classifier.

Results We quantified > 8000 proteins from 490 FFPE samples including 88 biological replicates (r = 0.77) and 66 technical replicates (r = 0.95). Using machine learning technology, we established a novel and practical protein-based classification system, containing the expression of about 10 ten proteins, for CRC prognosis which was further verified in an independent validation cohort.

Conclusion We demonstrated the practicality of PCT-DIA for analyzing large number of FFPE biopsy samples from multiple cohorts and established a novel and practical protein-based classification system CRCprot for CRC prognosis.

838

The interactome map of the chromatin organising protein CTCF

Mehdi M Sharifi Tabar[1], Chuck CG Bailey[1], john EJ Rasko[1]

1. The Centenary Institute of Cancer Medicine and Cell Biology,, SYDNEY, NSW, Australia

In eukaryotes, CTCF, a zinc finger DNA-binding protein, plays an important role in gene regulation through organising the three-dimensional architecture of the genome. Over the past two decades, there have been a large number of genomic studies investigating the role of CTCF in genome organisation of various cancer and normal tissues. However, the molecular mechanisms of CTCF function are not well understood. To understand the function of CTCF a map of interaction partners is needed. Here, we set up a two-step purification protocol coupled to liquid chromatography-tandem mass spectrometry to assess CTCF interaction partners in three different cancer cell lines, using label-free intensity-based absolute quantification (iBAQ). In addition, our interactome data suggest numerous proteins involved in gene regulation and chromatin remodeling as possible interaction partners of CTCF. To corroborate the high throughput interactome data, in vitro and cell-based pull-down experiments were performed.

839

Problems to be solved for more accurate SWATH analysis

SHUICHIRO SHIMADA[1], Bo Xu[1], Yoshitoshi Hirao[1], Keiko Yamamoto[1], Tadashi Yamamoto[1, 2]

1. Biofluid Biomarker Center, Niigata University., Niigata, NIIGATA, Japan

2. Shinrakuen Hospital, NIIGATA, JAPAN

Recent advances of MS technologies promote an increase in the number of proteins or peptides identified and quantified. For biomarker discovery the sequential window acquisition of all theoretical fragment ion spectra (SWATH) is a strong tool for label-free proteome quantification with data-independent acquisition (DIA). In the SWATH, a reference library created by data-dependent acquisition (DDA) runs is used for identification of proteins and peptides measured in the DIA by correlating the quantitative DIA data obtained from the fragment ion spectra data with the library information for peptide identification. Although the quality and coverage of the reference library are essential for SWATH data processing and the quality of data independent acquisition (DIA) is also important, there are several problems to be solved for the ideal accurate label-free proteome quantification by SWATH.

We analyzed proteins extracted from human urine and performed SWATH DDA and DIA analysis by LC-MS (TripleTOF 6600, SCIEX). The libraries were made by using different numbers of DDA data at different conditions. The SWATH analysis data obtained with these different libraries were compared each other to point out some problems for accurate quantification.

In this presentation, we will point out the several problems of the SWATH method and propose suggestions for improvement of the SWATH data accuracy. Our study may give an insight into how to make a better library and perform SWATH analysis more accurately.

840

Quantitative proteomic analysis of breast cancer formalin-fixed paraffin-embedded(FFPE) tissue proteins associated with distant metastasis.

Dongyoon Shin[1], Joonho Park[2], Hansuk Ryu[3], Youngsoo Kim[1, 2]

1. Biomedical science, Seoul National University, college of Medicine, Seoul, South Korea

2. Interdisciplinary Program for Bioengineering, Seoul National University, College of Engineering, Seoul, South Korea

3. Department of pathology, Seoul National University Hospital, Seoul, South Korea

Metastasis of breast cancer to other distant organs is fatal to patients. However, few studies have revealed biomarkers associated with distant metastatic breast cancer. Furthermore, accurately differentiating distant metastatic breast cancer from non-distant metastatic ones remains a diagnostic challenge with current biomarkers such as HER2, ER, and PR, necessitating the development of novel biomarkers. An integrated proteomics approach that combines filter-aided sample preparation, tandem mass tag labeling(TMT), high pH fractionation, and high resolution MS was applied to acquire in-depth proteome data of breast cancer FFPE tissue. Biological processes of differentially expressed proteins (DEPs) that may be involved in distant metastasis were identified through bioinformatics analyses such as gene ontology analysis and pathway analysis. In addition, antibody-based protein assays were performed to validate the differential regulation of biomarker candidates. A total of 9,190 and 8,564 proteins were identified from the series of TMT experiments. Bioinformatics analysis of DEPs drew several biological processes such as cell-cell adhesion, proteolysis involved in cellular protein catabolic process, microtubule-based process and positive regulation of protein kinase B signaling. In addition, distinct molecular features between breast cancer subtypes were investigated. This study suggests novel biomarker candidates and their functional characteristics of distant metastatic breast cancer.

841

Multiplexed immuno N-terminomics Enables Time-resolved Profiling of Protein Degradation

Sanghee Shin[1]

1. Seoul National University, Gwanak-gu, SEOUL, South Korea

Protein expression is maintained through the regulation of both translation and proteolysis. Since levels of degradation intermediates, i.e., partially degraded proteins, cannot be distinguished from those of stable proteins in global proteomics, which quantifies total protein levels, the mechanism underlying the regulation of protein degradation would be underrepresented using only global proteomics approaches. This study aimed to assess the unclear aspects of degradational regulation through a new multiplexed N-terminomics method involving selective isobaric labeling on protein N-termini and immunoaffinity capture of the labeled N-terminal peptides. Our method allows for not only identification of proteolytic cleavage sites, but also highly multiplexed quantification of degradation intermediates, enabling time-resolved profiling of protein degradation. The present method uniquely highlights the regulation of proteomic degradation rate during early embryogenesis in *Drosophila melanogaster*; furthermore, our method revealed that a group of zygotically expressed proteins are also expressed during embryogenesis but were stringently regulated through active degradation to maintain baseline expression levels.

842

Peptidomics and metabolomics approach to elucidate the proteolytic regulation of haemoglobin peptides within the malaria parasite

Ghizal Siddiqui[1], Amanda E Depaoli[1], Carlo Giannangelo[1], Nyssa Drinkwater[2], Sheena McGowan[2], Darren J Creek[1]

1. Drug Delivery, Disposition and Dynamics, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville Campus, Parkville, Victoria, Australia

2. Biomedicine Discovery Institute and Department of Microbiology, Monash University, Clayton, Victoria, Australia

Plasmodium falciparum parasite infects human red blood cells causing the most severe and life-threatening form of malaria in humans. In the human red blood cells, the parasite absolutely requires host haemoglobin digestion to supply amino acids for *de novo* protein synthesis. Digestion of haemoglobin occurs via a semi-ordered proteolytic cascade that is mediated by different proteases, and the exact interplay between the different classes of proteases and haemoglobin digestion is not clear. This limitation prompted us to develop an advanced metabolomics and peptidomics-based strategy to identify the 'signature peptide' libraries that are generated as a result of specific inhibition of each protease class involved in haemoglobin digestion.

The combined metabolomics and peptidomics analysis revealed that inhibition of specific haemoglobin digesting proteases resulted in unique changes in the abundance of haemoglobin derived endogenous peptides. Furthermore, metabolomics analysis of parasites cultured in media with labelled amino acids confirmed that the short chain peptides associated with protease inhibition were derived from haemoglobin. The results obtained could distinguish between different clans of the same class of protease. For example, inhibition of M1 aminopeptidase resulted in accumulation of peptides containing basic residues, while specific inhibition of M17 aminopeptidase resulted in accumulation of hydrophobic peptides.

In conclusion, this multi-platform approach provided an extensive coverage of endogenous peptides liberated during haemoglobin digestion within the parasite, and identified specific peptide signatures associated with inhibition of different classes of proteases involved in haemoglobin digestion.

843

Proteomics identification of radiation-induced changes of membrane proteins in the rat model of arteriovenous malformation in pursuit of targets for brain AVM molecular therapy

Margaret Simonian[1, 2, 3], Dyna Shirasaki[1], Vivienne Lee[2], David Bervini [4], Michael Grace[5], Rachel Ogorzalek Loo[1], Joseph Loo[1], Mark Molloy[6], Marcus Stoodley [2]

1. UCLA, Los Angeles, CALIFORNIA, United States

2. Department of Clinical Medicine, Faculty of Medicine and Health Sciences, , Macquarie University,, Sydney,, NSW, , Australia.

3. David Geffen school of medicine, UCLA, Los Angeles, CA, United States

4. Neurosurgery Department,, Bern University Hospital,, Bern,, Switzerland

5. Genesis Cancer Care, , Macquarie University Hospital , Sydney, , NSW,, Australia

6. Lawrence Penn Chair of Bowel Cancer Research,, Faculty of Medicine and Health, Northern Clinical School, Sydney, , NSW, Australia

Abstract

Background

Rapid identification of novel targets and advancement of a vascular targeting strategy requires a comprehensive assessment of AVM endothelial membrane protein changes in response to irradiation. This aim of this study is to provide additional potential target protein molecules for evaluation in animal trials to promote intravascular thrombosis in AVM vessels post radiosurgery.

Methods

We employed *in vivo* biotinylation methodology that we developed, to label membrane proteins in the rat model of AVM post radiosurgery. Mass spectrometry expression (MS^E) analysis was used to identify and quantify surface protein expression between irradiated and non irradiated rats, which mimics a radiosurgical treatment approach.

Results

Our proteomics data revealed differentially expressed membrane proteins between irradiated and non irradiated rats, e.g. Profilin-1, ESM-1, ion channel proteins, Annexin A2 and lumican.

Conclusion

This work provides additional potential target protein molecules for evaluation in animal trials to promote intravascular thrombosis in AVM vessels post radiosurgery

Key words

In vivo biotinylation, AVM molecular therapy, membrane proteins, radiosurgery

844

PiRT: cross-species retention time calibrants for plasma samples

Manika Singh[1], Selvam Paramasivan[2], Pawel Sadowski[1]

1. Central Analytical Research Facility, Queensland University of Technology, Brisbane, QLD, Australia

2. School of Veterinary Science, University of Queensland, Gatton, QLD, Australia

Peptide Retention Time Calibration Mixtures enable assessing chromatography and MS instrument performance. More importantly, they are paramount for normalizing results for variation in retention times and peak intensities between runs. Commercial mixes are expensive and cannot be added retrospectively in the case of datasets that were not spiked in the first place. It is, therefore, desirable to identify endogenous peptides that can be used for that purpose. Parker et al (2015) reported a set of peptide sequences that are conserved across eukaryotic species which they termed Common internal Retention Time standards (CiRT). We identified the limitation of CiRT which, being mostly derived from cytoskeletal and ribosomal proteins, are not detectable in blood plasma. Plasma is the most commonly used sample in both human biomedical research and for clinical veterinary applications and both animal and translational biomedical research, would benefit from the availability of calibration peptides. We have, therefore, conducted a bioinformatics study to develop a novel set of endogenous peptides which could serve the same purpose as CiRT but which are specifically present in sufficient abundance in plasma of multiple species including human.

Initially, we selected a list of 81 peptides from the in-house study on cross-species plasma proteome. We then compared this set of identified peptides with human and mouse plasma repository downloaded from PeptideAtlas. To further confirm the ubiquitous occurrence of peptides across a larger set of mammals we downloaded 12 species reference proteome from UniProt Knowledgebase. This resulted in the identification of a set of 13 peptides and called them as Plasma internal retention Time calibration peptides (PiRT). We have validated this new set of peptides by extracting them from plasma samples collected from 5 different and distantly-related species (human, mouse, cattle, sheep, giraffe) and showing that they can be used for retention time calibration.

845

Differential protein expression in *Pelistega indica* under anaerobic vs aerobic conditions

Om Prakash[1], Roger Mercer[1], Rakesh Singh[1]

1. Florida State University, Tallahassee, FLORIDA, United States

Recent studies related to human gut microbiota have revealed the direct or indirect role of gut microbes in human health and disease. Gut microbes are involved in various activities such as nutrient metabolism, degradation of xenobiotics and drugs, immunomodulation, prevention of pathogen invasion and maintenance of gut mucosal barrier. *Pelistega indica* is a Gram negative, non-spore-forming, facultative anaerobe isolated from human fecal material. We conducted comparative proteomic study of *Pelistega indica* grown under aerobic and anaerobic conditions on Q-Exactive HF. Whole cell protein was extracted from cultured bacteria and the proteome was analysed using nLC-MS/MS. We detected more than 1200 proteins with high confidence. The data will be presented indicating significant differences in specific pathways between aerobic and anaerobic conditions.

846

Multi-omics analyses reveal temporally distinct metabolic switches, carbon-nitrogen partitioning and oxidative signaling in chickpea seed

Arunima Sinha[1], Toshiba Haider[1], Kanika Narula[1], Sudip Ghosh[1], Niranjan Chakraborty[1], Subhra Chakraborty[1]

1. National Institute of Plant Genome Research, New Delhi, DELHI, India

Background

Nutrient dynamics in storage organs is a complex developmental process that requires coordinated interactions of environmental, biochemical, and genetic factors. It correlates with metabolically programmed progressive differentiation of genetically distinct compartments in seed. Nutrient signals and metabolic adaptations determine differentiation pattern and transition from maternally-controlled embryonic growth to maturation under filial regulation. Although sink organ developmental events have been identified, our understanding of transcriptional, translational, post-translational and metabolic regulation of reserve synthesis, accumulation and utilization is limited.

Method

Chickpea seeds were collected at different developmental stages (7-60 DAF) and germination stage. RNA-seq was performed using Illumina Hi-seq 2000 paired-end sequencing technology. Proteome and phosphoproteome were developed using 2-DE and subsequent Pro-Q Diamond staining. Further, TiO2 based phosphopeptide enrichment was done followed by identification of phosphoproteins using TripleTOF mass spectrometer. Integrated global network was built using cytoscaape. Furthermore, qRT-PCR analysis was performed to validate the omics datasets.

Results

To understand nutrient dynamics during embryonic and cotyledonary photoheterotrophic transition to mature and germinating autotrophic seeds, an integrated transcriptomics, proteomics and phosphoproteomics study in six sequential seed developmental stages in chickpea was performed. Differential gene expression analysis led to the identification of 6582 nutrient-associated transcripts predominantly involved in primary metabolism in synthesis phase, while downregulation of these pathways characterise the accumulation phase. Resume of central metabolism was observed in nutrient utilization phase. MS/MS analyses identified 175 and 78 nutrient-associated proteins/phosphoproteins (NAPs/NAPPs) related to metabolism, storage and biogenesis, and protein turnover. Identification of site-specific phosphorylation of amino acids indicated their possible effect in nutrient dynamics. Network analyses identified three significant modules centered around HSP70, vicilin, chalcone synthase and SBP65.

Conclusions

Our study identified several potentially interesting nutrient-associated transcripts and proteoforms of biological significance. Altogether, these findings demonstrate that nutrient signals act as metabolic and differentiation determinant governing storage organ reprogramming.

847

A new brain-centric atlas with expression maps in human, pig and mouse brain.

Evelina Sjöstedt[1], Wen Zhong[2], Linn Fagerberg[2], Max Karlsson[2], Nicholas Micsios[1], Per Oksvold[2], Fredrik Edfors[2], Agnieszka Limiszewska[1], Csaba Adori[1], Feria H Noraddin[3], Cecilia Lindskog[3], Fredrik Pontén[3], Yonglun Luo[4], Tomas Hökfelt[1], Mathias Uhlén[1, 2], Jan Mulder[1]

1. Karolinska institutet, Solna, STOCKHOLM, Sweden

2. Royal institutet of Technology (KTH), Solna, Sweden

3. Immunology, Genetics and Pathology, Uppsala university, Uppsala, Sweden

4. Center of regenerative medicine, BGI, Shenzhen, China

The Human Protein Atlas (HPA) was initiated in 2003 and aims to map all protein-coding genes in human tissues and cells. Transcriptomics data and antibody based protein localization is integrated and publicly available in an open access knowledge resource (www.proteinatlas.org). In 2015 we published an expression based tissue classification, identifying 1684 genes as brain elevated compared to all other tissue types. Since then, external data has been integrated and additional RNAseq analysis added, including the pig and mouse brains. These efforts aim to achieve a complete overview capturing the complex molecular organization of the brain. All integrated data of various sources, are now combined into a brain-centric sub atlas as part of the HPA. Quantitative RNA expression profiles provide regional overview of human, pig and mouse brain, enabling species comparison across the different brain regions. Global expression comparisons show a clustering of the brain regions related to the developmental origin, the regions of brainstem and cerebrum are for instance separated. By investigating regional specific

expression we learned that cerebellum is the most unique brain region from multiple perspectives, both between regions as well as between species. In addition, we also identified several species differences related to the olfactory system. Although most neurotransmitter systems are well preserved between the species we identified several receptor genes with a different regional expression profile when comparing species, highlighting the importance of species comparison and verification when selecting a relevant animal model. All expression data as well as protein profiles for selected targets can be found on gene specific pages at www.proteinatlas.org/brain (which is released the first week of September 2019). Additionally, interactive summary pages provide an exploratory platform to further investigate regionally elevated expression, cell type related proteins and brain elevated expression in the human body.

848

Size exclusion chromatography protein correlation profiling paired with crosslinking mass spectrometry for enhanced identification of yeast protein-protein interactions

Daniela-Lee Smith[1], Gene Hart-Smith[1], Marc Wilkins[1]

1. UNSW, Sydney, NSW, Australia

Whole proteome crosslinking mass spectrometry is now becoming feasible. The latest mass spectrometry approaches have dramatically improved crosslink identification in complex samples. However, for optimal large-scale analysis, major considerations remain how to decomplexify whole cell lysates to maximise crosslink identification and how to bring biological context to all the interactions that are found. Here we address both considerations by combining size exclusion chromatography and protein correlation profiling with crosslinking mass spectrometry. The application of size exclusion chromatography to intact native yeast complexes firstly decomplexifies the sample to enhance data acquisition and identifications within the mass spectrometer and, secondly, maintains biological context for protein-protein interactions. Wild-type yeast was subjected to offline size exclusion high-performance liquid chromatography, followed by crosslinking with the mass spectrometry cleavable crosslinker DSSO. 70 fractions were then analysed using LC-MS/MS CID+ETD/MS CID on a Fusion Lumos, for 180 min each fraction. Crosslink identification was performed using Proteome Discoverer 2.2 node, XlinkX. Across 70 fractions we detected 2,217 crosslinks representing 1,944 protein-protein interactions (FDR 5%). The yeast nucleosome, RNA polymerase, ribosome, oligosaccharyltranferase complex and vacuolar ATPase are examples of complexes detected in the experiment. Preliminary benchmarking analyses have shown that 208 of the total crosslinks found can be mapped onto known crystallographic structures. Of these, 90.4% are within the distance constraints of the crosslinker. Interestingly, mapped crosslinks exhibited a wide range of XlinkX scores (30-400). Therefore, score may not present an absolute metric for crosslink quality. Overall, our approach represents a novel approach for detecting crosslinks from intact complexes and has generated a foundational dataset of crosslinked protein-protein interactions in yeast.

849

Adapting EasyPep™ MS Sample Preparation and TMT labeling for 96-well Automated Liquid Handling Systems

Sergei Snovida[1], Ryan Bomgarden[1], Amarjeet Flora[1], Xinyu Zhang[2], Emily I. Chen[2], John Rogers[1], Kay Opperman[1]

1. Thermo Fisher Scientific, Rockford, IL, United States

2. Precision Medicine Science Center, Thermo Fisher Scientific, Cambridge, MA, United States

Introduction

Advances in mass spectrometry (MS) instrumentation have enabled routine analysis of complex protein samples. However, sample preparation methods are not standardized with many protocols taking 8-24hrs in addition to suffering from low peptide yields, poor digestion efficiency and limited scalability. Recently, we developed a new, simplified sample prep kit containing pre-formulated reagents and a standardized protocol for processing 10µg to 100µg protein samples in less than 2 hours. In this study, we adapted our chemistry to use an automated liquid handling system for sample preparation, TMT labeling, and a 96-well filter plate for peptide clean up.

Methods

Cellular protein extracts and human plasma samples were diluted in lysis buffer. A universal nuclease was added to cellular extracts to reduce sample viscosity. Protein samples were reduced/alkylated before digestion using a trypsin/LysC protease mixture. After labeling with Thermo Scientific™ TMTpro™ 16-plex reagents, a mixed mode peptide clean-up procedure using a 96-well filter plate format was evaluated using a vacuum manifold or automated positive-pressure system. Peptides were quantified and normalized using the Pierce™ Quantitative Colorimetric Peptide Assay prior to 1:1 mixing and LC-MS analysis using a Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap™ mass spectrometer.

Preliminary Data

Although our optimized protocol using in our EasyPep kit significantly reduces both hands-on and total sample processing time, peptide clean up using microcentrifuge spin columns is still time consuming with larger sample numbers. We developed a new 96-well filter plate format compatible with TMTpro reagents to support higher sample processing throughput for multiplexed quantitation. Compared to the manual labeling and spin column protocol, this format showed nearly identical performance and quality with scalability and better reproducibility among replicates.

Conclusion

We demonstrate that our chemistry is readily adaptable to automated liquid handling system which provides excellent reproducibility and greatly simplifies proteomic sample preparation.

850

Proteomic approaches in biomarker panel to diagnose mental illness: focus on major depression disorder, bipolar disorder, schizophrenia

Areum Sohn[1, 2], Sangjin Lee[3], Yongmin Ahn[3], Youngsoo Kim[1, 2]

1. Department of Biomedical Sciences, Seoul National University, Seoul, South Korea

2. Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, South Korea

3. Department of Neuropsychiatry, Seoul National University Hospital, Seoul, South Korea

Psychiatric disorders, such as major depression disorder (MDD), bipolar disorder (BP), and schizophrenia (SZ) are severe mental illnesses associated with morbidity and life-long disability for sufferers. Abnormal behavior and disturbed cognition, often assumed to represent psychiatric disorder, may actually result from some form of abnormal brain disease that can be detected by means of one or more biomarkers. However, heterogeneity of psychiatric disorder illness course complicates clinical decision-making. In recent years, the search for psychiatry-relevant biomarkers of major depression, bipolar disease and schizophrenia has intensified. In this study, quantitative targeted proteomics was performed on psychiatric disorder patients using liquid chromatography-mass spectrometry. We used plasma samples (40 normal control, 50 MDD, 50 BP, 50 SZ) for construction of biomarker panels for differential diagnosis of MDD, BP and SZ. In additions, we compared proteins expression levels of psychiatric disorder patients with that of normal control. For multi-marker panel, AUCs based on the machine learning algorithm in BP vs. SZ showed the highest value with 0.990, 0.882 in training and test set, respectively. These results confirm that clinically useful differential diagnosis in MDD, BP, and SZ, also improve on the conventional methods in developing such models.

851

viQC: Visual And Intuitive Data Quality Control For Bottom-Up Proteomics

Elizaveta Solovyeva[2, 1], Anna Lobas[2], Lev Levitsky[2], Mikhail Gorshkov[2]

1. Moscow Institute of Physics and Technology (State University), Dolgoprudny, Moscow Region, Russia

2. V.L. Talrose Institute for Energy Problems of Chemical Physics, Russian Academy of Sciences, Moscow, Russia

Background

Proteome analysis based on mass spectrometry becomes an essential part of many biological studies. The increasing value of the analysis results elevates the importance of proper quality control (QC) over the instrument performance and operation that becomes a key step in the experimental workflow. A large variety of earlier developed QC tools are typically loaded with numerous metrics, require complex mixtures and rely on extensive data pre-processing. Further, the results are often tricky for interpretation, especially for early-career scientists. Yet, rapid and simple assessment of the instrument's readiness for the analysis is often all one needs in practice. In this work, we developed an approach and the software viQC (visual&intuitive quality control) based on a few recently demonstrated metrics for quick assessment of data quality in bottom-up proteomics.

Methods

Three datasets (one protein digest, whole-cell lysates, and PNNL QC dataset [1]) obtained from five different Orbitrap-based instruments were used for developed QC method evaluation. Further, we used the method for large scale quality analysis of datasets from PRIDE database.

Results

The proposed QC method was compared with unsupervised approach developed recently [2] using 57 whole cell lysate experimental runs and shown high specificity (100%) and selectivity (one-sided 95%-CI from 86.6% to 100%). Moreover, viQC algorithm can be effectively used even for a simple mixture, such as a single protein digest. By optimizing the instrument parameters following the results of data characterization by viQC, multifold improvements in the number of identifications was demonstrated.

Conclusion

The Python-based software viQC for fast, intuitive, and visual quality control of bottom-up proteomic experiments were proposed and developed. The software takes less than a minute on a regular PC for analysis of typical experimental runs. Accessing the metric is not requiring time-consuming data processing, spectra identification, and can be easily implemented for day-to-day practice.

852

Plasma proteome profiling to detect novel biomarkers for cholangiocarcinoma

Chantragan Srisomsap[1], Kamolwan Watcharatanyatip[1], Daranee Chokchaichamnankit[1], Somchai Chutipongtanate[3, 2], Churat Weeraphan[4], Jisnuson Svasti[5, 1]

1. Laboratory of Biochemistry, Chulabhorn Research Institue, Bangkok, Thailand

2. Department of Pediatrics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Bangkok 10400, Thailand

3. Department of Cancer Biology, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

4. Department of Molecular Biotechnology and Bioinformatics, Faculty of Science, Prince of Songkla University, Songkhla, Songkhla 90110, Thailand

5. Applied Biological Sciences Program, Chulabhorn Graduate Institute, Bangkok, Thailand

The high incidence of cholangiocarcinoma (CCA) is well known in the Northeastern region of Thailand. The improvement of diagnosis, treatment and biomarker discovery for CCA are still needed and should greatly improve treatment outcome for patients. At present, the tumor markers for CCA are CA19-9 and CEA, which are not specific. Our previous studies on the search for CCA biomarkers since 2004 until present have found the list of possible candidates, together with our recent technique, the mass spectrometry-based label-free quantitative proteomics, was employed for 27 plasma samples (9 normal individuals, 9 CCA and 9 disease controls). Four proteins (Protein 1, 2, 3 and 4) were selected for immunoblot verification using ELISA immunoassay to examine the diagnostic performances in a larger cohort (63 normal, 26 CCA and 37 non-CCA). Receiver Operating Characteristics for protein 1 gave higher expression with the area under the curve (AUC) of 0.835 (80.8% sensitivity, 83.8% specificity) when compared to other 3 proteins. Among the used of indexed models, a combination of protein 1, protein 2 and protein 3 increased the diagnostic performance with the AUC of 0.849 (76.9% sensitivity, 89.2% specificity). Our result for the combination of 3 proteins showed good promise as a potential multiplexing biomarker for CCA, and the further validation should be explored in an independent cohort.

Supported by the Chulabhorn Research Institute

853

Proteogenomic study of vitronectin in breast cancer

Alakesh Bera[1], Madhan Subramanian[1], Michael Eklund[1], John Karaian[1], Harvey B Pollard[1], Hai Hu[2], Craig D Shriver[3], Meera Srivastava[1]

1. USU School of Medicine, Bethesda, MARYLAND, United States

2. Chan Soon-Shiong Institute of Molecular Medicine, Windber, Philadelphia, United states

3. Murtha Cancer Center, Uniformed Services University/Walter Reed National Military Medical Center, Bethesda, Maryland, United States

Breast Cancer is the most common form of cancer in women worldwide, impacting nearly 2.1 million women each year. In 2018, nearly 627,000 women died of breast cancer, which accounts for 15% of all cancer deaths. Identification of new biomarkers could be key for early diagnosis and therefore, increase the survival. Vitronectin, a glycoprotein that is abundantly found in serum, extracellular matrix, and bone, binds to integrin alpha V beta 3, and promotes cell adhesion and migration. Current studies indicate that patients with amplified Vitronectin copy-number have lower survival rates than patients without amplified Vitronectin. In this study, we focused on the role of vitronectin in breast cancer survival and its functional role as a non-invasive biomarker for early stage and stage specific breast cancer detection. To confirm that the expression of vitronectin is amplified in breast cancer, a total of 240 serum samples (n=240) from breast cancer patients were analyzed using the Reverse Phase Protein Array (RPPA) technique. Of the 240 samples, 120 samples were of African American descent, while the other 120 were of Caucasian American descent. Initial analysis of this data revealed that there were significant racial disparities in vitronectin expression level, specifically seen in the recurrent samples. Next, we tried to uncover the underling mechanism which plays critical role in Vitronectin expression. We analyzed the different protein markers in four breast cancer cell lines MCF7, MB231, MB468, and HCC1599. The data indicated that the PI3K/AKT axis is modulating the expression of vitronectin and associated survival rate in breast cancer.

854

Digging deeper in every direction: developing a next-generation platform for high-throughput multi-faceted protein characterization

AARON O BAILEY[1], JIMMY Y ZENG[1], GUANGHUI HAN[1], JOSHUA SILVEIRA[2], KRISTINA SRZENTIC[3], CHRISTOPHER MULLEN[2], JOHN E.P. SYKA[2], ROMAIN HUGUET[2]

1. BGI Americas, San Jose, California, United States

2. Thermofisher scientific, San Jose, California, United States

3. Thermo Fisher scientific, Cambridge, MASSACHUSETTS, United States

So-called “comprehensive” LC-MS-based characterization of therapeutic proteins requires attaining a detailed molecular knowledge of every residue and sidechain.We have designed a cutting-edge platform for addressing the versatility and throughput requirements of performing comprehensive characterization.We have developed advanced versions of multiple key characterization workflows using state-of-the-art LC-MS methods performed on a single Orbitrap mass spectrometer. We used the Fc-fusion protein etanercept to optimize intact mass and peptide mapping strategies for characterizing highly glycosylated therapeutic proteins.Specifically, we show that native SEC-MS intact mass strategy is significantly improved with the combined use of proton transfer charge reduction (PTCR) to provide additional separation of the broad isoform profiles at sequential charge state in the complex intact mass spectra.For peptide mapping, etanercept was digested using trypsin and AspN. Intact and IdeS-digested Fc and TNFR subunits were deglycosylated with combinations of PNGase F, O-glycosidase, and sialidase and analyzed by native SEC-MS followed by PTCR. LC-MS was accomplished using a Vanquish UHPLC connected to a Thermo Scientific™ Orbitrap Eclipse Tribrid mass spectrometer equipped with PTCR and extended mass range detection and isolation. Our studies showed highly complex isoform distributions consistent with previously published reports. We performed peptide mapping using data dependent acquisition (DDA) to fragment selected peptide ions by HCD and ETD. This effort resulted in 100% coverage of the etanercept amino acid sequence, including 3 N-glycan sites and 13 O-glycosylation sites. We analyzed etanercept in intact and subunit form after combined deglycosylation treatments and determined the main sources of glycoform heterogeneity for each preparation. For highly complex preparations we show that high mass isolation and PTCR are powerful tools to further separate sequential charge states.This form of charge reduction intact mass analysis enhanced our ability to accurately assess glycoform identity directly from a highly complex isoform background and can be implemented in a high throughput fashion.

855

An extracellular proteomics approach to understand the wheat phosphate starvation response at the plant-soil interface

Christiana Staudinger[1, 2], Bhagya Dissanayake[1], Owen Duncan[1], A Harvey Millar[1]

1. ARC Centre of Excellence in Plant Energy Biology, UWA, Crawley, WA, Australia

2. School of Biological Sciences, UWA, Crawley, WA, Australia

Extracellular proteins released by plant roots contribute to nutrient mobilisation from organic matter, pathogen defence and localised root growth adjustment to the surrounding environment. Thereby, they have the potential to remodel the extracellular matrix and play a pivotal role in plant nutrition and survival. However, little is known about the composition, abundance and functional potential of the extracellular proteome derived from plant roots at the plant-soil interface under nutrient stress conditions. In this project, we apply both untargeted and targeted proteomics approaches to characterise the localised response of wheat root tips and the associated extracellular proteome to low phosphate conditions. By using untargeted LC-MS/MS techniques, we first established a comprehensive map of the proteins released by wheat roots. More than 300 proteins were identified within the extracellular space. Relative to the root tip proteome, the proteome present in this environment was significantly enriched in proteins with predicted secretion tags and apoplastic localisation. The majority of proteins had a functional annotation related to stress and pathogen response, protein processing and cell wall biosynthesis. Second, we developed multiple reaction monitoring (MRM) assays for the detection of root-released phosphatases and other phosphate starvation responsive proteins within the extracellular space. Ultimately, the delineation of the wheat root tip and extracellular proteome responses to phosphorus starvation by targeted proteomics will deepen our understanding of spatial and time resolved phosphorus dynamics at the plant-soil interface.

856

Proteomic mapping of chemical warfare agent exposed plasma

Joel R Steele[1, 2], Craig S Brinkworth[2], Steven Torney[2], Tristan Kilmartin[2], Harry Rose[2], Rabih E Jabbour [3]

1. University of Technology, Sydney, East Hills, NSW, Australia

2. Land Division, Defence Science and Technology Group, Melbourne , VIC, Australia

3. CCDC-Chemical Biological Center, CBC, Aberdeen Proving Ground, MD, United States of America

Background:

Rapid detection and accurate identification of toxic chemical exposure are of paramount importance to various industries and government sectors. The use of analytical techniques to identify chemical warfare agents (CWAs), G-series, HD and VX in a known exposed area is commonplace, but there are currently no efficient analytical techniques available to specifically identify human exposure. The recent use of CWAs against civilians in war-torn countries stresses the need to develop accurate methods of detection and identification of CWAs among civilians and military personnel. This study aimed to establish a proteomic map of human plasma exposed to different CWAs and to identify potential biomarkers that could be used for the detection of specific CWA exposure.

Methods:

Whole plasma was exposed to individual CWAs and analysed via three distinct methods. Samples were subjected to liquid trypsin digestion and shotgun LC-MS/MS on two mass spectrometers, additionally utilising gas phase fractionation for enhanced proteome coverage. Exposed plasma was additionally analysed via 1D SDS-PAGE and two-dimensional difference gel electrophoresis (2D-DIGE) coupled with in-gel trypsin digestion and LC-MS/MS. Data analyses were performed using PEAKS Studio X bioinformatics software.

Results

Analysis revealed several potential biomarkers that can differentiate exposed versus unexposed plasma for various CWAs. Analysis of human serum albumin revealed sites of O-isomethylphosphonylation, O-pinacolylmethylphosphonylation and various phosphonylations sites indicative of chemical agent activity. Phosphorylation sites were removed in CWAs exposed plasma samples. The highest number of observed post-translational modifications occurred in plasma exposed to mustard gas.

Concluding Statement

This work has revealed unique and novel modifications sites in human plasma exposed to CWAs. These sites are potential biomarkers that hold promise for enabling future rapid and accurate identification of human CWA exposure.

857

Mapping hydroxylated tyrosine in the human brain proteome: The formation and incorporation of L-DOPA

Joel R Steele[1], Jake P Violi[1], Lisa Pu[1], Matt P Padula[1], Kenneth J Rodgers[1]

1. School of Life Sciences, University of Technology Sydney, Ultimo, NSW, Australia

Background

Parkinson's disease (PD) is a neurodegenerative disorder characterised by the loss of dopamine-producing neurons in the midbrain. Oxidative stress (OS) is thought to contribute to the initiation and progression of PD. Levodopa (L-DOPA), the primary treatment for PD, can increase OS and additionally be mistakenly incorporated into proteins. Conversion of L-tyrosine to L-DOPA can, therefore, occur from both hydroxyl radical attack on tyrosine residues and replacement of L-tyrosine with L-DOPA during protein synthesis. This study aimed to explore the presence of L-DOPA as a human protein constituent in addition to creating a synthetically modified proteome containing L-DOPA.

Methods

Tyrosinase was used to convert tyrosine residues in tryptic peptides of human neuroblastoma cell (SH-SY5Y) proteins to DOPA, creating a positive control for DOPA incorporation. Cells were also treated with L-DOPA to allow *in vitro* incorporation into proteins. The brain subset of the draft human proteome was reanalysed for sites of hydroxylated tyrosine as a negative control and compared to two publicly available Parkinson's disease datasets to quantitatively analyse the undocumented presence of this modification. PEAKS Studio X® was utilised for all data analyses.

Results

The tyrosinase reaction achieved a 10% conversion of detected tyrosine residues to L-DOPA. Analysis of L-DOPA treated cell cultures showed increased numbers of L-DOPA sites versus non-treated controls. Numerous sites of hydroxylated tyrosine and other residues in the draft human brain proteome were identified. It was noted that the substantia nigra contains a higher number of hydroxylated sites, even amongst controls.

Conclusion

This study has mapped tyrosine hydroxylation sites within the human brain proteome and generated a method for the specific conversion of tyrosine to L-DOPA. This work has additionally created a novel library of L-DOPA containing peptides for assaying clinical samples of Parkinson's disease and exploring hydroxylation in other human diseases and tissues.

858

The neurotoxin β-Methylamino-L-alanine and it's incorporation into proteins

Joel R Steele[1], Jake P Violi[1], Matt P Padula[1]

1. School of Life Sciences, University of Technology Sydney, Ultimo, NSW, Australia

Background

Approximately 10% of neurodegenerative diseases can be attributed to genetic disorders, while the remaining 90% are sporadic, suggesting that environmental factors may play a role in the iniation. An environmental neurotoxin β-Methylamino-L-alanine (BMAA) has been identified to potentially play a role in the formation of sporadic neurodegenerative diseases. Detection of BMAA in diseased brains has led to a hypothesis of BMAA incorporation into proteins potentially resulting in protein aggregation. This study aimed to explore methodologies that could identify whether BMAA is present within the detectable proteome.

Methods

Several neuroblastoma SH-SY5Y cell cultures were treated with a range of BMAA concentrations. Unfractionated and fractionated digests of treated SH-SY5Y cells were analysed by both data dependent and independent LC-MS/MS. Bioinformatic analyses were performed utilising PEAKS studio X.

Results

No incorporation of BMAA was detected with confidence in place of any amino acid. These results indicate that further research on the incorporation of BMAA into proteins is not feasible with traditional proteomic methodologies

Conclusion

Data collected in this study suggest that BMAA may not be incorporated into proteins by traditional protein synthesis but instead might be bound to proteins in another manner. Furthermore, this research indicates that if BMAA is incorporated into proteins, then detection of BMAA within proteins is not feasible with short term culture and simple proteomic methodologies.

859

Developing a robust proteomics workflow to unlock archival FFPE colorectal cancer tissue cohorts for biomarker discovery

Pascal Steffen[1], Mark Molloy[1]

1. The University of Sydney, St Leonards, NSW, Australia

Background

To extract proteins from formalin-fixed paraffin embedded (FFPE) tissue most protocols use high concentrations of SDS with heating for >1 hour to 100 °C. Two widely employed methods (SP3 bead capture and FASP) are able to mostly remove SDS, although we found that even minute amounts of remaining SDS alter chromatographic retention times. This is detrimental for run-run reproducibility, especially for data-independent acquisition (DIA) and could compromise quantitation from multi-batch cohort studies. Here, we developed an improved sample preparation workflow and DIA strategy aimed at supporting robust quantitative analysis of colorectal cancer (CRC) FFPE specimens.

Methodologies

CRC FFPE sections (5x5 µm per sample) were macro-dissected, lysed and proteins reduced and alkylated. Proteins were digested and purified according to a reported SP3 protocol (PMID: 29565595) with subsequent SDS precipitation using KCl. A pooled sample was used to acquire 6 small window DIA gas-phase fractions for chromatogram library generation. Wide window DIA data were acquired using an overlapping window scheme. All data were acquired on a Thermofisher QExactive HFX mass spectrometer with 140min LC time using a 50cm x 75µm self-packed pulled column.

Results

Using this method we observed highly reproducible chromatography from FFPE extracted samples. We could measure all FFPE samples reproducibly and with minimal RT shifts from the library (median deltaRT: 1.30 ±0.16 min). The gas-phase fractionation library consisted of over 6000 proteins and when applied to the EncyclopeDIA workflow (PMID: 30510204) allowed for the Identification and quantification of >4000 proteins in each of the 12 test specimens from ~1ug on column load.

Conclusion

Combining SP3 bead clean-up with a subsequent precipitation of residual SDS led to highly reproducible RT across FFPE extracted samples. Using the EncyclopeDIA workflow we were able to acquire chromatogram libraries with similar depth to that of high pH fractionation with less sample handling.

86

Integrative multi-omics analysis from minimally invasive colorectal cancer FFPE tissue

Mark P Molloy[1], Pascal Steffen[1], Jun Li[1], Mahsa Seyed Ahadi[2], Anthony Gill[2], Alexander Engel[3]

1. Bowel Cancer and Biomarker Laboratory, The University of Sydney, St Leonards, NSW, Australia

2. Dept Anatomical Pathology, Royal North Shore Hospital, St Leonards

3. Dept Colorectal Surgery, Royal North Shore Hospital, St. Leonards

Background: Clinical tumour specimens are routinely preserved as formalin-fixed paraffin embedded (FFPE) tissues for histopathological analysis. These specimens are richly annotated with clinicopathological and prognostic information. However, molecular analysis of early stage FFPE cancer specimens is far more challenging than analysis using fresh tissue, or late staged FFPE specimens which tend to be larger. We have investigated integrative omics analysis from Stage I and IIIA colorectal cancer FFPE specimens by proteomic mass spectrometry, Nanostring gene expression analysis, DNA sequencing and CyToF immune cell profiling.

Methods: For proteomics 5x5 µm sections were macro-dissected on slide, processed using trypsin and SP3 purification. An EncyclopeDIA narrow window gas-phase fractionated chromatogram library empirically corrected with PROSIT was produced. Each sample was acquired via overlapping wide window DIA. All data were acquired on a QExactive HFX mass spectrometer over 140min using a 50cm x 75µm column. For Nanostring and DNA sequencing, mRNA and DNA was extracted from freshly cut 10x5 µm sections and quality checked by Bioanalyzer. Samples were run on Cancer Progression kit profiling 770 genes. We explored whole exome sequencing and targeted panel sequencing. For imaging mass cytometry a Hyperion CyToF system was used with 40 CD antigen markers on each 7 µm section.

Results: The extensive peptide chromatographic library enabled quantitative measurement of ~5300 proteins across all 12 samples. Statistical analysis identified proteins with roles in metastasis and cell migration when we dichotomised specimens based on lymph node positivity. Nanostring gene expression analysis identified factors absent from the proteomics analysis and revealed the importance of AP-1 TF activation in promoting metastasis. WES produced poor data due to fragmented DNA. Imaging mass cytometry was successfully implemented from archival sections and is under review.

Conclusion: Learnings for integrative omics analysis of minimally invasive FFPE CRC tumours has been established to underpin larger cohort studies.

860

Aging in testis: proteome alterations in senescent testicular peritubular cells

Jan Bernd Stöckl[1], Thomas Fröhlich[1], Nina Schmid[2], Florian Flenkenthaler[1], Rüdiger Behr[3], Artur Mayerhofer[2], Georg J. Arnold[1]

1. Gene Center Munich, LMU München, München, BAYERN, Germany

2. Biomedical Center (BMC), LMU München, München, Bayern, Deutschland

3. Platform Degenerative Diseases, German Primate Center, Göttingen, Germany

Aging of the human testis, conjointly with a decline in reproductive function, is still underexplored and the fundamental molecular mechanisms are poorly understood. Testicular peritubular cells (TPCs) are smooth muscle-like cells, building the wall of seminiferous tubules in men. TPCs transport sperm through the seminiferous tubules and are supposed to be important for spermatogenesis by secretion of proteins to the spermatogonial stem cell (SSC) niche. Therefore, senescence of TPC may also contribute to the age-related decline of testicular functions in men. Due to the limited availability of standardized human samples, we established marmoset (*C. jacchus*) TPCs (MKTPC) as a model for human TPCs (HTPC). To investigate the degree of similarity between HTPCs and MKTPCs, we performed a LC-MS/MS analysis of TPCs from both species and could find a high degree of sequence homology and a considerable correlation of protein abundances between both species. The similarity of their proteomes introduces MKTPCs as a compelling model for the human system. To trigger senescence of MKTPCs, cells of young individuals (n = 5) were repeatedly passaged. After approximately 10 passages, cells stopped dividing and showed hallmarks of cellular senescence. Proteomes and secretomes of passaged and control MKTPCs were then compared using a label-free approach. Specific proteome alterations in passaged MKTPCs included an increased abundance of mitochondrial and a decreased abundance of smooth muscle cell proteins. Furthermore, proteins of the endoplasmic reticulum and proteins related to mRNA regulatory processes were found to be decreased in passaged MKTPCs. These findings suggest that cellular senescence may impair sperm transport through a reduction of contractility. Additionally, the detected alterations in the secretomes of aged MKTPCs imply changes in the protein composition of the SSC niche.

1. Mayerhofer, A. (2013). Human testicular peritubular cells: more than meets the eye. Reproduction, 145(5), R107-116. doi:10.1530/REP-12-0497
2. Schmid, N., Stockl, J. B., Flenkenthaler, F., Dietrich, K. G., Schwarzer, J. U., Kohn, F. M., . . . Mayerhofer, A. (2018). Characterization of a nonhuman primate model for the study of testicular peritubular cells - comparison with human testicular cells. Mol Hum Reprod. doi:10.1093/molehr/gay025

861

Real-time, high density monitoring of pTyr signaling targets in human tumors using SureQuant heavy peptide triggered targeted quantitation

Lauren E Stopfer[1, 2], Cameron T Flower[1, 2], Aaron S Gajadhar[3], Bhavin Patel[4], Sebastien Gallien[5], Romain Huguet[3], Graeme McAlister[3], Derek Bailey[3], Shannon Eliuk[3], Markus Kellmann[6], Tabiwang N Arrey[6], Alexander Harder[6], Daniel Lopez-Ferrer[3], Andreas Huhmer[3], Forest M White[1, 2]

1. Massachusetts Institute of Technology, Cambridge, MA, United States

2. Koch Institute for Integrative Cancer Research, Cambridge, MA, United States

3. Thermo Fisher Scientific, San Jose, CA, United States

4. Thermo Fisher Scientific, Rockford, IL, United States

5. Precision Medicine Science Center, Thermo Fisher Scientific, Cambridge, MA, United States

6. Thermo Fisher Scientific, Bremen, Germany

Tyrosine phosphorylation (pTyr) plays a pivotal role in signal transduction and is commonly dysregulated in cancer. Profiling the tumor pTyr proteome may reveal therapeutic insights critical to combating disease. Existing discovery-based mass spectrometry methods to monitor pTyr networks favor broad coverage while sacrificing reproducibility, while targeted methods attempt to address reproducibility but are focused on a limited subset of sites in the entire pTyr network. To achieve high coverage, high reproducibility analysis of the network, we applied a novel, high-density global targeted approach that leverages isotopically-labeled trigger peptides to reliably quantify several hundred commonly dysregulated pTyr targets in a cohort of 30 human colorectal tumor samples.

Tumor specimens were lysed, and proteins were reduced, alkylated, and digested to peptides. A mixture of stable isotopically-labeled (SIL) peptides, corresponding to several hundred pTyr targets, was spiked into each tryptic tumor digest. The endogenous and SIL forms of pTyr-containing peptides were isolated with a 2-step immunoprecipitation and IMAC enrichment, and LC-MS analysis was performed with IonOpticks Aurora column (250mm x 0.075mm) and an EASY-nLC™ 1200 coupled to an Orbitrap Exploris™ 480 mass spectrometer. To ensure reproducible measurement of selected pTyr targets, a SureQuant™method adapted from the internal standard triggered parallel reaction monitoring method (IS-PRM) was performed where real-time heavy peptide detection triggered selective and sensitive measurements of endogenous pTyr targets. Data analysis was performed using Proteome Discoverer and Skyline software.

SIL-triggered targeted pTyr analysis quantified over 300 unique pTyr across patient tumor samples, revealing quantitatively distinct proteomic signatures. In some cases, pTyr profiles align with proteomic & transcriptomic molecular subtypes previously reported. However, pTyr profiling also revealed putative patient specific oncogenic driving mechanisms not captured in global proteomics. Heavy peptide triggered and guided acquisition maximizes the efficiency of targeted quantification by enhancing the detectability of targets, significantly improving measurement reproducibility across analysis.

862

Development of a fully automated magnetic workflow for phosphoproteome profiling

Stoyan Stoychev[2, 1], Previn Naicker[2], Ireshyn Govender[2], Sindisiwe Buthelezi[2], Sipho Mamputha[2], Isak Gerber[2], Jusitn Jordaan[1]

1. Proteomics, Resyn Biosciences, Pretoria, Gauteng, South Africa

2. CSIR, Pretoria, GAUTENG, South Africa

The research aims to implement a seamless fully automated workflow for phosphoproteome profiling, covering all steps from extracted protein to the mass spectrometry analysis. Automation of this workflow increases throughput and reproducibility. Each component of the pipeline was evaluated individually, and subsequently integrated to form the seamless workflow. We started with protein isolation from a complex lysate using different strategies, employing three bead chemistries for solid phase extraction, strong anion exchange (SAX), hydrophilic affinity (HILIC), and amine for on-bead protein aggregation capture (PAC/SP3). Based on 200ug starting material the three approaches exhibited similar recoveries. However, their performance varies based on sample source, amount of protein, concentration, extraction conditions, and additives. We further plan to assess the ability to perform peptide fractionation using SAX, potentially allowing for deeper phosphoproteome mining. For phosphopeptide enrichment we evaluate the complementarity of Ti-IMAC, Zr-IMAC and TiO2. We provide alternate buffering conditions that can result in increased phosphopeptide recovery. The potential complementary of the phosphopeptide enrichment chemistries is explored, where various bead combinations are tested in an attempt to provide more comprehensive phoshopeptide coverage in a single enrichment. All of the above-mentioned methods including protein capture, clean-up and on-bead digestion as well as phosphopeptide enrichment have been adapted for use in a magnetic handling stations (KingFisher Duo) allowing for fully automated sample processing. The methods are readily transferable to a range of liquid handling robots fitted with magnetic handling platforms.

863

Loss of protein *N*-glycosylation influences peptidoglycan structure in *Campylobacter jejuni*

Zeynep Sumer-Bayraktar[1], Joel A Cain[1], William Klare[1], Paula Niewold[1], Lok Man[1], Stuart J Cordwell[1, 2]

1. School of Life and Environmental Sciences, Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia

2. Sydney Mass Spectrometry, The University of Sydney, Sydney

Campylobacter jejuni is a Gram-negative microaerophilic bacterium that is the leading cause of food-borne gastroenteritis in developed countries. The common route of human infection is *via* consumption of poorly prepared or under-cooked poultry, in which *C. jejuni* is an asymptomatic commensal. Human disease is characterised by mild to severe diarrhoea, vomiting and inflammation. The molecular basis for *C. jejuni* infection includes initial adherence to, followed by invasion of, human intestinal epithelium. Bacterial peptidoglycan (PG) is an integral component of the cell wall/membrane that is involved in pathogen colonization, host-pathogen interactions, cell structure and morphology. Disruption of PG integrity results in cell lysis making it the major target for existing and novel antibiotics. *C. jejuni* contains a protein *N*-glycosylation gene (*pgl*) cluster that modifies membrane-associated proteins and deletion of genes in this cluster results in reduced pathogenesis. Proteome analysis revealed several glycoproteins in the PG biosynthesis pathway were impacted by oligosaccharyltransferase (*pglB*) gene deletion. We therefore investigated whether glycosylation influenced the peptidoglycome. Wild type and *pglB* mutant strain PG structures were isolated and analysed using LC-ESI tandem mass spectrometry. The structural features and relative abundance changes of peptidoglycans in each strain were determined. In addition to the already known peptidoglycans, we discovered several unique structural modifications on the *N*-acetylneuraminic acid residues and showed that these changes are modulated differently in *pglB* mutant strains compared to wild-type. The observed variations in the PG structure impacted lysozyme resistance, supporting the role of wild-type PG composition for *C. jejuni* host survival. In summary, our results show that changes in protein glycosylation impact the structural composition of the PG layer commensurate with changes to PG-associated proteins observed at the proteome level. Unravelling PG structures and the factors involved in their biosynthesis may help in defining why *N*-glycosylation is critically required for *C. jejuni* pathogenesis.

864

Discovery and Multi-center Verification of Prostate Cancer Protein Biomarkers using Single-shot Short Gradient Microflow Scanning SWATH and MRM^{HR} Mass Spectrometry

Rui Sun[1], Christie Hunter[2, 3, 4], Chen Chen[5], Weigang Wei[1], Nick Morrice[2], Qiushi Zhang[1], Xue Cai[1], Bo Wang[6], Xiaoyan Yu[7], Xiaodong Teng[6], Lirong Chen[7], Shaozheng Dai[8], Jian Song[1], Zhongzhi Luan[8], Changbin Yu[1], Rudolf Aebersold[9], Yi Zhu[1], Tiannan Guo[1]

1. Westlake University, HANGZHOU, ZHEJIANG, China

2. SCIEX, Redwood city, USA

3. SCIEX, Redwood city, USA

4. SCIEX, Redwood city, USA

5. SCIEX, Shang Hai, China

6. Department of Pathology, The First Affiliated Hospital of College of Medicine, Zhejiang University, HANGZHOU, ZHEJIANG, 中国

7. Department of Pathology, The Second Affiliated Hospital, Zhejiang University School of Medicine, HANGZHOU, ZHEJIANG, 中国

8. School of Computer Science and Engineering, Beihang University, Beijing, China

9. Department of Biology, Institute of Molecular Systems Biology, ETH, Zurich, Switzerland

Publish consent withheld

1. 1. Christoph Messner, V.D., Nic Bloomfield, Gordana Ivosev, Fras Wasim,Aleksej Zelezniak, Kathryn Lilley, Steven Tate and Markus Ralser. ScanningSWATH enables ultra-fast proteomics using high-flow chromatography and minute-scale gradients. (2019).

865

Statistical Characterization of Peptide Fragmentation Behaviors in the Negative Ion Mode Using High-Resolution Mass Spectrometry

Rui-Xiang Sun[1], Mei-Qing Zuo[1], Meng-Qiu Dong[1]

1. National Institute of Biological Sciences, Beijing, China

High-throughput protein identification in most proteomic researches relies primarily on mass spectrometry (MS) analysis of the large-scale digested peptides, which has routinely been conducted in the positive ion mode on protonated peptides. Nevertheless, there is still a subset of proteins or peptides rich in acidic amino acids which are underrepresented in current popular strategy. Negative-ion MS has provided a potential to achieve a full proteome coverage largely due to the complementarity to its positive-ion partner. However, to date the available knowledge of the fragmentation behaviors of deprotonated peptides is insufficient towards to an automatic interpretation as that done for protonated peptides in bottom-up proteomics. Here, we statistically characterized the fragmentation patterns of 36 synthesized peptides, containing 4-16 amino acids, in the negative ion mode using CID and HCD under a variety of MS conditions, including a series of normalized collision energies (NCE). In light of the statistical analyses on all MS/MS spectra acquired in our study, the relationships between peptide fragmentation efficiency and its NCE were depicted for the first time. To facilitate communication, we modified the nomenclature of fragment ions in the negative ion mode in reference to that in the positive ion mode. Our studies find that *y*-, *c*- and *z*-type ions are the most dominant species in both CID and HCD spectra of deprotonated peptides, accompanied by abundant neutral loss peaks. Furthermore, HCD is generally better than CID for peptide sequencing in the negative ion mode, since HCD generates more backbone cleavage products whereas CID produces more side-chain neutral loss peaks. For disulfide-bonded peptides and C-terminally amidated peptides, specific fragmentation patterns are observed and analyzed. These statistically significant fragmentation behaviors of deprotonated peptides reported in our study will promote further investigation of the fundamental mechanisms and facilitate algorithmic development for peptide sequencing in the negative mode.

866

SimPhospho 2: phosphorylation site validation in multiply phosphorylated peptides using simulated spectral libraries

Veronika Suni[1, 2], Susumu Imanishi[3], Pekka Haapaniemi[1], Laura Elo[1], Garry Corthals[4]

1. Turku Bioscience Centre, University of Turku and Åbo Akademi University, Turku, Finland

2. TUCS Turku Centre for Computer Science, Turku, Finland

3. Faculty of Pharmacy, Meijo University, Nagoya, Japan

4. Faculty of Science, Van 't Hoff Institute for Molecular Sciences, Amsterdam, The Netherlands

Background

Site-specific identification of phosphorylation is one of the main computational challenges in phosphoproteomics. Earlier we proposed a method based on simulated phosphopeptide spectral libraries [1] that together with our software, SimPhospho [2], facilitates highly sensitive and accurate phosphosite assignments. Here, we will investigate how well multiply phosphorylated peptides can be simulated.

Methods

The principle of the simulation algorithm is as follows. Based on the spectra of dephosphorylated or nonmodified peptides, we predict spectra of phosphorylated peptides using the information about present ion peaks and their intensities as reference. Different types of ions, as well as their 2+, 3+ charged ions, and neutral losses are taken into account. Simulated spectra are collected in spectral libraries and used for identification of MS/MS spectra of phosphopeptide samples using SpectraST. For method development, hundred phosphopeptides (singly-, doubly- and triply-phosphorylated) were selected to be synthesized. These peptides were divided into three pools making sure that the phosphopeptide isoforms were separated. Analysis was done by Orbitrap HCD in Thermo Scientific Q Exactive mass spectrometer.

Results and discussion

The earlier version of SimPhospho was biased towards detection of singly phosphorylated peptides. This limitation is addressed by implementing simulation of doubly- and triply-phosphorylated peptides in SimPhospho 2. We will compare various combinations of simulation parameters in terms of False Localization Rate of phosphorylation sites we are able to obtain, when performing spectral library search of samples with known phosphorylation sites.

1. Suni et al. Confident site localization using a simulated phosphopeptide spectral library. 2015 Journal of Proteome Research
2. Suni et al. SimPhospho: a software tool enabling confident phosphosite assignment. 2018 Bioinformatics

867

DIAtools: Metaproteomics by data-independent acquisition mass spectrometry

Tomi Suomi[1], Sami Pietilä[1], Laura L. Elo[1]

1. University of Turku, Turku, Finland

For metaproteomic analyses in microbiome research, a commonly used technique is the data-dependent acquisition (DDA) mass spectrometry. A key limitation of this method is that multiple repeated runs are required for each sample to obtain a stable set of detected peptides. Technically that is mainly because of undersampling and that the MS/MS spectra are taken outside the elution peak. More recently data-independent acquisition (DIA) methods have been successfully proposed to overcome some of the limitations, but there remains a lack of free and easy-to-use software for the analysis of DIA mass spectrometry proteomics. This is especially true in the context of metaproteomics, which proposes additional requirements for the software. To this end, we have developed a free and open-source software package for analyzing complex DIA metaproteomic data. It enables accurate and consistent quantification and we demonstrate its feasibility in gut microbiota metaproteomics using laboratory assembled microbial mixtures as well as human fecal samples.

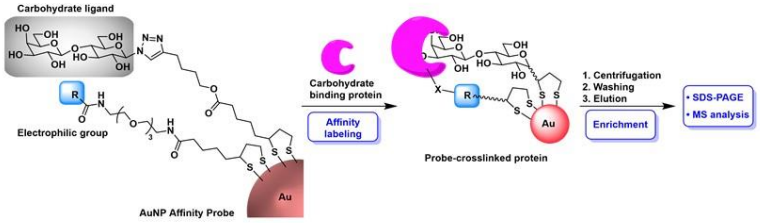
868

Development of gold nanoparticle-based affinity labeling probes for identification of carbohydrate-binding proteins

Nanako Suto[1], Kaori Sakurai[1]

1. Tokyo University of Agriculture and Technology, Graduate School of Engineering, Koganeishi, TOKYO, Japan

Carbohydrate-protein interactions mediate various important cellular signaling. However, it is difficult to explore novel carbohydrate binding proteins because carbohydrate ligands often display low affinity. Photoaffinity labeling provides a promising strategy for target identification because of its ability to capture proteins in situ by covalent bonds. Despite the potential utility, low labeling efficiency of photoaffinity labeling and the need for the subsequent enrichment process hamper its routine use. To address this issue, we have developed gold nanoparticle (AuNP) -based photoaffinity probes that multivalently display carbohydrate ligands and photoreactive groups at high density. AuNPs enable rapid optimization of the probe design due to the ease of preparing the AuNP-based probes and easy protein purification by centrifugation. It facilitated purification and dramatically improved the affinity and labeling efficiency. Nevertheless, more efficient protein labeling is desired for target identification. Various electrophilic groups are known as protein labeling reagent, however, there are few examples used for affinity labeling because of control of selectivity. In this presentation, we show design and synthesis of new AuNP-based affinity labeling probes bearing lactose, a model carbohydrate ligand and various electrophilic groups for efficient capture of carbohydrate binding proteins. We also prepared control probe presenting lactose only and evaluate its binding affinity to a lactose binding lectin PNA, to find K_d is 16 nM, which verified successful affinity enhancement by multivalent effect. We then conducted screening of the new probes for labeling efficiency of known lectins, PNA, ECA and RCA. Our results suggested that a judicious selection of a protein reactive group enables efficient affinity labeling of target lectins.



869

Role of Cytoskeleton Proteins in Trans-Endothelial Permeability: An *In Vitro* Model for Endothelial Dysfunction in Dengue Virus Infection

Aroonroong Suttitheptumrong[1], Sa-nga Pattanakitsakul[1]

1. Mahidol University Faculty of Medicine Siriraj Hospital, Thailand, Bangkoknoi, BANGKOK, Thailand

Dengue virus (DENV) infection is one of the global arboviral epidemic; most common in the South East Asia. The severity of DENV infection ranges from dengue fever (DF), dengue hemorrhagic fever (DHF) to the most severe dengue shock syndrome (DSS). Plasma leakage and vascular permeability was evident in DHF/DSS, with marked symptoms of endothelial dysfunction. During DENV infection, cytokine and chemokine responses were reported to regulate the functional cytoskeleton protein networking, however, the exact role of individual cytoskeleton proteins is still unclear. The pro-inflammatory cytokine, Tumor Necrosis Factor-a (TNF-a) which was markedly elevated in patients with DHF/DSS, we recently identified the synergistic effect of DENV infection with TNF-a treatment causing the trans-endothelial permeability in human endothelial cells (EA.hy926). In the present study, our aim was to characterize the role of cytoskeleton proteins, in DENV-infected human endothelial cells with TNF-a activation. Mass spectrometry analysis (GeLC-MS/MS) was used to characterize the cytoskeleton proteomes in the DENV-mediated endothelial dysfunction. A total of 450 altered cytoskeleton proteins was found; where moesin was the most remarkably reduced cytoskeleton protein. Western blot analysis and immunofluorescence staining was used to further characterize the functional role of moesin in DENV-induced endothelial dysfunction. Further studies identified the structural reorganization of other two cytoskeleton proteins, including F-actin and vinculin in DENV-induced endothelial dysfunction. Trans-endothelial electrical resistant (TEER) assay was applied to measure the vascular permeability in the DENV-infected endothelial cells with TNF-a induction; and, the results were correlated with the expressions of moesin, F-actin and vinculin. Our results suggest the reduced expression of moesin regulate the endothelial permeability in severe forms of DENV infection. In conclusion, we identified the pivotal role of cytoskeleton protein in an *in vitro* model of DENV-induced endothelial dysfunction; this would paves in identifying therapeutic strategies towards the severe forms of DENV infection.

87

Understanding liver regeneration. Proteomics and phosphoproteomics analysis

Lorena Carmona[1], Matías Avila[3, 2], Antonio Ramos[4], Rosana Navajas[1], Fernando Corrales[1]

1. National Centre of Biotechnology, CSIC, Madrid, MADRID, Spain

2. Centro de Investigación Biomédica en Red en el área temática de Enfermedades Hepáticas y Digestivas (CIBERehd), Pamplona, Spain

3. Division of Hepatology and Gene Therapy, Center for Applied Medical Research, University of Navarra, Pamplona, Spain

4. Proteobotics SL, Darwin 3, Madrid, Spain

The liver is able to orchestrate a regenerative process to compensate the loss of functional tissue. Resection of 70% of liver parenchyma (partial hepatectomy, PH) in mice is a well stablished model to analyze the mechanisms underlying liver regeneration. Three phases can be defined in liver regeneration: sensitization of hepatocytes to mitogenic factors, proliferative response and regeneration termination. To dissect in detail liver regeneration mechanisms after PH we have generated a dynamic framework integrating the changes in protein abundance and phosphorylation along the three phases of the process. Proteomic and phosphoproteomics analysis in the starting regeneration phase shows a total of 4300 proteins identified, of which 2800 quantified being 219 differentially regulated in the PH / sham contrast (215 up-regulated and 4 down-regulated, qvalue <0.05) with 24 peptides differentially phosphorylated (21 up-regulated and 3 down-regulated, qvalue <0.05). Principal cellular pathways were deregulated during the sensitization phase, such as Inflammation, sirtuin Signaling, PXR/RXR,NRF2-mediated Oxidative Stress Response, Planar cell polarity (PCP) signaling, PI3K/AKT and RhoA signaling. Moreover, transcriptional factors including HNF4 (Hepatocyte nuclear factor 4 alpha) that plays a central role in liver differentiation and FGF19 (Fibroblast growth factor 19) which regulates the synthesis of bile acids through the FGFR4 / Klotho-β complex in the liver, were also impaired. The activation of FGF19 was confirmed since the expression of its target genes (Cyp7a1, Idh3a, Cytc, Atp5b, Pepck, G6pase and Pgc) was significantly modified. Our preliminary data provide an integrated framework to understand in detail the mechanisms associated to liver regeneration and may prove to be translatable to humans and benefit liver diseased patients.

870

Proteomic profiling reveals key cancer progression modulators in shed microvesicles released from isogenic human primary and metastatic colorectal cancer cell lines

Wittaya WS Suwakulsiri[1], Alin AR Rai[2], Rong RX Xu[1], Maoshan MC Chen[3], David DG Greening[2], Richard RS Simpson[1]

1. Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia

2. Molecular Proteomics, Baker IDI Heart and Diabetes Institute, Melbourne, Victoria, Australia

3. Australian Centre for Blood Diseases, Alfred Hospital, Monash University, Melbourne, Victoria, Australia

Extracellular vesicles (EVs) comprise two main classes - exosomes and shed microvesicles (sMVs) with distinct mechanisms of biogenesis. Whilst much is known about exosome cargo content and functionality, sMVs are poorly understood. Here, we describe the large-scale purification of sMVs released from primary (SW480) and metastatic (SW620) human isogenic colorectal cancer (CRC) cell lines using a combination of differential ultracentrifugation and isopycnic iodixanol density centrifugation. The yield of SW480-sMVs and SW620-sMVs was 0.75 mg and 0.80 mg, respectively. Both SW480-/SW620-sMVs are heterogeneous in size (100-600 nm diameter) and exhibit identical buoyant densities (1.10 g/mL). We show that sMVs, unlike exosomes, are ALIX$^-$,

TSG101$^-$, CD63$^-$ and CD9$^-$. Quantitative mass spectrometry identified 1295 and 1300 proteins in SW480-sMVs and SW620-sMVs, respectively. Gene Ontology enrichment analysis identified 'cell adhesion' (CDH1, OCLN, CTN families), 'signalling pathway' (KRAS, NRAS, MAPK1, MAP2K1), and 'translation/RNA related' processes (EIF, RPL, HNRNP families) in both sMV types. Strikingly, SW480- and SW620-sMVs exhibit distinct protein signatures. sMVs from primary tumour cell lines are enriched in ITGA/B, ANXA1, CLDN7, CD44 and EGFR/NOTCH signalling networks, while metastatic tumour-derived-sMVs are enriched in PRKCA, MACC1, FGFR4 and MTOR/MARCKS signalling networks. Both SW480- and SW620-sMVs are taken up by NIH3T3 fibroblasts and exhibit fibroblast invasion capability, suggesting a role of sMVs in intercellular communication. We report, for the first time a comprehensive analysis of a hitherto undescribed subpopulation of EVs. We anticipate our i*n vitro* findings will be a starting point for more sophisticated studies aimed at elucidating the biochemical and functional properties of EV subpopulations *in vivo*. Furthermore, the emerging roles of specific EV subpopulations in the tumour microenvironment we believe will alter our view of cancer biology and undoubtedly present new targets for therapeutic intervention.

871

Proteome and phosphoproteome changes during left ventricular dysfunction in post-myocardial infarction

Hannah LF Swa[1], Sunitha Nair[2], Christopher J Charles[3], Carolyn SP Lam[4], Arthur M Richards[3], Jayantha Gunaratne[1]

*1. Institute of Molecular and Cell Biology, A*STAR, Singapore, SINGAPORE*

2. National University Health System, Singapore

3. National University of Singapore, Singapore

4. Duke-NUS, Singapore

Having mechanistic insights into early post-myocardial infarction (MI) is key to developing specific diagnostic and therapeutic strategies for early detection and intervention in heart failure with reduced ejection fraction (HFrEF). The study here aims to understand molecular event/s specific to early progression of HFrEF. We hypothesize that unbiased systemic screening of proteins and phosphoproteins in relevant left ventricular (LV) tissues can illuminate pathophysiological processes and underlying mechanisms implicated in HFrEF. We performed mass spectrometry-based multiplexed and label-free quantification for protein and phosphoprotein analysis respectively on LV tissues of normal (healthy pigs), normally perfused region (control) and viable myocardium surrounding the MI region (peri-infarcted) of HFrEF pig models. With over 3000 proteins quantified in the multiplexed experiments, network analysis of differentially-regulated proteins (mapped unto the human UniProt accession numbers) in the different LV regions highlighted extracellular matrix (ECM) organization as the main significant network involved in early HFrEF. ECM has been proposed to contribute to overall LV remodelling process in post-MI; and to further focus on the signalling events that may lead to this, we performed phosphoprotein pathway analysis based on confident differentially-regulated phosphosites in the different LV regions. With a total of over 14,000 phosphosites identified and mapped to over 3000 proteins, preliminary analyses revealed phosphorylation changes in cytoskeletal remodeling mediated by Rho GTPase effectors and PKA, and adrenergic signalling. Detailed network and pathway analyses of the phosphoproteins will be carried out to confirm this initial finding, which will subsequently shed light unto mechanistic details for further experimental validation. Understanding and scrutinizing the molecular events at both the protein and phosphorylation levels using this HFrEF pig LV tissues can potentially elucidate potential diagnostic and therapeutic target/s to reverse the underlying LV remodelling.

872

Discovering the N-linked glycome of the important zoonotic parasite Fasciola

Jaclyn Swan[1], Kim Johnson [1], Tony Bacic [1], Travis Beddoe[1]

1. Animal, plant and soil sciences , La Trobe University, Bundoora, Victoria , Australia

Fasciola hepatica is a globally distributed zoonotic parasite which causes a large economic impact on the agricultural sector and human health. Due to the increasing prevalence of drug resistance, alternative treatment methods are urgently required. Many protein antigens have been trialled as vaccine candidates, with promising but variable success. It has been a common observation that native vaccines have far more efficacy than their recombinant orthologues and this is likely due to altered post-translational modifications with the recombinant proteins. The tegument of the parasite, which comes in direct contact with the host immune system, is highly glycosylated and a major source of potential vaccine antigens. Hence, glycosylation patterns of *F. hepatica* glycoproteins warrant further investigation. To better predict potential glycan structures that may be present, *in silico* methods have been used to identify potential glycosyltransferases that are present in *F. hepatica* and the closely related species *F. gigantica*. This has demonstrated many similarities between the two species, as well as highlighting a novel xylosyltransferase not previously known to exist in *Fasciola* spp. It is known that N-glycan structures can be quite heterogenic, and therefore to understand if this occurs in *F. hepatica*, we have used hydrophilic interaction liquid chromatography followed by high-resolution tandem mass spectrometry with complementary and triggered fragmentation to reveal a number of glycoproteins containing a range of N-glycans. The majority of these N-glycans have an oligomannose or paucimannose structure. These N-glycans possibly play a role in host recognition and/or protein folding. This information will provide insight into the natural occurrence of N-glycan heterogeneity on *F. hepatica* glycoproteins providing new information that can be considered when engineering a recombinant expression system for potential vaccine candidates in the future.

873

Comprehensive profiling of neural retina proteins in C57BL/6 mouse with S-Trap and high-pH peptide fractionation by mass spectrometry

Ying Hon Sze[1], Ka Wai Cheung[1], King Kit Li[1], Chuen Lam[1]

1. Laboratory of Experimental Optometry, Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Purpose: The mechanism of axial elongation in the myopic eyeball remains to be elucidated. The retina is a delicate tissue that receives visual input and transmits the molecular signals to guide the eye growth. Reliable ocular biometric data have provided evidence to support the use of C57BL/6 mouse as the animal model susceptible to experimental myopia. We performed biometric measurements of the eyes in normal growing C57BL/6 mouse and used proteomics approach to establish a proteomic database of neural retina, revealing expressed proteins related to the developmental pathways and providing the basis for examining the molecular mechanisms underlying the development of myopia in future study. **Methods**:Three C57BL/6 mouse retina were obtained at postnatal day 46. Retinas were lysed with 5% SDS buffer. Proteins were tryptic digested with suspension traps (S-Traps). Peptides in each retina were separated into 6 fractions with high-pH peptide reversed-phase fractionation. In total, 18 fractions of peptides were analyzed with the data-dependent acquisition by using Sciex TripleTOF® 6600 mass spectrometer followed by bioinformatics analysis. **Results**: A total of 7122 non-redundant proteins (n=3, 1% FDR) were identified in the normal fractionated mouse retina. High-confidence proteins were identified in each fraction, with 909, 1780, 1674, 1876, 1620 and 1576 proteins respectively (n=3, 2 technical replicates). In result of 1612 fraction-specific proteins. KEGG pathway analysis revealed protome with high coincidence to reported signalling pathways(number of proteins), such as MAPK(91), ras(78), mTOR(64), axon guidance(64), cAMP(56), insulin(51); synapses, including glutamatergic(50), Dopaminergic(45), GABAergic(40) and cholinergic(40). **Conclusion**: Mass spectrometry enabled efficient screening of pathways involved in experimental condition. S-Trap protocol provides highly efficient and less complicated sample preparation for mouse retinal proteomics study. The comprehensive C57BL/6 mouse retinal proteome database with enriched pathways interested in myopia study, shall take advantage with the use of SWATH-MS acquisition to quantify multiple pathways in single experiment.

874

Cys-DIA – Cysteine specific DIA increase the proteome coverage

Muhammad Tahir[1], Arkadiusz Nawrocki[1], Martin R. Larsen[1]

1. University of Southern Denmark, Odense, FYN, Denmark

Background

Recent advances in mass spectrometry still faces a key challenge to maximize the complex proteome coverage. Although, data independent acquisition (DIA) is a very powerful tool to perform both protein identification and quantification of the complex samples still the ordinary approaches require more improvements to further extend the proteome in-depth analysis.

Cysteine is the second rarest amino acid but present in about 97% of the human proteome. By using a robust method for enrichment of cysteine-containing peptide in combination with DIA we should be able to increase the proteome coverage compared with normal DIA

In our study, we present Cysteine-DIA (Cys-DIA) approach that allows a successful and comprehensive characterization of the complex proteome with high dynamic range of protein abundances.

Methods

All experiments were performed on Q Exactive HF-X mass spectrometer using a high resolution MS1 based quantitation data-independent acquisition (HRMS1-DIA). Chromatographic separations were performed using EASY-nLC 1000 system (Thermo Fisher Scientific) and Evosep (Biosystems). HeLa and breast cancer cell line were used for analysis. Performed cysteine enrichment followed by high pH fractionation to generate sample specific spectral libraries. For quality control and retention times normalization, reference peptides were spiked into all samples. Spectral libraries generation and DIA data analysis were performed in Spectronaut PulsarX (Biognosis).

Preliminary Data

The Cys-DIA raw files were searched against the HeLa cysteine and total spectral library and we identified 5027 protein groups whereas for the total HeLa 3827 were identified. The cancer cell line 21 min of Cys-DIA search analysis resulted in the identification of about 2800 out of which about 2000 proteins were quantified across all replicates and all the three different cell line with average CV of 11%.

Novel Aspect

The cysteine-DIA approach reduces the proteome complexity and allows identification of a larger number of proteins in short time analyses.

875

In the lupin: Exploring solvent extraction methods for characterising lupin allergens by LC-MS

Arineh Tahmasian[1], Mitchell G Nye-Wood[1], Angela Juhasz[1], Michelle L Colgrave[1, 2]

1. Edith Cowan University, Joondalup, WA, Australia

2. CSIRO Agriculture and Food, St Lucia, QLD, Australia

The valuable health-promoting benefits and technological characteristics of lupin have led to its growing use as an ingredient in a broad range of food products. However, as a result of growing consumption of this valuable legume in the food industry, the number of allergic reactions to lupin proteins has reportedly increased, with noted cross-reactivity to other legumes, particularly peanut. To protect public health and safety of the allergic individuals, lupin declaration on the ingredient list has become

mandatory in many countries including Australia, hence, it is crucial to design reliable methods for accurate detection of lupin proteins in food products.

In this study the proteins of lupin flakes produced from Australian narrow-leaved lupin (NLL) (*Lupinus angustifolius*) have been identified through a sensitive bottom-up proteomics approach and bioinformatic analysis. The proteome of NLL resulted from using three different extraction buffer composition were identified and compared. In addition, the effects of the pre-extraction defatting step on the number of identified proteins were investigated. Discovery proteomics was performed on a TripleTOF 6600 (SCIEX) mass spectrometer which was followed by the identification of peptides and proteins using ProteinPilot software employing a genome-derived protein database.

The highest number of protein identifications was achieved using a urea-based extraction method which led to characterisation of many seed storage globulin proteins (conglutins) that exhibited high sequence similarity to known reference food allergens of NLL. A pre-extraction defatting step did not significantly increase the protein identifications.

The choice of sample preparation method influences the composition of the mass spectrometry-derived lupin proteome.

876

Assessment of drug-associated HLA antigens involved in co-amoxiclav induced liver injury.

Arun A Tailor[1], James C Waddington[1], Kareena Adair[1], John Farrell[1], Jane Hamlett[1], Gordon Dear[2], B. Kevin Park[1], Xiaoli Meng[1], Dean J Naisbitt[1]

1. MRC Centre for Drug Safety Science, University of Liverpool, Liverpool, UK

2. GlaxoSmithKline, Ware, UK

Background: Exposure to the antibiotic amoxicillin is associated with the development of idiosyncratic adverse drug reactions. Amoxicillin-specific T-cells have been detected in patients with severe liver injury, suggestive of an immune disease aetiology. This theory has gained credence with the discovery of a number genetic associations including the HLA DRB1*15:01-DQB1*06:02 haplotype. β-lactam antibiotics form drug-protein adducts by conjugating with lysine residues that are postulated to activate T-cells, following protein processing and the liberation of peptide epitopes. *In vitro* studies have characterised drug-modified proteins in patient sera, however, the exact epitopes being presented on T-cells have not yet been identified. This study involves a parallel approach to identify the nature of class II epitopes including a designer drug-modified peptide methodology and an interrogation of the naturally eluted immunopeptidome. T-cell models were employed to assess the immunogenicity of candidate drug-associated antigens.

Methods & Results: Designer peptides were synthesised containing (1) anchors for the HLA DRB1*15:01-DQB1*06:02 haplotype and (2) amoxicillin-bound lysine residues in several locations representative of possible drug-TCR contact sites. The immunoaffinity capture of MHC class II antigens was carried out on B-cell lines derived from healthy volunteers homozygous for the risk alleles. Both designer peptides and MHC eluted peptides were purified using HPLC and characterised using proteomics/ immunopeptidomics methodologies on a TripleTOF 6600 mass spectrometer (AB Sciex). Candidate peptides were incubated with PBMCs from hypersensitive patients positive for the risk haplotype to generate antigen-responsive T-cell lines. Amoxicillin-modified peptide-specific T-cells proliferated and secreted cytokines such as IFNγ in a dose dependent manner with high specificity to the antigen showing no cross reactivity with unmodified peptides or with positional derivatives at different TCR contact sites. T-cell responses were restricted to the HLA DRB1*15:01-DQB1*06:02 risk alleles.

Conclusion: Here we demonstrate an approach to elucidate and validate drug-associated-antigens which may be implicated in drug hypersensitivity reactions.

877

The advanced immuno-MS work flow for serum biomarker quantification using the antibody-immobilizing magnetic beads

Yoshio Takahata[1], Misato Hara[2], Kouhei Nishino[1], Naohiro Hanyu[2], Takao Kawakami[1]

1. Medical ProteoScope Co., Yokohama, Japan

2. Tamagawa Seiki Co., Iida, Japan

Immuno-MS, mass spectrometric peptide measurement combined with immunoaffinity purification of protein, is a powerful method for quantitative analysis of low abundant proteins in biological specimens. In the procedures it is crucial to collect specifically and efficiently the target protein antigens from the antigen-antibody complex generated on the surface of the nano-scaled carrier beads. The optimized conditions of the immunoaffinity purification include elution of the target protein with an aqueous organic solvent.

As an experimental model, we used serum alpha-fetoprotein (AFP), one of the hepatocellular carcinoma biomarkers, and its specific antibody immobilized covalently on the magnetic beads. The human serum was spiked with known amounts of the standard AFP. The antibody-coated beads were incubated with the spiked serum for antigen-antibody reaction. AFP was then eluted from the antibody-coated magnetic beads with 0.1% TFA aqueous solution containing organic solvents. The eluted protein was hydrolyzed with a trypsin/LysC mixture. Stable isotope-labeled standard peptides were added to the hydrolysate to quantitate the eluted AFP by LC-MS/MS.

Various organic solvents in the elution solution facilitated AFP recovery, which was maximized with an aqueous solution containing 50% acetonitrile and 0.1% TFA. Moreover, using the optimized workflow, quantitative analyses showed a correlation between the amounts of AFP spiked into serum (0–100 ng/ml) and the corrected ion intensities of the tryptic peptides from AFP ($R^2 > 0.99$). The concentration of the endogenous AFP was calculated as 2.3 ± 0.6 ng/ml from the standard curve regression equation. This result is consistent with previous reports for AFP concentration in healthy human sera (< 10 ng/ml). The present immuno-MS workflow is easily applicable to detection and quantitation of the other low abundant biofluid biomarkers.

878

Proteome-wide analysis of USP14 substrates revealed its role in hepatosteatosis via stabilization of FASN

Minjia Tan[1], Yan Lu[2], Xiaoying Li[2]

1. Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Zhangjiang, SHANGHAI, China

2. Department of Endocrinology and Metabolism, Fudan University, shanghai, China

Publish consent withheld

879

SWATH-MS reveals functional differences between glycogenin 1 and 2 in yeast metabolism

Xinle Tan[1], Ben Schulz[1]

1. school of chemistry and molecular bioscience, University of Queensland, Brisbane, QLD, Australia

Yeast accumulate glycogen as a major carbon and energy reserve to buffer against environmental starvation conditions. Glycogen is a branched polysaccharide of glucose that is initiated from glycogenin, a self-glucosylating priming protein. There are two homologs of glycogenin in human and yeast, Glg1p and Glg2p, both of which are capable of priming glycogen synthesis. While Glg1p and Glg2p have distinct sequence characteristics, differences in their molecular function and physiological roles have not been investigated. Here, we created yeast strains lacking either or both glycogenins, either genetically deficient or complemented. Analysis of glycogen content showed differences between these yeast strains depending on the presence of Glg1p or Glg2p. We then performed global SWATH-MS analysis of these strains, which identified large and significant changes in their proteomes. In particular, we identified changes in central carbon metabolism, protein stress response, and trehalose regulation. Our results support a critical role for glycogen in general stress response beyond a carbon storage molecule, and suggest different specific functions for Glg1p and Glg2p.

88

Proteomic dynamics of colorectal cancer evolution identifies PLOD2-mediated microenvironmental regulation as a novel drug target

Kailun Xu[1], Yingkuan Shao[1], Xi Zheng[1], Biting Zhou[1], Shaozheng Dai[2], Huanhuan Gao[3, 4], Guan Ruan[3, 4], Tiansheng Zhu[3, 4], Yaoting Sun[3, 4], Yi Zhu[3], Shu Zheng[1], Tiannan Guo[3, 4]

1. Second Affiliated Hospital, and the Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education, Zhejiang University College of Medicine, Hangzhou, zhejiang Province, China

2. School of Computer Science and Engineering, Beihang University, Beijing, China

3. School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

4. Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China

Background The molecular evolution of colorectal cancer (CRC) from normal colon (N), hyperplastic polyp (P), adenomas (A) to carcinoma (C) remains largely unclear, preventing effective diagnosis and drug target discovery. Here we aim to profile the proteomic dynamics of CRC evolution using primary tissue samples and discover novel drug targets.

Methods We procured 180 FFPE tissue samples from 90 patients (two biological replicates), including 20 normal (N), 18 polys (P), 22 adenoma (A), 15 tubular adenocarcinoma and 15 mucinous adenocarcinoma (C). The samples were processed using pressure cycling technology (PCT) followed by data independent acquisition (DIA) mass spectrometry. Selected proteins were further verified using parallel reaction monitoring (PRM). PLOD2 was further investigated for potential as drug targets in CRC cell lines and mouse model.

Results We quantified 4858 SwissProt proteotypic proteins from 180 samples with high technical reproducibility (r =0.97), and identified 155 proteins consistently up- or down-regulated along the CRC evolution process, as well as 491 up-regulated proteins in the early carcinogenesis (adjusted p value < 0.05, fold change > 2), most of which are involved in inflammation. The transition from A to C is accompanied by up-regulation of 63 proteins mainly involved in metabolism. Then we further verified the expression of 31 dysregulated proteins using PRM, and focused on a microenvironmental protein lysyl hydroxylase 2 (PLOD2) as a novel drug target. PLOD2 knocked-out HCT116 and HT-29 were created using Crispr-Cas9, exhibting low proliferation and migration. Finally we observed the effectiveness of administration of an FDA approved PLOD2 inhibitor Minoxidil in 3 CRC patient-derived tumor xenograft.

Conclusion We demonstrated the effectiveness and practicality of PCT-DIA in studying the proteome evolution of CRC samples based on FFPE biopsy tissue samples, and identified microenvironmental regulation, driven by PLOD2, as a critical contributor that can be exploited as novel therapeutics for CRC.

880

Mass++ ver.4: an open-source, simple and extensible MS data viewer

Satoshi Tanaka[1], Masaki Murase[2], Masaki Kato[2], Tsuyoshi Tabata[2], Maiko Kusano[3], Shin Kawano[4], Susumu Goto[5], Yasushi Ishihama[2], Akiyasu C Yoshizawa[2]

1. Trans-IT, Kaminokawa-machi, TOTIGI [TOCHIGI], Japan

2. Grad. School of Pharma. Sci., Kyoto Univ., Kyoto, Japan

3. Res. Inst. of Environ. Med., Nagoya Univ, Nagoya, Japan

4. Toyama Intl. Univ., Toyama, Japan

5. DBCLS, DS, ROIS, Kashiwa, Japan

Background

We had developed a software platform “Mass++,” which integrates multiple functions including viewers and file-open functions for multiple kinds of mass spectrometry raw data on Windows environment (J Proteome Res., 13 (8): 3846-3853 (2014)). Now we are rewriting the entire software without OS-depending libraries for the multiple-OS-platforms as an open-source package; we have completed to implement the data viewer, and continue to develop function for proteome analysis.

Methods

Mass++ ver.4 is written in Java instead of C++, which was used for past versions, for effective development of cross-platform software for the maintainability. The plug-in structure, the characteristics of past versions of Mass++, remain in this version.

Results

The current build has the following features; the development is still continuing.

- Plug-in structure. Easy to extend and/or release function when needed as same as the past versions.
- Cross-platforms. The current build can run not only on Windows but also on Mac OS/X and Linux, in contrast with the previous versions that could run only on Windows.
- Currently mzML files, the standard and portable file formats, are readable.
- Simple-designed viewer. The viewer can display fundamental data for mass spectrometry: mass spectra, chromatograms, and heatmaps.
- Utilizing external functions. Mass++ can read result files from the external program, Comet, for database search and display the list of identified peptides and the annotation of the identified peaks on the heatmap.

Conclusion

Our present purposes of developing Mass++ ver.4 is to cooperate easily with external software and visualize intermediate results during the analysis process; these are necessary for the effective analysis and rarely implemented in freeware. Mass++ ver.4 is released as a BSD 3-claused open-source software at http://www.mspp.ninja/.

881

Glycosylation of N-Glycan Branching: Implications for Disease Onset, Biomarker and Therapeutics

Naoyuki Taniguchi[1]

1. Osaka International Cancer Institute, Osaka, JAPAN (JPN), Japan

Our group has been interested in glycosyltransferases such as GnT-III, GnT-IV, GnT-Va, Vb (GnT-IX), Fut8 and St6Gal1 and discovered several target proteins on which these enzymatic products carry (1). Core fucose, a product of α 1, 6 fucosyltransferase (Fut8), is a cancer biomarker target (1) and is also implicated in COPD (chronic obstructive lung disease), a progressive and inflammatory airway disease due to cigarette smoking and environmental chemicals. . We found that a keratan sulfate disaccharide designated L4 ([SO3--6]Galβ1-4[SO3--6] GlcNAc) showed protective effects in two murine COPD models. L4 attenuated alveolar destruction, reduced neutrophil influx and inflammatory cytokines, inactivated matrix metalloproteinase and myeloperoxidase in bronchoalveolar lavage fluid (2). We have also identified receptor protein of L4 and the underlying mechanism by which L4 suppresses inflammation.

Bisecting GlcNAc, a GnT-III product was high in Alzheimer’s disease patients (2). Analysis of knockout mice of GnT-III revealed that decreased cleavage of APP (Aβ-precursor protein) by BACE1 (β-site amyloid precursor protein cleaving enzyme-1) as well as decreased Aβ plaque. The lack of this modification directs BACE1 to late endosomes/lysosomes where it is less co-localized with APP, leading to accelerated lysosomal degradation.

GnT-Va producesb1,6 GlcNAc structure is known to play a key role in EMT (epithelial to mesenchaymal transition) and cancer metastasis due to modification of adhesion molecules such as E-cadherin and integrins. . Recently crystal of GnT-Va was obtained and bi-substrate analogue of GnT-Va is a good inhibitor for Gnt-V.

In conclusion new findings of various glycosyltransferase functions will open a new avenue toward novel and promising druggable candidates.

882

Defining the protein *N*- and *O*-glycosylation associated with human monocyte-to-macrophage transition

Hannes Hinneburg[1], Morten Thaysen-Andersen[1], Erdmann Rapp[3, 2], Alexander Pralow[3], Nilesh Bokil[4], Jessica Pedersen[4], Rebeca Sakuma[1], Falko Schirmeister[5], Bernadette Saunders[4, 6]

1. Macquarie University, NSW, Australia

2. glyXera GmbH, Magdeburg, Germany

3. Max Planck Institute for Dynamics of Complex Technical Systems, MPI Magdeburg, Magdeburg, Germany

4. Faculty of Science, University of Technology, Sydney, NSW, Australia

5. None, Berlin, Germany

6. Centennary Institute, Sydney, NSW, Australia

Monocytes are abundant mononuclear precursors to macrophages, immune cell populations essential for our innate immune system. Protein glycosylation is recognised to impact the function of monocytes and macrophages, but the exact glycophenotypes underpinning the monocyte-to-macrophage transition remain undocumented. Herein, we use label-free PGC-LC-MS/MS-based glycomics to determine the protein *N*- and *O*-glycosylation associated with human monocyte-to-macrophage transition. Primary CD14+ monocytes from blood of three healthy donors were differentiated and polarised *ex vivo* to classical M1-macrophages using a conventional seven-day GM-CSF differentiation protocol with longitudinal sampling at day 0, 1, 3, 5 and 7. Accompanying confocal microscopy and proteomics were used to confirm the expected morphological changes and the down-regulation of known monocyte-specific proteins including myeloperoxidase, neutrophil elastase, and cathepsin G and altered protein pathways impacted by macrophage maturation. Glycomics surprisingly demonstrated that monocytes and macrophages display a similar glyco-phenotype comprising predominantly paucimannosidic ($Man_{1-3}GlcNAc_2Fuc_{0-1}$) (22-31%) and oligomannosidic ($Man_{5-9}GlcNAc_2$) (29-35%) *N*-glycans as well as complex α2,3- and α2,6-sialylated bi- and tri-antennary *N*-glycans with and without core fucosylation (31-43%). Relatively large donor-specific *N*-glycome variations were observed possibly reflecting a considerable physiology-dependent heterogeneity of monocytes and macrophages. Only few consistent *N*-glycome changes were observed to correlate with the monocyte-to-macrophage transition across all donors including a decrease in core fucosylation and a slightly reduced expression of mannose-terminating *N*-glycans (paucimannosidic/oligomannosidic) on mature macrophages. The less heterogeneous *O*-glycome also largely remained unchanged over the seven-day macrophage maturation, findings that were supported by a largely unchanged *N*- and *O*-glycosylation machinery as evaluated by quantitative proteomics. This high-resolution system-wide map of the protein *N*- and *O*-glycome associated with healthy monocyte-to-macrophage transition, the most detailed to date, aids our understanding of the molecular makeup pertaining to key innate immune cells and forms an essential reference library of value to future glycoimmunological studies involving human monocytes and macrophages.

883

Exendin-4 protects pancreatic *BTC-6* cells against metabolic stress: a proteomic study

Dhanya Madhu[1], Maha Hammad[1], Sina kavalakatt[1], Abdelkrim Khadir[1], Ali Tiss[1]

1. Dasman diabetes Institute, Kuwait, KUWAIT

Background and objectives: Lipotoxicity is an important factor in the pathogenesis of type 2 diabetes resulting in defective β-cell proliferation and increased apoptosis. Increase in glucagon-like peptide-1 (GLP-1) activity has recently emerged for the treatment of type 2 diabetes (T2D) by enhancing the glycemic control and helping in decreasing body weight of most patients. Recent studies have suggested beneficial effects of these peptides on insulin-responsive tissues such as adipose tissue, muscles and liver aside with pancreatic β-cells. We investigated here the potential beneficial effects of a GLP-1 mimetic on mouse pancreatic cells under stressing levels of palmitic acid (PA).

Methods: Using mouse pancreatic cell line (BTC-6), we investigated the effect of a GLP1 mimetic (Exendin-4) on the protein expression pattern using Mass Spectrometry approaches (LC-MS/MS Orbitrap system and label-free quantification) in the presence of stressing amounts of PA. We also investigated the impact of Exendin-4 on MAPKs using RT-PCR and Western blot. Cell viability assay and lipid droplet accumulation assessment were also performed.

Results: We showed that ERK MAP-Kinase phosphorylation was highly increased by Exendin-4 both in presence and absence of PA. Furthermore, cell viability assays have shown that Exendin-4 significantly alleviated the PA-induced cell death. This was further confirmed with proteomics analysis where various cellular functions were improved in presence of Exendin-4, including cell growth, cellular assembly and organisation. Moreover, proteomics analysis highlighted a panel of interconnected heat shock proteins (HSP) that have been modulated by Exendin-4. This was further confirmed by Western blot where the heat-shock inducible HSP72 was significantly increased by PA and attenuated in presence Exendin-4.

Conclusion: Our results suggest that GLP-1 mimetics alleviate the lipotoxicity-related cellular stress in pancreatic cells and enhance heat shock response thus restoring normal cellular homeostasis.

884

Unusual Site- and Granule-Specific *N*-Glycosylation of Human Myeloperoxidase from Resting Neutrophils

Harry Tjondro[1], Julian Ugonotti[1], Sayantani Chatterjee[1], Ian Loke[2], Si-Yun Chen[3], Hannes Hinneburg[1], Benjamin L Parker[4], Vignesh Venkatakrishnan[5], Regis Dieckmann[5], Weston Struwe[3], Anna Karlsson[5], Morten Thaysen-Andersen[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Department of Biological Science, National University of Singapore, Singapore

3. Department of Chemistrty, University of Oxford, Oxford, South East England, United Kingdom

4. Department of Physiology, University of Melbourne, Melbourne, VIC, Australia

5. Department of Rheumatology and Inflammation Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Myeloperoxidase (MPO), which predominantly resides in the primary granules of resting neutrophils but reportedly also in other neutrophil compartments at lower levels, is an important glycoprotein in neutrophil-mediated innate immunity. Most structural features of human neutrophil MPO and its capacity to produce anti-microbial hypochlorous acid are well-documented, but the exact site- and granule-specific glycosylation of MPO remains undocumented, despite the recognised importance of *N*-glycosylation for its structure-function. Herein, we performed a deep structural characterisation of the *N*-glycosylation of human MPO from healthy donors using LC-MS/MS-based glycomics, intact glycopeptide and glycoprotein profiling. Quantitative glycomics and glycopeptide analyses of MPO from neutrophil lysates revealed that all five *N*-glycosylation sites of the heavy chain MPO monomer are conjugated with heterogeneous *N*-glycans albeit with strong site-specific differences. Paucimannosidic *N*-glycans were predominantly found to be linked to Asn323 (47%) and Asn483 (56%), oligomannosidic *N*-glycans were mainly carried by Asn355 (97%) and Asn391 (64%) while Asn729 was largely unoccupied (44%) or modified with chitobiose core *N*-glycans (33%). Importantly, the glycoprofiles of the intact MPO monomer and dimer obtained using high-resolution top-down mass spectrometry recapitulated the heterogeneity observed at the bottom-up level as supported by matching experimental and theoretical intact glycoprotein masses. Further, glycoprofiling of MPO from isolated neutrophil granules intriguingly demonstrated that the *N*-glycosylation of Asn355, Asn391 and Asn729 varied dramatically whereas Asn323 and Asn483 were similarly glycosylated on MPO across the studied neutrophil compartments. Particularly the abundant primary granule-resident MPO displayed unique glycosylation signatures comprising a higher proportion of oligomannosidic *N*-glycans on the three variable glycosylation sites. In conclusion, this study represents the most detailed structural characterisation of MPO *N*-glycosylation to date. The complex site- and granule-specific *N*-glycosylation of MPO, shown here for the first time, adds fundamental knowledge that aid our understanding of the fascinating glycobiology underpinning neutrophil-mediated immunity.

885

Immunomodulation of *N*-glycolylneuraminic Acid Glycans Using Nanotechnology Platform

Suraj Toraskar[1], Preeti Chaudhary[1], Raghavendra Kikkeri[1]

1. Indian Institute of Science Education and Research, PUNE, MAHARASHTRA, India

The sialic acid (Sia) family of sugars are terminally present on the cell surface with *N*-glycolylneuraminic acid (Neu5Gc) and *N*-acetylneuraminic acid (Neu5Ac) as the predominant form on most mammalian cells. Neu5Gc is deficient in humans due to an inactivating deletion in the *CMAH* gene encoding the hydroxylase responsible for the conversion of CMP-Neu5Ac to CMP-Neu5Gc. On the contrary, Neu5Gc is metabolically incorporated into human tissues from dietary sources (principally red meat), and also detected at even higher levels in human tumors. The up-regulation of sialylation in cancer might also explains, why ingested Neu5Gc preferentially accumulates in cancer tissues. Commensal bacteria incorporate dietary Neu5Gc into lipopolysaccharides, this leads to the generation of antibodies in human, ranging widely in levels among individuals. Currently known monoclonal antibody is unable to detect Neu5Gc in its native form and polyclonal antibody has less specificity and high background reactivity. Therefore generation of specific antibodies against Neu5Gc is essential which could be used as a potential marker for cancer. Herein we proposed the nanotechnology based platform to generate the specific antibodies against Neu5Gc glycans. We designed and synthesized 2,3 and 2,6 linked Neu5Gc trisaccharides, which were further conjugated to CRM protein and spherical gold nanoparticles of different sizes. These glycan functionalized nanoparticles and CRM protein was injected in mice and sera were collected at different time intervals. This serum was analyzed for antibody response using glycan microarray, which showed particularly IgG immune response against 2,3 linked Neu5Gc glycan.

886

Integrating MALDI Imaging Mass Spectrometry with shotgun proteomics for the studies of neuropathology of Alzheimer's disease

Yumiko Toyama[1], Nobuto Kakuda[1], Ryo Kajita[2], Takashi Nirasawa[2], Tomohiro Miyasaka[1], Shigeo Murayama[3], Yasuo Ihara[4], Masaya Ikegawa[1]

1. Department of Life and Medical Systems, Doshisha University, Kyotanabe, Kyoto, Japan

2. Bruker Japan K.K., Yokohama, Kanagawa, Japan

3. The Brain Bank for Aging Research, Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology , Tokyo, Japan

4. Graduate School of Brain Science, Doshisha University, Kyotanabe, Kyoto, Japan

Neuropathology of Alzheimer's disease (AD) is characterized by the accumulation and aggregation of Amyloid β (Aβ) peptides into extracellular plaques of the brain. The Aβ peptides are generated from amyloid precursor proteins by β- and γ-secretases. Aβ deposited not only in cerebral parenchyma but also in leptomeningeal and cerebral vessel walls. This has been known as cerebral amyloid angiopathy.

Here, we adopt MALDI-imaging mass spectrometry (MALDI-IMS) on autopsied brain tissues to obtain a comprehensive protein mapping. Human cortical specimens for IMS were obtained from brains that were removed, processed, and stored at -80 ˚C within 8 h postmortem at the Brain bank at Tokyo Metropolitan Institute of Gerontology. Each brain specimen was taken from the occipital cortex of five AD patients and five non-pathological controls. This study was approved by the ethics committee at each hospital or institute. Cryosections were cut and transferred to Indium-Tin-Oxide coated glass slides. Spectra were acquired using the rapifleX in positive linear mode, whereas ions were detected in a mass range of *m/z* 2,000-20,000 with spatial resolution of 20-100 *μm*. Matrix was uniformly deposited on the slide using the HTX-sprayer. Visualization and statistical analysis were used flexImaging and SCiLS Lab.

The current analysis clarifies that Aβ1-42 and Aβ1-43 were selectively deposited to senile plaque and shorter Aβ peptides were deposited to leptomeningeal blood vessels. In order to deepen proteomic information with the current specimen, we have dissected a small piece of tissues from leptomeningeal vessels as well as parenchymal area with laser micro dissection and were applied to LC-TIMS-TOF-MS/MS analysis. Data analysis was done with Proteinscape. From single vascular structure and adjacent cortical parenchyma dissected with 0.25-0.5 mm^2 yields a thousand of protein annotation. Further analysis including detected peptide fragments as well as immunohistochemistry with a specific epitope recognition will validate the current strategy.

887

Quantification of histone proteoforms through top-down proteomics

Denise P Tran[1], Muhammad A Zenaidee[1], Clara T Cano[1], Chen Cheng[1], Joseph A Loo[1]

1. UCLA, Los Angeles, CALIFORNIA, United States

Top-down proteomics is an emerging mass spectrometric technique with the ability to identify intact proteoforms. Currently, there is no method to concurrently quantify proteins and post-translational modifications (PTMs) through top-down methods. Histones have a diverse range of proteoforms stemming from their tails which are susceptible to chemical modifications and directly influence the compaction of the chromatin. Due to the positively-charged and unstructured termini, quantification of histone PTMs remains a challenge for bottom-up histone analysis. The typical analysis relies on propionylation that allows proteolysis to only occur at the C-terminal of arginine residues. The peptides then require assembly and reconstruction to provide information on the histone and this can produce ambiguous results. Top-down proteomics, where intact proteoforms are analysed is therefore a better fit for quantifying multiple PTMs on histone isoforms. We have developed a process to investigate the quantitative changes of PTMs on histones grown in varying media conditions.

Histones grown in HeLa cells were acid extracted and dissolved in milliQ H_2O and 1% FA and 2 µg of each sample was loaded onto a C8 HPLC column and subjected to on-line fractionation coupled to a 15-Tesla Solarix FT-ICR mass spectrometer. Raw files were exported and subjected to an in-house developed program where theoretical intact proteoform masses are matched to the deconvoluted LC-MS data and the relative quantification of the proteoforms are extracted. Current results show significant differences in patterns of PTMs found for histones grown in HeLa cells in media with and without amino acids.

The development of quantifying algorithms for high resolution top-down proteomics will determine the magnitude of different PTMs that occur simultaneously. Here, the identification and quantification of concurrent PTMs on histone isoforms will add great value to the knowledge of histone mechanisms.

888

Direct Oligosaccharide Profiling using Thin-Layer Chromatography Coupled with Ionic Liquid-Stabilized Nanomatrix-Assisted Laser Desorption-Ionization Mass Spectrometry

Mei-chun Tseng[1], Elias Gizaw Mernie[2], Leta Deressa Tolesa[2], Yu-Ju Chen[1]

1. Institute of Chemistry, Academia Sinica, Taipei, Taiwan

2. Department of Chemical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

The in-depth characterization of glycan structures is crucial to understanding their structure-function relationships and their effects on health and various diseases. Despite advances in rapid analysis, the utility of matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) is limited for complex mixtures of carbohydrates due to their low ionization efficiency and the difficulty in separating oligosaccharides because of their high structural similarity. In this study, we developed an ionic liquid (IL)-stabilized, nanomatrix-decorated, thin-layer chromatography (TLC)-MALDI MS method for simultaneous and rapid separation, detection and identification of oligosaccharides to achieve efficient profiling. The IL demonstrated good dispersion and stabilization for the spin coating of dihydroxybenzoic acid-functionalized magnetic nanoparticles (DHB@MNPs) on the TLC plate with spot homogeneity, which contributed to the observed high reproducibility (<20% CV) and 12- and 28-fold signal enhancement. Although the TLC was not able to separate isomeric glycans, the DHB@MNPs generate diagnostic glycosidic and cross-ring cleavage ions, enabling on-spot structural elucidation of composition, sequence, branching, and linkage of glycans in each separated spot. Without chemical derivatization of glycan samples, glycan visualization by TLC and tandem MS, our integrated platform, allowed the identification of 25 oligosaccharides from human milk, and heatmap analysis revealed the variability in the oligosaccharide abundance in samples from individual donors at different lactation times, which may provide insight into the microbiota and immunity of infants. With the demonstrated simplicity of our sample preparation method along with the achieved separation and in-depth structural characterization, our approach can be used for the rapid screening of other oligosaccharide-rich samples.

889

Imaging Mass Spectrometry of the kidneys from autoimmune type I diabetes rat model

Konomi Uchida[1], Yume Mukasa[1], Kazuhiko Ishibashi[2], Yuki Kuzuhara[1], Takashi Nirasawa[3], Ryo Kajita[3], Hiroyuki Yanagi[2], Nobuto Kakuda[1], Masaya Ikegawa[1]

1. Faculty of Life and Medical Systems, Doshisha University, Kyotanabe, Kyoto, Japan

2. ONO pharmaceutical Co., Ltd, Hukui, Japan

3. Bruker Japan K.K, Yokohama, Japan

Insulin-dependent (type I) diabetes mellitus (IDDM) is an autoimmune disease with multifactorial etiology. Komeda Diabetes-Prone (KDP) rats are excellent animal models to apply for this type of study. Here we demonstrate a comprehensive matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry (IMS) approach to study the molecular distribution of proteins in type I diabetic kidney from KDP rats. Especially, we ask if the minimal changes of diabetic kidney with an autoimmune etiology at early stages of the disease progression can be traced with MALDI-IMS.

Kidneys from KDP and KND (Komeda non-diabetic) rats at 7 weeks of age were resected and snap-frozen in liquid nitrogen. 10 μm cryosections were cut and transferred to Indium-Tin-Oxide (ITO) coated glass slides. a-Cyano-4-hydroxycinnamic acid (HCCA) was uniformly deposited on the slide using the ImagePrep device and measured using rapifleX tissuetyper with a spatial resolution of 50 μm in linear mode. Ions were detected in a mass range of *m/z* 800 to 3000. For Protein ID experiments, trypsin was sprayed at room temperature using the same apparatus. The multivariate analysis was done for obtained data.

In this study, in order to find early proteomic markers of minimal diabetic renal insufficiency, we have analyzed KDP rats at 7 weeks of age by comparing with those from KND rats. Deeper analysis using supervised statistical evaluations in combination with an in-depth shotgun proteomics will reveal a number of candidate markers and marker proteomic signature directly off-tissue that may be related to early histologic changes in this model. Comprehensive proteomics IMS analysis on kidney sections from KDP rats will clarify early type I diabetic nephrotic changes at peptide and protein level.

Keywords Imaging Mass Spectrometry, type I diabetes model, KDP rat, kidney

89

Spliced epitopes are abundant and highly immunogenic components of the melanoma immunopeptidome

Pouya Faridi[1], Katherine Woods[2], Nathan P. Croft[1], Ralf Schittenhelm[1], Andreas Behren[2], Anthony W. Purcell[1]

1. Infection and Immunity Program, Biomedicine Discovery Institute & Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC, Australia

2. Olivia Newton-John Cancer Research Institute, Melbourne, Australia

Antigen-recognition by CD8+ T cells is governed largely by the pool of peptide antigens presented on the cell surface in the context of HLA class I complexes. Recent studies have shown not only a high degree of plasticity in the immunopeptidome, but also that a considerable fraction of all presented peptides are generated through proteasome-mediated splicing of non-contiguous regions of proteins to form novel peptide antigens.

We have developed a novel workflow for the identification of cis and trans-spliced peptide antigens. We have applied our approach to p-HLA derived from multiple melanoma cell line and identified~ 30% cis and trans HLA-I spliced peptides . Of note, more than more 100 spliced peptides were derived from melanoma-associated antigens (MAA) and ~40% of known MAA were only represented by spliced peptides. We have confirmed the authenticity of a series of MAA spliced peptides by corresponding synthetic peptides. Moreover, immunogenicity studies of a subset of the MAA shown that several of these peptides were shown to be immunogenic in unrelated melanoma patients.

These observations highlight the breadth and complexity of the repertoire of immunogenic peptides that may be exploited therapeutically and suggest that spliced peptides may be a major class of tumour antigens. We found spliced peptides may yield more immunogenic epitopes than are available from the viral/cancer genome. Moreover, some antigens lack high-affinity HLA-ligands and the peptide splicing mechanism can generate higher affinity neoepitopes for interaction with host HLA allomorphs. Understanding the nature and abundance of spliced peptides has a high relevance for our understanding of potential novel targets of T cell immunity and will have significant implications for further immunotherapeutic approaches.

890

Multisialylated LacdiNAc structures on PSA (PSA G-Index®) as a highly specificity-enhanced secondary biomarker for prostate cancer

Yoshimi Haga[1], Motohide Uemura[2], Kengo Takeuchi[3], Norio Nonomura[2], Koji Ueda[1]

1. Cancer Precision Medicine Center, Japanese Foundation For Cancer Research, Tokyo, Japan

2. Department of Urology, Osaka University Graduate School of Medicine, Osaka, Japan

3. Division of Pathology, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan

Prostate cancer (PCa) is one of the most commonly diagnosed cancer in men worldwide. Serum prostate specific antigen (PSA) is a powerful biomarker widely used for diagnosing PCa. However, high false positive rate of PSA screening is a great issue to be resolved.

In this study, we performed comprehensive and quantitative profiling of glycan structures on serum PSA using energy resolved oxonium ion monitoring (Erexim) technology (*Anal Chem*, (2012) 84, 9655) (Patent US8653448, etc.) to improve the specificity and preclude false positive diagnoses of traditional PSA test.

The Erexim parameters of LCMS-8060 triple quadrupole mass spectrometer (Shimadzu) were optimized for quantify multiple glycan structures on PSA. In total 67 glycan structures on PSA, even from 0.1% content structures, were quantitatively monitored in 25 minutes run without enzymatic glycan release or chemical labeling.

As a result of analysis using sera from 15 prostate cancer or 15 benign prostate hyperplasia (BPH) patients whose PSA levels were in “gray zone” (4.0-10.0 ng/ml), abundance of multisialylated LacdiNAc (GalNAcβ1-4GlcNAc) structures were significantly upregulated in the prostate cancer group compared to the BPH group. A couple of these glycoforms were then extracted and subjected to establish a novel prostate cancer-specific diagnosis model (PSA G-Index®). When the diagnostic power was assessed using an independent validation sample set (15 PCa and 15 BPH patients in the PSA gray zone), an AUC of PSA G-index was 1.00, while that of total PSA or PSA f/T ratio was 0.50 or 0.60, respectively. Moreover, both PSA glycoforms showed significant correlation with Gleason scores (*Anal Chem*, (2019) 91, 2247).

The analytical basis of PSA G-Index® was already transferred to LSI Medience Corporation to realize the first mass spectrometric cancer diagnostics. Here we’d like to demonstrate intermediary results of the recently-launched larger-scaled validation study to evaluate clinical benefits of PSA G-Index® test.

891

Deciphering the biosynthesis of paucimannosidic proteins in human neutrophils

Julian Ugonotti[1], Harry C. Tjondro[1], Sayantani Chatterjee[1], Ian Loke[2], Yuqi Zhu[3], Zeynep Sumer-Bayraktar[4], Meral Topçu[5], Sriram Neelamegham[3], Morten Thaysen-Andersen[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Department of Biological Sciences, National University of Singapore, Singapore

3. Department of Chemical and Biological Engineering, University at Buffalo, Buffalo, New York, United States of America

4. School of Life and Environmental Sciences, Charles Perkins Centre, University of Sydney, Sydney, NSW, Australia

5. Department of Child Neurology, Hacettepe University, Ankara, Turkey

We have previously documented that human neutrophils are actively expressing immune-related glycoproteins carrying paucimannosidic glycans ($Man_{1-3}GlcNAc_2Fuc_{0-1}$), a functional yet under-represented class of *N*-glycosylation in the human glycobiological literature. Whilst inference from the paucimannose-rich invertebrates and plants points to a β-*N*-acetyl-hexosaminidase (Hex)-driven production of paucimannosidic proteins from immature β1,2-linked *N*-acetylglucosamine-terminating glycoprotein intermediates in the mammalian glycosylation machinery, this hypothesis remains experimentally unsupported for human neutrophils. We therefore sought to obtain evidence for the involvement of the Hex isoenzymes, which exist in homodimeric (αα and ββ) and heterodimeric (αβ) variants arising from two coding genes, the *HEXA* (α) and *HEXB* (β), in the biosynthesis of paucimannosidic proteins in human neutrophils. Our approach was to perform quantitative PGC-LC-MS/MS-based *N*-glycomics of several CRISPR-Cas9-edited Hex-deficient undifferentiated neutrophil-like HL-60 mutant cell lines (one *HEXA*-/- and three *HEXB*-/-) relative to unedited HL-60 control cells. Accurate disruption of the two target genes and absence of off-target genetic mutations were validated using next-generation sequencing and a conventional 4-methylumbelliferyl-(6-sulfo-)*N*-acetyl-β-glucosaminide-based Hex activity assay. Importantly, the *N*-glycome of all four Hex-deficient mutants displayed a significant reduction of paucimannosidic *N*-glycans (2.2-3.6%) relative to unedited HL-60 (13.8%, $p<0.05$). In particular, $Man_{2-3}GlcNAc_2Fuc_{0-1}$ displayed a significant reduction in the Hex-deficient mutants relative to unedited HL-60. Interestingly, an *N*-glycomics follow-up analysis of mature blood neutrophils from a patient diagnosed with early onset Sandhoff disease (*HEXB*-/-) recapitulated a reduced expression level of paucimannosidic *N*-glycans (20.7%) relative to levels found in mature blood neutrophils from a healthy age-matched donor (40.5%). This study provides evidence to support that several Hex isoenzyme variants encoded by both *HEXA* and *HEXB* are directly responsible for the biosynthesis of paucimannosidic proteins in human neutrophils. These well-characterised Hex-deficient mutant cell lines displaying low protein paucimannosylation may also be useful to further our understanding on the functional roles of paucimannosidic proteins in neutrophil glycoimmunology.

892

Calculating well-adjusted spectrum e-values using cloud approaches

Julian Uszkoreit[1], Dirk Winkelhardt[1], Katrin Marcus[1], Martin Eisenacher[1]

1. Ruhr University Bochum, Bochum, NRW, Germany

The currently most widely used approach to match MS/MS spectra to peptides is by using a database search engine. These algorithms take as input a protein database, which mostly is specific for only one species or a small group of taxa, and a set of MS/MS spectra. After the actual search engine's peptide identification, often a strategy containing the target-decoy-approach to estimate the false discovery rate (FDR) is applied. While this strategy worked well for many years, new high-resolution mass spectrometers with precursor and fragment mass errors in the lower ppm respective mmu range exhibit problems. Firstly, the essential decoys are no longer identified, as their theoretical mass spectra do not fit the measured data. With this effect, the traditional FDR estimation is no longer possible. Furthermore, almost all search engines perform well in distinguishing which given peptide matches a spectrum best. But the differentiation, whether the match of one spectrum is better than another spectrum's match, is often not possible when sing the algorithm's scores. Many search engines have for example a tendency to score heavier, longer peptides higher than lighter, smaller sequences.

To overcome both problems, we modified a compute-intensive strategy introduced almost five years ago, which becomes now feasible using cloud technology approaches. Instead of matching only the relatively few peptides in the precursor tolerance to each respective spectrum, we additionally match 1-10 thousand decoy peptides per spectrum, which are created to match the spectrum's tolerance. This amount of peptide spectrum matches per spectrum allows us to calculate well-calibrated e-values per spectrum, which are comparable between spectra and hopefully require no additional FDR estimation. As a side-effect of our strategy we can allow searches with very large databases - up to the complete UniProt KB - without exhibiting the FDR problems, which lead to lower sensitivity.

893

Targeted mass spectrometry quantification of IFN-γ induced ADP/ATP translocase 2 suppression related to cancer cell metabolism

Marie Vajrychova[1], Barbora Salovska[2], Ivo Fabrik[1], Miroslav Pribyl[2], Blanka Mrazkova[2], Vojtech Tambor[1], Zdenek Hodny[2]

1. Biomedical Research Centre, University Hospital Hradec Kralove, Hradec Kralove, Czech Republic

2. Department of Genome Integrity, Institute of Molecular Genetics of the ASCR, Prague, Czech Republic

Background: Cellular senescence represents a primary tumorigenesis barrier. On the other hand, senescent cells may participate on aging as well as cancer process development due to secretion of inflammatory cytokines like interferon (IFN) type I and II and TGF-β. We have recently shown that IFN-γ suppresses ADP/ATP translocase 2 (ANT2), which may result in elevation of reactive oxygen species, DNA damage and DNA damage response, persistent activation of cell cycle checkpoints and induction of senescent-like phenotype in cancer cell lines (1). Nevertheless, the mechanism of ANT2 growth-promoting role as well as detailed mechanism of its transcriptional suppression during exit from cell growth is unclear. Moreover, there is a lack of methods to accurately measure cellular concentration of different proteins from ADP/ATP translocases family. For illustration, it is not possible to distinguish between ANT1, ANT2, and ANT3 isoforms by Western blot. Thus, the aim of this work was to develop a method enabling us to quantitatively estimate levels of individual ATP/ADP translocators in mitochondria of normal and cancer cells from tissue samples and cell cultures undergoing various treatments.

Methodology: Firstly, unique ANT2 peptides to quantify protein level were carefully chosen and used for targeted quantification of ANT isoforms. The level of ANT2 peptide was determined using parallel reaction monitoring method upon normalization to synthetic internal ANT2 peptide standards.

Principal findings: Based on unique ANT2 peptides and synthetic peptide standards, we developed and optimized method for targeted quantification of ANT2. We applied this approach to assess ANT2 suppression in cellular senescence induced by IFN-γ.

Conclusions: Given the high similarity of amino acid composition of all human ANTs, we optimized a MS-based targeted method for quantification of ANT2 as an alternative to immunoblotting. Subsequently, we assessed the ratio of ADP/ATP translocators in normal and cancer cells on relative level.

1. Hubackova, S., et al. (2015) Oncogene 35, 1236-1249

894

Proteomic analysis of drug metabolising enzymes in extracellular vesicles

Madelé Dr van Dyk[1], Andrew A Prof Rowland[1]

1. Flinders University, Bedford Park, SOUTH AUSTRALIA, Australia

Introduction:

Variability in drug exposure as a result of variability in drug absorption, distribution, metabolism and excretion can be accounted for by understanding the enzyme activity and expression. Small extracellular vesicles (sEVs) are released into the bloodstream by organs, containing functional proteins and nucleic acids, and reflect the functional state of that organ. This study aims to quantify activity and expression of Cytochrome P450s (CYPs) and UDP-glucuronosyltransferases (UGTs) in sEVs derived from blood as a source for potential biomarkers.

Methods:

For peptide screening, in-gel trypsin digestion was performed. Peptides were separated by liquid chromatography (LC) with a 45 min acetonitrile gradient (BSciexEkspert400nanoHPLC). Column elutant was monitored by an AB Sciex 5600+ triple time of flight mass spectrometer (MS). De novo sequencing was performed on raw MS data (Peaks Studio v7.0 software).

Endogenous and labelled peptides were separated by LC (Agilent 1290 Infinity II HPLC) with a 17 min 0.1% formic acid in acetonitrile gradient. Column eluant was monitored by an Agilent 6495B Triple Quadrupole MS (ESI+ mode). Multiple reaction monitoring was performed with a single quantifier and two qualifier ion transitions. Endogenous peptide identities were confirmed by comparison of retention time, and quantifier/qualifier transition ratios of the respective labelled peptide standards.

Results:

188 unique peptides originating from CYP 1A2, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1, 2J2, 3A4 and 3A5, and UGT 1A1, 1A3, 1A4, 1A6, 1A9, 2B4, 2B7, 2B10 and 2B15 were detected. The number of unique peptides detected for each protein ranged between 2 and 19, with a mean of 9.65. By way of example, mean (range) CYP2D6 and CYP3A4 protein abundances in sEV were 192 (79 to 347) fmol/mL and 1094 (713 to 1523) fmol/mL, respectively.

Discussion

This study demonstrated the quantification of CYPs and UGTs in sEVS derived from blood which may be used as a potential source of clinical biomarkers. Additionally, it may complement existing drug probe-based approaches, while possibly circumventing the need for tissue biopsy.

895

Exercise induced muscle fatigue and damage- in search of a diagnostic biomarker

Jessica van Haeften[1], Daniel Broszczak[1], Tony Parker[1], Jonathan Peake[1]

1. Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, QLD, Australia

Athletes training at higher than typical intensities without appropriate preparation or adequate recovery periods may become overtrained. This debilitating condition is characterised by persistent fatigue and an inability to maintain high-performance athletic outputs. Proteins and metabolites harvested from body fluids during this process can be used to monitor the state of stress and recovery when high intensity training has caused exercise-induced fatigue, stress or injury. In addition, the identification of biomarkers indicating the onset of overtraining syndrome would allow for improved monitoring and determination of key points of intervention during athletic training to prevent long-term and career threatening performance decrements.

This discovery project will utilize quantitative protein and metabolite profiling to search for novel low abundant proteins in samples sourced from a study of highly trained athletes, pre- and post-high-intensity exercise [1]. The resulting –omics based data will be integrated and analysed through a data analysis pipeline to generate a more informative view of the underlying physiology. Results generated will be compared to a control cohort that was normally trained under the same conditions.

A differential abundance of biological markers may be detectable between excessively trained and normally trained athlete cohorts. This changed biomolecule profile will be indicative of an altered physiological state based upon effected biological pathways post-overtraining and better elucidate the mechanisms that underpin the processes of post-exercise muscle fatigue.

896

Proteins contained in outer membrane vesicles produced by *Rhizobium etli* grown in the presence of the *nod* gene inducer naringenin

María del Carmen Vargas-Lagunas[1], Hermenegildo Taboada[1], Michael Dunn[1], Andrés Andrade-Dominguez[2], Niurka Meneses[3], Sergio Encarnación-Guevara[1]

1. University Nacional Autónoma de México, Cuernavaca, MORELOS, Mexico

2. Centro Novatec, 4Noxgen Biotech, Atlatlahucan, Morelos, México

3. Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland

In this work, we compared the proteomic profiles of outer membrane vesicles (OMVs) isolated from *Rhizobium etli* CE3 grown in minimal medium (MM) with and without exogenous naringenin. One hundred seven proteins were present only in OMVs from naringenin-containing cultures (N-OMVs), 57 proteins were unique to OMVs from control cultures lacking naringenin (C-OMVs) and 303 proteins were present in OMVs from both culture conditions (S-OMVs). Although we found no absolute predominance of specific types of proteins in the N-, C- or S-OMV classes there were categories of proteins that were significantly less or more common in the different OMV categories. Proteins for energy production, translation and membrane and cell wall biogenesis were overrepresented in C-OMVs relative to N-OMVs. Proteins for carbohydrate metabolism and transport and those classified as either general function prediction only, function unknown, or without functional prediction were more common in N-OMVs than C-OMVs. This indicates that naringenin increased the proportion of these proteins in the OMVs, although NodD binding sites were only slightly more common in the promoters of genes for proteins found in the N-OMVs. In addition, OMVs from naringenin-containing cultures contained nodulation factor. Part of this work was supported by CONACyT grant 220790 and DGAPA-PAPIIT grants IN213216 and IN207519

897

The immunoglobulin superfamily receptome reveals novel functional and cancer-associated networks

erik verschueren[1], bushra hussain[1], kobe yuen[1], yi sun[2], sairupa paduchuri[1], yasin senbabaoglu[1], isabelle lehoux[1], tia arena[1], blair wilson[1], steve lianoglou[3], corey bakalarski[1], yvonne franke[1], pamela chan[1], athena wong[1], lino c gonzalez[4], sanjeev mariathasan[1], shannon j turley[1], jennie r lill[1], nadia martinez-martin[1]

1. Genentech, Inc., South San Francisco, CALIFORNIA, United States

2. University of Birmingham, Birmingham, United Kingdom

3. Denali, south san francisco, CA, United States

4. 23 and me, south san francisco, CA, United States

Cell surface receptors and their interactions fundamentally determine physiological and pathological signaling. Despite its clinical relevance, the Immunoglobulin Superfamily (IgSF) remains remarkably uncharacterized and many receptors are orphan. We present the first systematic extracellular protein map, the IgSF Interactome. Using a high throughput technology to interrogate most single transmembrane receptors for binding to 445 IgSF proteins, we identify over 500 interactions, 85% previously unreported, and confirm new interactions for immune checkpoints and orphan proteins using orthogonal assays. Our study reveals functionally related protein communities and the landscape of dysregulated receptor-ligand crosstalk in cancer, including selective loss-of-function for tumor-associated mutations. Investigation of the IgSF Interactome in a large cohort of cancer patients enrolled in a phase 2 clinical trial reveals expression signatures of interacting proteins associated with poor response to immunotherapy. The IgSF Interactome represents a unique resource to fuel biological discoveries and a framework for understanding the functional organization of the surfaceome during homeostasis and disease, ultimately informing therapeutic development.

898

SubCellBarCode: Proteome-wide Mapping of Protein Localization and Relocalization.

Mattias Vesterlund[1], Lukas M Orre[1], Yanbo Pan[1], Taner Arslan[1], Yafeng Zhu[1], Alejandro Fernandez-Woodbridge[1], Oliver Frings[1], Erik Fredlund[1], Janne Lehtiö[1]

1. Department of Oncology and Pathology, Karolinska Institutet, Science for Life Laboratory, Solna, STOCKHOLM, Sweden

Background: The architecture of eukaryotic cells is highly dynamic with protein complexes constantly being formed and resolved and proteins shuttling between different parts of the cell to carry out biological effects. Subcellular localization is one of the main regulatory levels of protein activity and protein interactions, however, a global view of cellular proteome organization remains relatively unexplored.

Purpose of the study: To develop a robust mass spectrometry based analysis pipeline to generate a proteome-wide view of subcellular localization.

Materials and Methods: We fractioned five different cell lines individually into five separate subcellular fractions and utilized high-resolution iso-electric focusing coupled to LC-MS/MS to generate in-depth proteomics data. Fractionation profiles were analyzed by tSNE and clustered by finite gaussian mixture models, finally, we used machine learning algorithms to classify proteins into subcellular compartments.

Results: Our analysis gives a highly accurate classification of proteins mapping to 12,418 individual genes into cytosolic, nuclear, secretory and mitochondrial locations and further into 15 specific subcellular compartments. We also investigated the impact of alternative splicing and protein domains on localization, complex member co-localization as well as protein relocalization after growth factor inhibition. Additionally, we have combined the method with peptide level phospho-enrichment which allows for identification of PTM-dependent localization.

Conclusion: Our analysis provides knowledge about the cellular architecture and the complexity of the spatial organization of the proteome, as we show that the majority of proteins have a single main subcellular localization, that alternative splicing rarely affects subcellular localization and that cell types are best distinguished by expression of proteins exposed to the surrounding environment.

899

Triglyceride-Based Nanodroplets and Their Interaction with Lipid-Specific Proteins

Valerija Vezočnik[1], Vesna Hodnik[1], Halil I. Okur[2], Simona Sitar[3], Magda Tušek-Žnidarič[4], Ksenija Kogej[1], Kristina Sepcic[1], Sylvie Roke[2], Ema Žagar[3], Peter Maček[1]

1. University of Ljubljana, Ljubljana, Slovenia

2. École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

3. National Institute of Chemistry, Ljubljana, Slovenia

4. National Institute of Biology, Ljubljana, Slovenia

Understanding of the interactions between proteins and natural and artificially prepared lipid membrane surfaces and embedded nonpolar cores is important in studies of physiological processes and their pathologies and is applicable to nanotechnologies. In particular, rapidly growing interest in cellular droplets defines the need for simplified biomimetic lipid model systems to overcome in vivo complexity and variability. In this study, we present a protocol for the preparation of kinetically stable nanoemulsions with nanodroplets composed of sphingomyelin (SM) and cholesterol (Chol), as amphiphilic surfactants, and trioleoylglycerol (TOG), at various molar ratios. To prepare stable SM/Chol-coated monodisperse lipid nanodroplets (LDs), we modified a reverse phase evaporation method and combined it with ultrasonication. Lipid composition, ζ-potential, gyration and hydrodynamic radius, shape, and temporal stability of the LDs were characterized and compared to extruded SM/Chol large unilamellar vesicles (LUVs). LDs and LUVs with theoretical SM/Chol/TOG molar ratios of 1/ 1/4.7 and 4/1/11.7 were further investigated for the orientational order of their interfacial water molecules using a second harmonic scattering technique, and for interactions with the SM-binding and Chol-binding pore-forming toxins equinatoxin II and perfringolysin O, respectively. The surface characteristics and binding of these proteins to the LDs SM/Chol monolayers were similar to those for the SM/Chol bilayers of the LUVs and SM/Chol Langmuir monolayers, in terms of their surface structures. We propose that such SM/ Chol/TOG nanoparticles with the required lipid compositions can serve as experimental models for monolayer membrane to provide a system that imitates the natural lipid droplets.

9

Multidimensional proteomic study identifies decreased protein synthesis and increased histone 2A ubiquitylation during aging

Lu Yang[1], Zaijun Ma[1], Nan Liu[1], Yaoyang Zhang[1]

1. Interdisciplinary Research Center on Biology and Chemistry, Chinese Academy of Sciences, Shanghai, SH, China

The long-lived proteome constitutes a pool of exceptionally stable proteins with limited turnover. Previous studies on ubiquitin-mediated protein degradation primarily focused on relatively short-lived proteins; how ubiquitylation modifies the long-lived proteome and its regulatory effect on adult lifespan is unclear. Here we profile the age-dependent dynamics of long-lived proteomes in *Drosophila* by mass spectrometry using stable isotope switching coupled with antibody-enriched ubiquitylome analysis. Our data describe landscapes of long-lived proteins in somatic and reproductive tissues of *Drosophila* during adult lifespan, and reveal a preferential ubiquitylation of older long-lived proteins. We further used pulsed metabolic stable isotope labeling analysis to characterize protein synthesis during aging. Interestingly, this study determines an age-modulated decline in protein synthesis with age, particularly in the pathways related to mitochondria, neurotransmission, and proteostasis. Importantly, we identify an age-modulated increase of ubiquitylation on long-lived histone 2A protein in *Drosophila*, which is evolutionarily

conserved in mouse, monkey, and human. A reduction of ubiquitylated histone 2A in mutant flies is associated with longevity and healthy lifespan. Together, our data reveal proteomic dynamics during aging and an evolutionarily conserved biomarker of aging that links epigenetic modulation of the long-lived histone protein to lifespan.

900

Evaluation of FAIMS technology for mass spec analysis of chemical cross-linked peptides.

Rosa Viner[1], Leigh Foster[2], Ryan Bomgarden[2], Michael W Belford[1], Romain Huguet[1], Eloy R Wouters[1]

1. Thermo Fisher Scientific, San Jose, CA, United States

2. Thermo Fisher Scientific, Rockford, IL, USA

Background

Chemical cross-linking in combination with mass spectrometry is a powerful method to determine protein-protein or nucleic acid-protein interactions. This method has been applied to recombinant and native protein complexes and, more recently, to whole cell lysates or intact unicellular organisms in efforts to identify protein-protein interactions on a global scale. However, this method suffers from low identification rates as the typical yield of cross-linked peptides is less than 1 % of total identified peptides. In this study, we compared widely used enrichment/fractionation techniques and the newly developed High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) device for cross-linked peptides.

Methods

Amine-reactive, homo-bifunctional crosslinkers, including DSS, DSSO, and DSBU were used to crosslink protein and protein complex standards. Cross-linked samples were reduced, alkylated and digested with trypsin for MS analysis. Cross-linked peptides were pre-fractionated using SCX. Samples were analyzed by LC-MS/MS on the Orbitrap mass spectrometers with/out the Thermo Scientific™ FAIMS Pro™ Interface. Data were analyzed using Thermo Scientific™ Proteome Discoverer™ 2.3 and XlinkX node 2.0.

Results

The identification of cross-linked peptides by LC-MS/MS presents significant analytical challenges due to their low abundance and higher charge state distribution compared to tryptic peptides. Selective enrichment/fractionation of cross-linked peptides by SCX fractionation using an offline LC approach is widely used for improved interaction sites identification; however, identification rates only increase by a maximum of 2-3 folds. To further enhance the detection of cross-linked peptides, we evaluated the use of a FAIMS device for gas-phase fractionation. Different compensation voltages (CV) between -40 and -90V were tested with 5V resolution. Analysis of the identical samples with a FAIMS device in place, using optimized methods, produced the same number of identified cross-linked peptides as after enrichment/fractionation.

Conclusion

FAIMS alone or in combination with enrichment/pre-fractionation significantly improved identification rates of cross-linked peptides in all samples.

901

An integrated atlas of protein expression in human cancer based on public proteomics data

Andrew F. Jarnuczak[1], Hanna Najgebauer[1], Mitra Barzine[1], Deepti J. Kundu[1], Fatemeh Ghavidel[2], Yasset Perez-Riverol[1], Irene Papatheodorou[1], Alvis Brazma[1], Juan Antonio Vizcaino[1]

1. EMBL-European Bioinformatics Institute, Hinxton, CAMBRIDGESHIRE, United Kingdom

2. University of Bergen, Bergen, Norway

DNA and RNA-based omics technologies have been successfully applied to profile primary tumours and their corresponding cell line models. While these studies often include hundreds of samples, proteomics studies are usually much smaller in scale. However, given the amount of MS-based datasets in the public domain, we can now employ *in-silico* analyses to integrate and reuse this valuable data. Here, we describe the generation of an integrated atlas of protein expression in human cancer using public data.

We collected and manually curated 7,171 raw files coming from 11 large-scale quantitative cancer proteomics datasets, most of them publicly available in the PRIDE database. The raw data was reprocessed in two combined analyses using MaxQuant. Quantification values were normalised through a multi-step procedure to obtain a complete integrated matrix of protein expression across all samples. Proteomics data was also integrated with public RNA expression information. All results are made available in the EMBL-EBI resources Expression Atlas and PRIDE.

Protein expression values were obtained for 191 cancer cell lines, 246 clinical tumour samples and 35 non-malignant tissues. These covered 13 different cancer lineages, including breast, colorectal, ovarian and prostate.

By exploring this integrated resource, we found that baseline protein expression in cell lines was generally representative of clinical tumour samples. However, as a key point, some differences in this overall trend were detected between cancer subtypes, as exemplified in the breast lineage. Furthermore, integration of RNA-seq data suggested that the level of transcriptional control in cell lines changed significantly depending on their lineage. Additionally, in agreement with previous studies, we found that variation in mRNA levels was generally a poor predictor of changes in protein abundance.

This work constitutes the first comprehensive meta-analysis including cancer cell lines and tumour proteomes, providing a highly valuable resource.

902

Comprehensive proteome analysis of the plant pathogen *Xanthomonas campestris* pv. campestris with focus on virulence

Julia Voß[1], Louise Schelletter[1], Karsten Niehaus[1], Vera Ortseifen[1]

1. Bielefeld University, Bielefeld, NRW, Germany

Xanthomonas campestris pv. campestris (*Xcc*) are Gram negative phytopathogens listed under the top five important phytopathogenic bacteria. They infect *Brassicaceae* and cause increasingly crop failure in agricultural. Additionally, the strain is of industrial interest due to its capability of producing the secondary metabolite xanthan which also plays an important role in phytopathology. In our study we worked with the wild type strain *Xcc* B100, of which the whole genome sequence was revealed in 2008 and refined 2017. Moreover, extracellular proteins as well as first analysis of the phospho proteome have been investigated. Till now a comprehensive proteome analysis of Xcc B100 is missing.

Proteome isolation was carried out using different isolation and nLC MS protocols to identify the largest possible number of proteins. With trifluoroethanol-separation the cytosolic proteome was isolated. Therefore, also a reversed phase fractionation was done. Isolation of the exoproteins was done with FASP-preparation. Additionally, the periplasmatic and membrane proteome was isolated. All samples were prepared in five biological replicates and measured on a nanoLC system coupled to a QExactive Plus. Data analysis was done with Proteome Discoverer ™.

In total more than 120 nLC MS measurements were done leading to a total number of 1,442,553 PSMs which can be associated 3,077 when minimal 2 unique peptides were used for identification. This presents 69% of whole putative proteome of *Xcc* B100. In this list many proteins located in clusters involved in virulence, like *hrp* or DFS could be identified.

With 3,077 detected proteins and therefore 69 % of the whole putative protein repertoire of Xcc B100, this work represents the most comprehensive proteome analysis in this strain to date. The identification of a variety of different proteins known to be involved in virulence could give a great groundwork for further analysis.

903

Cell membrane transporters facilitate the accumulation of drug-protein antigens, generating a localized pool of MHC epitopes associated with drug-induced liver injury

James C Waddington[1], Xiaoli Meng[1], Patricia T Illing[2], Arun A Tailor[1], Seart-E Ali[1], Jane Hamlett[1], John Farrell[1], Anthony Purcell[2], Dean J Naisbitt[1], B. Kevin Park[1]

1. MRC Centre for Drug Safety Science, University of Liverpool, Liverpool, UK

2. Infection and Immunity Program, Monash Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia

Introduction. Flucloxacillin is a β-lactam antibiotic associated with a high incidence of idiosyncratic drug-induced liver reactions. Although expression of HLA-B*57:01 increases susceptibility, little is known of the pathological mechanisms involved in the induction of the clinical phenotype. Irreversible protein modification is suspected to drive the reaction through the modification of peptides that are presented by the risk allele. In this study, the binding of flucloxacillin was characterized in terminally differentiated liver like HepaRG cells. Here, the membrane transporters involved in drug localisation were investigated. Additionally, due to the immune element of the onset of hypersensitivity, the precise nature of the MHC peptides presented by HLA-B*57:01 were identified.

Methods. Flucloxacillin protein binding was assessed in HepaRG and C1R-B*57:01 cells; B-lymphoblast cells transfected with HLA-B*57:01, using immunofluorescence microscopy. To identify drug modified MHC peptides presented by the risk allele, C1R-B*57:01 cells, used as a model for antigen presentation, were incubated with flucloxacillin for 48h. HLA peptide complexes were subsequently eluted and processed for mass spectrometric analysis.

Results. Direct modification of multiple proteins was observed, which could lead to neo-antigens being presented. Flucloxacillin was shown to disrupt transporter activity in HepaRG cells with localization appearing in bile canaliculi regions. The localization of flucloxacillin was likely mediated primarily by the membrane transporter MRP2. Of the peptides eluted from flucloxacillin treated C1R-B*57:01 cells, 6 peptides were fully annotated to show flucloxacillin-lysine covalent binding. Changes in the overall peptide repertoire were also observed.

Conclusions. We have demonstrated that localization of flucloxacillin occurs in the site of clinical disease during flucloxacillin-induced liver injury. We also demonstrated that neo-antigens, including drug-modified peptides, are presented by HLA-B*57:01. Further investigation into the immunogenicity of haptenated proteins and MHC peptides in the onset of iDILI is ongoing.

904

*i*HPDM: *in silico* human proteome digestion map with proteolytic peptide analysis and graphical visualizations

Jen-hung Wang[1], Wai-Kok Choong[1], Ching-Tai Chen[1], Ting-Yi Sung[1]

1. Institute of Information Science, Academia Sinica, Taipei, Taiwan

To detect missing proteins and protein isoforms in the human proteome, shotgun proteomics experiments are usually conducted, in which proteins are first proteolytically digested into peptides. It is desirable to use a protease that can yield more unique peptides with properties amenable for mass spectrometry analysis. Though trypsin is currently the most widely used protease, some proteins can yield only a limited number of unique peptides by trypsin digestion due to some hindering effect on cleavage activity. For example, trypsin does not cleave at a lysine or arginine residue when a negatively charged amino acid, such as aspartic acid and glutamic acid, or phosphorylated serine or threonine, is located nearby. Moreover, it has been reported that 19 types of PTMs occurring at lysine may hinder trypsin cleavage efficiency.

Other protease and multiple proteases have been applied in reported studies to increase the number of identified proteins and protein sequence coverage. To facilitate the selection of protease, we have developed a web server, called *in silico* Human Proteome Digestion Map (*i*HPDM), which contains a comprehensive proteolytic peptide database constructed from human proteins including isoforms in neXtProt digested by 15 protease combinations of single or two proteases. *i*HPDM provides convenient functions, including *Protein Query*, *Multi-protease Comparison* and *Isoform Digestion*, and graphical visualizations for users to examine and compare the digestion results of different proteases. Notably, it also supports users to input filtering criteria on digested peptides, e.g., peptide length and uniqueness, in consideration of MS detectability to select suitable proteases.

In summary, with the capability to compare and examine different combinations of proteases for protein digestion in MS experiments, *i*HPDM can facilitate the selection of proteases for shotgun proteomics experiments to identify missing proteins, protein isoforms and single amino acid variant peptides.

905

SIRT5 contributes to colorectal cancer growth by regulating T cell activity

KE WANG[1], CUIPING ZHANG[1], LUJIE YANG[1], PENGYUAN YANG[1], HONGXIU YU[1]

1. Institutes of Biomedical Sciences & Department of Systems Biology for Medicine, School of Basic Medical Sciences, Fudan University, Shanghai, SHANGHAI, China

Over the past several years, SIRT5 has attracted considerable attention in metabolic regulation. However, the function of SIRT5 in immune cells is poorly understood. We discovered that proteome and network analysis revealed SIRT5 was important in T cell receptor signaling pathways. Additionally, we determined that a deficiency in SIRT5 induced T cell activation. Furthermore, our results demonstrated that SIRT5 played a pivotal role in regulating the differentiation of regulatory T (T_{reg}) cells and T-helper 1 (Th1) cells. An imbalance in the lineages of immunosuppressive T_{reg} cells and the inflammatory Th1 subsets of helper T cells leads to the development of inflammatory diseases and cancer. Furthermore, we found that *Sirt5* knockout mice were resistant to AOM and DSS-induced colitis-associated colorectal tumorigenesis.

906

A time-resolved multi-omic atlas of the developing mouse liver

Tongqing Gong[1], Chunchao Zhang[2], Xiaotian Ni[1, 3], Xianju Li[1], Chen Ding[4], Jun Qin[1, 4], Yi Wang[1, 2]

1. State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Changping District, BEIJING, China

2. Alkek Center for Molecular Discovery, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX, USA

3. Department of Life Sciences, East China Normal University, Shanghai, China

4. State Key Laboratory of Genetic Engineering, Human Phenome Institute, Institute of Biomedical Sciences, and School of Life Sciences, Fudan University, Shanghai, China

Liver organogenesis and development is composed of a series of complex, well-orchestrated events. Identification of key factors and pathways governing liver development will help understand this process and may also provide insights for other physiological and pathological processes including cancer. For this purpose, we conducted multi-dimensional omics measurements including profiling of protein, mRNA, and transcription factor (TF) DNA binding activity from mouse liver tissues collected at embryonic day E12.5 to postnatal week W8, encompassing the major developmental stages to provide a molecular and integrative panoramic view of mouse liver development. These datasets revealed dynamic changes of core liver functions as well as canonical signaling pathways governing development at both mRNA and protein levels, and identified novel RNA splicing variants that are confirmed at the peptide level. The TF DNA binding activity dataset provided the first glimpse of liver development from four waves of TF activations and major TFs that may be responsible to drive the transcriptional program to govern liver development. A comparison between mouse liver development and human hepatocellular carcinoma (HCC) proteomic profiles revealed that more aggressive tumors are characterized with the activation of early embryonic development pathways, whereas less aggressive ones maintain liver-function related pathways that are elevated in the mature liver. This work provided a rich resource for liver development research community for future in-depth functional characterization.

907

Understanding the information content in proteomic profiles of late pleistocene to iron age humans from the tropics

Valerie C Wasinger[1, 2], Darren Curnoe[3, 2], Ray Mendoza[2], Lewis Adler[1], Andy Baker[2]

1. Mark Wainwright Analytical Centre, The University of NSW, Kensington, NSW, Australia

2. Palaeontology, Geobiology and Earth Archives Research Centre, The University of NSW, Sydney, NSW, Australia

3. ARC Centre of Excellence for Australian Biodiversity and Heritage, The University of NSW, Sydney, NSW, Australia

Recovery of proteins preserved over time is challenging. Here we present the challenges unique to extracting information from samples preserved in tropical humid environments from ~ 2000 years B.P (iron-age teeth sample) collected from Long Long Rak rock shelter in Northwest Thailand; and from two ~35, 000 years B.P (late Pleistocene skull and mandible sample) from anatomically modern humans inhabiting West Mouth of Niah Caves in Borneo.

Evidence for lifestyles in challenging environments such as tropical rain forests is essential for understanding the adaptations of early anatomically modern humans. Tropical rain forests are considered difficult surroundings to inhabit long term and require biological and cultural adaptations to survive. The uniquely crystalline structures of these samples has preserved sufficient proteomic content to explore protein pathways, potential illness and the association of diets and lifestyle consequences.

We demonstrate that human proteomes can be successfully recovered from individuals inhabiting a tropical setting extending well into the Late Pleistocene. We also show the two important amelogenin protein isoforms allowing the determination of sex from preserved teeth samples. We describe a novel MRM method using the isoforms of both X-related and Y-related forms of the amelogenin protein. Together, these different samples provide valuable insight into societies living in pre-history.

908

Network analysis of proteogenomics data in lung cancer cell lines

Yu Watanabe[1], Yiwei Ling[1], Hayato Uemura[2], Akiyasu C Yoshizawa[2], Yasushi Ishihama[2], Shujiro Okuda[1]

1. Niigata University, Niigata, NIIGATA, Japan

2. Kyoto University, Kyoto, KYOTO, Japan

Background

Molecular actions derived from gene mutations are still unclear although many of large scale cancer genomics datasets have been produced and analyzed recently. Proteogenomics is necessary to elucidate the dynamics of such gene mutations and the resulting protein expression. Therefore, it is important to establish the analysis method for the dynamics of protein expression caused by gene mutation.

Methods

We obtained protein-protein interaction data from STRING (https://string-db.org/) and loaded them into the Neo4j (https://neo4j.com/) graph database. Subsequently, we calculated all shortest paths between each protein node pair. We also obtained gene mutation information from lung adenocarcinoma cell lines stored in DBKERO (http://kero.hgc.jp/). Then, we measured the proteome expression of the cell lines with LC/MS/MS. Combining both data, we mapped up/down-regulated proteins on the shortest paths and extracted linearly up/down-regulated paths which all nodes are simultaneously co-regulated. Finally, we combined extracted paths to draw whole co-regulated network and traversed the network started from a mutated gene. As a result, continuously co-regulated pathways started from a mutated gene could be obtained.

Results

We successfully obtained all continuously co-regulated protein pathways linked to mutated genes for each cell line sample. These pathways have different characteristics for each gene mutation subtype.

Conclusions

Our novel network analysis method is helpful for interpretation of proteogenomic data and it makes possible to characterize cancer subtypes. Our method can explore novel causative routes by gene mutations that it is difficult to find from the existing knowledge-based pathway databases.

909

Phosphoproteome profiling of isogenic cancer cell-derived exosome reveals HSP90 as a potential marker for human cholangiocarcinoma

Churat Weeraphan[1], Amornrat Pongdara[3, 2], Parunya Chaiyawat[4], Penchatr Diskul-Na-Ayudthaya[5], Daranee Chokchaichamnankit[6], Chris Verathamjamras[6], Pukkavadee Netsirisawan[6], Yodying Yingchutrakul[7], Sittiruk Roytrakul[7], Voraratt Champattanachai[6], Jisnuson Svasti[4, 6], Chantragan Srisomsap[6]

1. Department of Molecular Biotechnology and Bioinformatics, Faculty of Science, Prince of Songkla University, Songkhla, Thailand

2. Faculty of Medical Technology, Prince of Songkla University, Songkhla , Thailand

3. Center for Genomics and Bioinformatics Research, Faculty of Science, Prince of Songkla University, Songkhla , Thailand

4. Applied Biological Sciences Program, Chulabhorn Graduate Institute, Bangkok, Thailand

5. Soonchunhyang Institute of Medi-bio Science (SIMS) , Soonchunhyang University, Chungcheongnam-do, Republic of Korea

6. Laboratory of Biochemistry, Chulabhorn Research Institute, Bangkok , Thailand

7. Proteomics Research Laboratory, National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency, Pathum Thani, Thailand

The Northeastern region of Thailand is well known to have a high incidence and mortality of cholangiocarcinoma. Protein phosphorylation status has been reported to reflect a key determinant of cellular physiology, but identification of phosphoproteins can be a problem due to the presence of phosphatase. Growing evidence indicates that exosomes are stable towards circulating proteases and other enzymes in human blood and can be recognized before the onset of cancer progression.Here an *in vitro*metastatic model of isogenic cholangiocarcinoma cells was used to provide insight into the phosphorylation levels of exosomal proteins derived from highly invasive cells. Gel-based and gel-free proteomics approaches were used to reveal the proteins differentially phosphorylated in relation to tumor cell phenotypes. Forty-three phosphoproteins were identified with a significant change in phosphorylation level. Phos-tag western blotting and Immunohistochemistry staining was then employed to validate the candidate phosphoproteins. Heat shock protein 90 was successfully confirmed as being differentially phosphorylated in relation to tumor malignancy. Importantly, the aberrant phosphorylation of exosomal proteins might serve as a promising tool for the development of a biomarker for metastatic cholangiocarcinoma.

1. Weeraphan C, Phongdara A, Chaiyawat P, Diskul-Na-Ayudthaya P, Chokchaichamnankit D, Verathamjamras C, et al. Phosphoproteome Profiling of Isogenic Cancer Cell-Derived Exosome Reveals HSP90 as a Potential Marker for Human Cholangiocarcinoma. Proteomics. 2019
2. Chen IH, Xue L, Hsu CC, Paez JS, Pan L, Andaluz H, et al. Phosphoproteins in extracellular vesicles as candidate markers for breast cancer. Proc Natl Acad Sci U S A. 2017;114(12):3175-80.

910

Development of a highly sensitive lectin suspension microarray system for profiling glycans in serum glycopeptidome from patients with intrahepatic cholangiocarcinoma

Liming Wei[1], Hong Li[1], Lei Zhang[1], Guoquan Yan [1], Pengyuan Yang[1]

1. Fudan University, Shanghai, SHANGHAI, China

Glycomics is the study of comprehensive structural elucidation and characterization of all glycoforms found in nature and further their dynamic spatiotemporal changes are associated with biological processes [1]. Relative to genomics and proteomics, glycomics is just growing out of infancy with great potential in biomedicine for biomarker discovery, diagnosis, and treatment. Lectin microarrays are increasingly used as versatile platform with a high potential for rapid and high throughput glycosylation profiling of biology sample, without the need for the release of glycan. However, weak interaction between lectin and glycan causes low sensitivity of the approach. Here we firstly conjugated lectin on the surface of microspheres assisted by poly-L-lysine (PL) and PEG-diglycolic acid (PEG). The introduction of PL and PEG chain could form a hydrophilic shell to enhance the dispersibility of suspension microarray in aqueous solution and reduce the steric hindrance among the bound lectin for high conjugated amount of lectins. In this report, we present the performance of novel high density lectin suspension microarray in the analysis of standards glycoproteins such as asialofetuin (ASF), fetuin (FET), HRP, ribonuclease B (RNB). Compared with the direct conjugation of lectin suspension microarray, the novel platform shows better dispersibility with 10% (CV) and three to five more sensitivity than first generation suspension microarray [2]. With the high sensitivity and good dispersibility, the novel suspension microarray was applied for serum glycopeptidome analysis from intrahepatic cholangiocarcinoma (ICC) patients. Serum from normal (n=20) and ICC partients (n=20) was analyzed and found that the signature glycan expression in ICC, including α-Fuc, α-Glc, etc. The signature can be explored further as potential biomarker for discovery, diagnosis, and prognostic evaluation of ICC.

[1] Nat. Rev. Cancer 2015, 15(9), 540-555.

[2] Proteomics 2014, 14, 78–86.

911

Non-enzymatic oxidative post-translational modifications in peptides – Key to understand the functional consequences of cold atmospheric plasma (CAP)

Sebastian Wenske[1], Jan-Wilm Lackmann[1], Thomas von Woedtke[1], Kristian Wende[1]

1. Leibniz Institute for Plasma Science and Technology, Greifswald, GERMANY, Germany

Cold atmospheric plasma (CAP) sources can be used to produce a multitude of different reactive species in a spatially and time resolved manner leading to a wide variety of modifications on nearly all amino acids [1]. CAP thus has the ability to alter the activity of proteins or enzymes in cellular signalling pathways, which could be one mechanism in the treatment of chronic wounds that is strongly modulated by reactive species signalling [2].

The aim of this study was to determine which oxidative PTMs on three model peptides are generated by the treatment with two different plasma sources (kINPen [3] and COST jet [4]). Nano-LC/High resolution mass-spectrometry was performed and data were analysed by a software algorithm allowing to search for various PTMs simultaneously. Significant differences in the modification pattern were observed. While the COST jet modified predominantly aromatic amino acids (AA), the kINPen led also to the oxidation of aliphatic AA and showed a substantial nitration of aromatics. The nitration of such an aromatic AA like tyrosine in the enzyme aldolase A can lead to a regulation of mast cell degranulation, an important regulator in wound healing processes [5]. Additionally, the chemical environment of a given AA determined its reactivity. Trioxidations and the modification 3O-NHCO [6] can only be found on histidine during direct argon treatments. Phenylalanine in the middle of a peptide shows massively increased oxidations and dioxidations compared to another phenylalanine, which is closer to the C-terminus.

Ten further peptides were designed in a way that each AA occurs with various neighbouring AA providing a different chemical environment found to modulate the observable modifications. Thus, a comprehensive library of plasma-derived oxidative modifications will be created that consequently will be used to predict and identify modifications of proteins in complex settings such as chronic wounds under oxidant conditions.

1. [1] E. Takai, T. Kitamura, J. Kuwabara, S. Ikawa, S. Yoshizawa, K. Shiraki, H. Kawasaki, R. Arakawaw, K. Kitano, Journal of Physics D: Applied Physics, 2014
2. [2] A. Schmidt, S. Bekeschus, Antioxidants, 2018
3. [3] S. Reuter, T. von Woedtke, K.-D. Weltmann, J. Phys. D: Appl. Phys., 2018
4. [4] J. Golda et al, J. Phys. D: Appl. Phys., 2016
5. [5] Y. Sekar, T.C. Moon, C.M. Slupsky, A.D. Befus, J. Immunol., 2010
6. [6] E. Takai et al, J. Phys. D: Appl. Phys., 2014

912

N-linked glycan branching and fucosylation are increased directly in hepatocellular carcinoma tissue as determined through in situ glycan imaging and are associated with poor outcome

Connor A West[1], Mengjun Wang[1], Richard R Drake[1], Anand S Mehta[1]

1. Medical University of South Carolina, Charleston, SOUTH CAROLINA, United States

Hepatocellular carcinoma (HCC) remains the fifth most common cancer worldwide and accounts for approximately 700,000 deaths annually. Changes in serum glycosylation have long been associated with HCC but the source material is unknown and

direct glycan analysis of HCC tissues has been limited. Our laboratory previously developed a method of *in situ* tissue based N-linked glycan imaging that bypasses the need for microdissection and solubilization of tissue prior to analysis. We used this methodology in the analysis of 138 HCC tissue samples and compared the N-linked glycans in cancer tissue with either adjacent untransformed or cirrhotic tissue. Post analysis, there were over 75 N-linked glycans found to be associated with HCC. Ten glycans were found significantly elevated in HCC tissues compared to cirrhotic or adjacent tissue and fell into two major classes: increased fucosylation and increased branching. In analysis of 90 tissue pairs (HCC and adjacent un-transformed tissue), these classes were associated with poor outcome with high levels of branched and singly or doubly fucosylated glycans resulted in a median survival time of 25 months, as compared to patients with low levels having a median survival time of 35 months and 32 months respectively. When patients with early HCC were examined, the five-year survival rate of those with high levels of the tetra-antennary glycan with a single fucose residue was 23% less than those with low levels, and a more striking difference of 40% less when looking at two-year survival times. These 90 tissue pairs were then evaluated for core versus outer arm fucosylation via Endoglycosidase F3 imaging to determine the relevance of fucosylation linkages on survival and outcome data while maintaining spatial localization. This work marks one of the first spatially localized glycomic evaluations of HCC tissue directly as compared to un-transformed adjacent tissue with included survival outcomes.

913

Samples to features to enrichment - Rapid and interactive visualization of proteogenomics data

Jakob Willforss[1], Erik Andreasson[2], Aakash Chawade[3], Fredrik Levander[1]

1. Department of Immunotechnology, Lund University, Lund, Sweden

2. Plant Protection, Swedish University of Agricultural Sciences, Alnarp, Sweden

3. Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

Proteogenomics studies involve multiple layers of omics data, from samples to features to sets of features such as functional enrichments. If not analyzed correctly, there is a risk of missing out on biological insights or overlook technical limitations. We developed nOmics with the intent of enabling rapid exploration across different data layers.

The software is developed in RShiny and leverages the Bioconductor ecosystem for its visualizations. It was used for exploration of a proteogenomics dataset in oat where the protein expression of two oat varieties infected by *Fusarium graminearum* was compared using transcriptome-based references to identify features distinguishing their infection response.

nOmics allows inspecting the data both globally using high-level interactive illustrations such as PCA, clustering and histograms and locally by zooming in on individual proteins to examine their expression patterns and sequence variations across references. Subsets of the data can easily be sliced out and used for calculating gene ontology enrichment and for identifying proteins of interest. nOmics was employed in the oat dataset where it was successfully demonstrated to reveal and handle technical anomalies and to identify both enriched gene ontology terms and proteins potentially linked to differing infection response across oat varieties.

The study demonstrates how nOmics allows for rapid inspection of multiple layers of proteogenomics data. It provides a nuanced overview of omics datasets, allowing for pinpointing of technical issues and rapid examination of functional relations.

914

A fast and universal sample preparation proteomic workflow for cancer biopsies sourced from different embedding methods

Steven G Williams[1], Asim Anees[1], Srikanth Manda[1], Jennifer M S Koh[1], Clare Loudon[1], Belinda Serafin [1], Brett Tully[1], Roger Reddel[1], Qing Zhong[1], Peter G Hains[1], Phil J Robinson[1]

1. The Childrens Medical Research Institute, Westmead, NSW, Australia

The main solid samples available for proteomics of cancer biopsy tissues include: Fresh-frozen (FF), OCT polymer embedded FF (FF-OCT), or formalin-fixed paraffin embedded (FFPE). Sample preparation prior to mass spectrometry (MS) requires different approaches for each. First is the time-consuming washing steps to remove OCT or paraffin. Next the tissues are lysed and proteolytically digested, and finally impurity removal by solid phase extraction (SPE). Different SPE approaches may be required for FF-OCT or FFPE since standard SPE on C18 resin does not remove polymers. In contrast, tissue lysis and digestion can be performed in common in a Barocycler.

We aimed to produce a near-universal sample preparation workflow across tissue sources. Jennifer Koh (see HUPO 2019 poster) developed an SPE method on mixed cation exchange (MCX) resin that removes OCT, allowing elimination of FF-OCT sample washing. Our first aim was to reduce the washing steps for FFPE samples. We developed a 10 min procedure for de-waxing in heated heptane then methanol that is sufficient to remove the wax, reducing washing steps by hours. Reversal of cross-linking was performed during a standard Barocycler digestion. This effectively eliminates sample washing for all typical cancer solid biopsies.

We next determined the optimum SPE method. C18 cannot be used with OCT, but both MCX and silica resins are alternatives for peptide clean-up. With all three sample sources and five rat organs representing a broad tissue density range, we found that all three SPE methods performed well. Both MCX and silica removed OCT, while all three cleaned up FF and FFPE samples. Each method was tested for reproducibility in multiple replicates.

The results show near-universal sample washing and peptide clean-up procedures, that are robust and suitable for any solid tissue source regardless of tissue fixation method.

915

SiTrap: rapid and high throughput clinical multiomics

John P Wilson[1], Alexandre Zougman[2], Darryl J Pappin[3]

1. ProtiFi, LLC, Farmingdale, NEW YORK, United States

2. University of Leeds, Leeds, UK

3. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States

Sample processing in omics analyses is obligatory and has traditionally been the largest source of variability in proteomics and metabolomics data. Resultantly, many biomarker “discoveries” are subsequently traced to batch or run order effects. Such errors unfortunately continue to the present day, thus compelling for clinical deployment of omics technologies first serious improvements in reproducibility especially in sample processing. Second, for widespread use, sample throughput should be as high as possible, necessitating extension to an automated platform. Moreover, to enable eventual point of care (POC) deployment of omics techniques, sample processing time should be as rapid as possible. Third, as we do not necessarily have a priori knowledge of which biomolecule will prove to be useful as diagnostic, prognostic, therapeutic or predictive signals for a given biological condition, sample processing should ideally produce from one sample different fractions for multiple omics interrogations. Fourth, such a solution should be affordable enough to be accessible to the majority of labs.

Here, we present SiTrap, a new high-throughput solution to address all necessities for clinical implementation of omics analyses. A new development over the original S-Trap[1-4] sample processing platform, SiTrap yields from the same sample both metabolomic and proteomic fractions with extraordinary inter- and intrarun reproducibility over time. It requires only minutes of processing time per sample and can be executed either in loose spin columns or in an automated 96-well format (4 - 96 samples simultaneously). Importantly, automation and its commensurate high reproducibility can be achieved on an inexpensive Tecan A200 positive pressure workstation, a general automation platform accessible – due to its low cost – to the majority of research and clinical labs. We anticipate that SiTrap will become an essential omics tool in laboratory and clinical settings and will enable novel discoveries, thereby helping to usher in a new era of clinical proteomics.

1. Zougman A, Selby PJ, Banks RE. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. Proteomics. 2014 May;14(9):1006-0.
2. Ludwig KR, Schroll MM, Hummon AB. Comparison of in-solution, FASP, and S-Trap based digestion methods for bottom-up proteomic studies. Journal of proteome research. 2018 May 13;17(7):2480-90.
3. HaileMariam M, Eguez RV, Singh H, Bekele S, Ameni G, Pieper R, Yu Y. S-Trap, an Ultrafast Sample-Preparation Approach for Shotgun Proteomics. Journal of proteome research. 2018 Aug 16;17(9):2917-24.
4. Elinger D, Gabashvili A, Levin Y. Suspension trapping (S-trap) is compatible with typical protein extraction buffers and detergents for bottom up proteomics. Journal of proteome research. 2019 Mar 1.

916

Determination of total plasma Coenzyme Q10 by liquid chromatography-tandem mass spectrometry to help diagnosis to possible mitochondrial disorder.

Yi Ling Wong[1], Wai Yuen Yu[2], Chun Yiu Law[2], Ching Wan Lam[2, 1]

1. Department of Pathology, University of Hong Kong, Hong Kong, China

2. Division of Chemical Pathology, Queen Mary Hospital, Hong Kong, China

Background: Coenzyme Q10 (CoQ_{10}), also named ubiquinone is an essential cofactor in the mitochondrial respiratory chain, which is responsible for oxidative phosphorylation. It acts as an electron carrier and an antioxidant in the human body. Endogenous synthesis of CoQ_{10} is the major source of CoQ_{10} in human and > 12 genes had been identified in the biosynthetic pathway. Defective CoQ_{10} biosynthesis could lead to primary CoQ_{10} deficiency and mitochondrial disorder. In this study, we hypothesize plasma CoQ_{10} is a biomarker in mitochondrial disorder due to defective CoQ_{10} synthesis. Since CoQ_{10} measurement is not universally available, a simple and robust liquid-chromatography tandem mass spectrometry (LC-MS/MS)-based method was introduced and described in this work.

Methods: Test was performed using AB SCIEX 5500 QTRAP coupled with ExionLC™ System. Fresh plasma samples were extracted by protein precipitation using ethanol which included coenzyme Q_{10}-[D9] as an internal standard. 1, 4-benzoquinone was added as an oxidizing agent to convert all reduced form of CoQ_{10} into the oxidized form for the measurement of total CoQ_{10}. The ammonium adduct ion with MRM transition of 881.0/197.0 was used for quantification.

Results: Our assay showed a linear range of 5-2000 µg/L with R-squared greater than 0.99. The limit of quantification is 10 µg/L in plasma. Accuracy was checked by using commercial plasma QC controls in three different concentrations with less than 5% deviation from target value. The Intra- and inter-assay coefficients of variation were less than 10%. There was no matrix suppression or enhancement in the elution time of CoQ_{10}. The reference range was verified to be 320-1,558 µg/L for age group less than 18 years old and be 433-1532 µg/L for those greater than or equal to 18 years old. A marked reduction in plasma CoQ_{10} was noted in patients with defective CoQ_{10} biosynthesis due to *COQ4* defect who presented with mitochondriopathy.

Conclusions: This is an efficient and a selective method for determination of CoQ_{10} concentrations in human plasma for helping diagnosis of CoQ_{10} deficiency in mitochondrial disorders and monitor its status during treatment.

1. 1) Miles MV, Horn PS, Tang PH, et al, Clin Chim Acta 2004, 34, 139-144; 2) Claessens AJ, Yeung CK, Risler LJ, Phillips BR, Himmelfarb J, Shen DD, Ann Clin Biochem., 53, 265-273; 3) Quinzii CM, DiMauro S, Hirano M., Neurochem Res., 32, 723-727

917

Characteristics of MALDI-imaging on a new dual ion source QTOF with TIMS separation

Arne Fuetterer[1], Juergen Suetering[1], Janina Oetjen[1], Nils Goedecke[1], Stephanie Kaspar-Shoenenfeld[1], Scarlet Koch[1], Pierre-Olivier Schmit[2], Shannon Cornett[3], Alice Ly[1], Lucy Woods[1], jens Fuchser[1], Oliver Raether[1], jens Hoendorf[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Bruker France, Wissembourg, France

3. Bruker Daltonics Inc, Billerica, MA, USA

Background

MALDI Imaging has a broad range of applications in Omics research. However, a gap exists between desired mass resolution capabilities and the acquisition speed of current instrumentation. We present initial results from the timsTOF flex system; a timsTOF Pro QTOF mounted with a high-throughput, high spatial resolution MALDI source and stage.

Methods

A timsTOF Pro was mounted with a MALDI source and 10 kHz smartBeam 3D laser featuring electronically controlled spot positioning and beam profile for imaging. Performance in ESI mode was evaluated by analyzing a commercially available HeLa digest (Pierce) using DDA PASEF approach. Tissue samples for MALDI-MSI were mounted on conductive glass slides, and coated with matrix using standard protocols on a TM Sprayer (HTX Technologies, Chapel Hill, NC, USA). MALDI ion mobility imaging experiments were acquired on the system at a mobility resolution of 150 1/K0.

Findings

High spectral quality MALDI Imaging data could be acquired at a rate of up to 20 pixels/second in both positive and negative mode. A sagittal rat brain section consisting of approximately 370,000 pixels took ~5 hours to measure. Spatial resolution of 20 µm was confirmed by matching ion signals to specific cells and structures in rat brain. In experiments designed to stress the system, 20 hours of image acquisition or ~1.5 million pixels showed no decline in imaging dataset quality and a mass deviation of RMS 2.06 without lock mass. Trapped ion mobility imaging measurements removed isobaric interferences in lipid imaging.

Proteomics analysis s was used to assess if the dual source design and MALDI Imaging experiments affected LC-MS/MS performance. Injections of 200ng HeLa revealed over 5000 protein groups identified; this figure is maintained over the course of measuring 20 million MALDI pixels.

Conclusion:

The timsTOF fleX allows for fast, high-spatial resolution MALDI acquisition, and robust ESI performance.

918

High throughput plasma proteomics with PASEF and 4D feature alignment

Thomas Kosinski[1], Raphael Heilig[2], Nicolai Bache[3], Ole Bjeld Horning[3], Pierre-Olivier Schmit[4], Lucy Woods[1], Roman Fisher[2], Heiner Koch[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

3. Evosep BioSystems, Odense, Denmark

4. Bruker France, Wissembourg, France

Background

Blood analysis is one of the most commonly performed procedures in medicine, where clinical parameters are used for diagnosis and for decision on treatment options. If LC-MS/MS based proteomics has long been a powerful research tool but has not provided the robustness and throughput to decipher new biomarkers in large cohort studies of blood plasma. Here, we have combined the robustness and speed of a tims-Q-TOF instrument operated in PASEF mode with PASEF together with the a robust Hight-Trouhghput EvoSep One LC system and a 4D feature alignment-based processing

Methods

192 blood plasma samples from patients with a severe infection were digested with trypsin after depletion of the 12 most abundant proteins. 100ng of each digest were separated on an EvosepOne system at a sampling rate of 100 runs per day (11.5 min gradient). After every 10th we include QC sample. LC-MS/MS measurements were performed on a timsTOF Pro system using a reduced cycle time (0,5 seconds instead of 1.1) PASEF method adapted to short gradients. Post processing was performed with PEAKS studio (version X, Bioinformatics Solutions Inc.). Search results were corrected to 1% PSM FDR and quantification was performed based on MS1 feature intensities. A match of feature vectors in retention time, m/z, 1/K0 was applied to reduce missing values and to transfer protein identifications between runs.

Findings

We found high quantitative reproducibility across the study (R2 > 0.97, median CV 9.3%). Using the 4D feature alignment-based processing allowed to improve the number of quantified proteins from 188 to 500 protein groups on average per run. In the sample set we found several proteins of intermediate or lower abundance (CRP, PSA, IFN-γ)

Conclusion

The combination of these latests technologies offers a robust and high throughput LC-MS/MS solution for plasma screening.

919

Sexual dimorphism and epileptogenesis in developing and adult rat brains

Merridee A Wouters[1], Jonas Albinas[1], Madeleine J Otway[1], Peter G Hains[1], Martin Larsen[2], Pablo Espinosa[1], Terence O'Brien[3], Nigel Jones[3], Qing Zhong[1], Phil J Robinson[1]

1. Children's Medical Research Institute, The University of Sydney, Westmead

2. University of Southern Denmark, Odense

3. Monash University, Clayfield

Absence epilepsy (AE) develops in the somatosensory cortex in children, with a bias towards females. Previous work has shown elevation of GABA during seizures, and changed expression of actin, profilin 1 and α-synuclein. We performed a dual study of gender differences and epileptogenesis in developing rat brains by comparing two rat models of each gender for genetic absence epilepsy (GAERs, WAG/Rij) to non-epileptic control (NEC) rats using mass spectrometry-based proteomics. The GAERs have a deficit in Cacng2, a T-type voltage-dependent calcium channel. The WAG/Rij rats, have hyperpolarized HCN channels.

The study design consisted of four rats of each sex from the three rat types, studied over 5 timepoints: at weaning (3 weeks), pre-puberty (6 weeks), post-puberty (10 weeks) and young adults (16 & 24 weeks). Although both rat models display AE, the age of onset differs. GAERs have regular seizures from the onset of puberty at 10 weeks while WAG/Rij rats typically begin around 16 weeks. Post-mortem samples from the somatosensory cortex were analysed with SWATH-mass spectrometry, a new type of data-independent acquisition (DIA) methodology. A generalized linear model was fitted.

Independent of the AE, the data revealed large differences in the proteome of the somatosensory cortex of male rats compared to females. AE changes were more restricted, with a small, specific set of proteins altered in the somatosensory cortex, independently of gender. This core set of AE proteins were differentially abundant from the earliest 3-week timepoint. Thus, proteome changes in the AE models precede seizure development. We also identified additional proteins with altered abundances that were specific to the two different epileptic rat models. The data show both that deep proteomic brain analysis needs to consider gender and that protein changes in the brain can occur well before seizure onset.

92

Is the mitochondrial protein processing system robust? Lessons from a combined N-terminomics and shotgun proteomics approach on human cells treated with rapamycin or zinc

Thierry Rabilloud[1], Joanna Bons[2], Charlotte Macron[2], Catherine Aude-Garcia[1], Sebastian Alvaro Vaca-Jacome[2], Magali Rompais[2], Sarah Cianferani[2], Christine Carapito[2]

1. Chemistry and Biology of Metals, UMR CNRS-CEA-UGA 5249, Grenoble, RA, France

2. IPHC/LSMBP, Strasbourg, Alsace, France

All but thirteen mammalian mitochondrial proteins are encoded by the nuclear genome, translated in the cytosol and then imported into the mitochondria. For a significant proportion of the mitochondrial proteins, import is coupled with the cleavage of a presequence called the transit peptide, and the formation of a new N-terminus. Determination of the neo N-termini has been investigated by proteomic approaches in several systems, but generally in a static way in order to compile as many N-termini as possible. In the present study, we have investigated how the mitochondrial proteome and N-terminome react to chemical stimuli that alter mitochondrial metabolism, namely zinc ions and rapamycin. To this end, we have used a strategy that analyzes both internal and N-terminal peptides in a single run, the dN-TOP approach. Rapamycin and zinc induced different changes in the mitochondrial proteome. However, convergent changes to key mitochondrial enzymatic activities such as pyruvate dehydrogenase, succinate dehydrogenase and citrate synthase were observed for both treatments. Other convergent changes were seen in components of the N-terminal processing system and mitochondrial proteases. Investigations into the generation of neo N-termini in mitochondria showed that the processing system is robust, as indicated by the lack of change in neo N-termini under the conditions tested. Detailed analysis of the data revealed that zinc caused a slight reduction in the efficiency of the N-terminal trimming system and that both treatments increased the degradation of mitochondrial proteins.

In conclusion, the use of this combined strategy allowed a detailed analysis of the dynamics of the mitochondrial N-terminome in response to treatments which impact the mitochondria.

920

Biological implications of gene modularisation in the virus φX174

Bradley W Wright[1], Paul R Jaschke[1], Mark P Molloy[2]

1. Macquarie University, North Ryde, NSW, Australia

2. Kolling Institute, Sydney University, Sydney, NSW, Australia

Modularisation of native genomic coding elements has been a feature in recent synthetic biology designs. The rationales behind these efforts is to produce gene pathways more amenable to manipulation, and to facilitate more predictable and efficient genome engineering outcomes. However, efforts can be hampered by the discovery of overlapping genes, leading to time and resource intensive characterisation and redesign to off-put losses in system efficiency and fidelity as a result of their disruption. To uncover the biological impacts of gene modularisation we characterise a synthetic φX174 virus (termed *decompressed*) in which all instances of gene overlap has been removed. Our results indicate that despite *decompressed* possessing the same coding repertoire as the wild-type, there are deficiencies in attachment, progeny production, and capsid stability. Furthermore, preliminary work at the proteome level through the use of targeted mass spectrometry has facilitated for the first time, identification of all known phage proteins. This will enable us to quantitatively measure and compare phage protein production between the wild-type and its genome decompressed variant. We aim to identify the causative reasons for loss in viability as a result of

modularisation through the linkage of protein expression, and phenotype measurements to that of *in silico* genome structure, and translation rate predictions. We will provide preliminary results of the above.

921

Discovery of new human protein coding genes in GENCODE using evolutionary signatures and mass spectrometry

James C Wright[1], Lu Yu[1], Jonathan Mudge[2], Irwin Jungreis[3], Jyoti Choudhary[1]

1. Institute of Cancer Research, London, LONDON, United Kingdom

2. Vertebrate Genomics, European Bioinformatics Institute, Cambridge, UK

3. Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, US

Despite nearly twenty years of intense study the exact portion of the human genome that is translated into protein remains to be ascertained. The GENCODE gene annotation is created by a consortium of manual annotation, computational biology and experimental groups. The objective of the GENCODE consortium is to create in the GENCODE geneset a foundational reference genome annotation encompassing all functional elements, including protein coding genes for Mouse and Human,. This annotation is released through Ensembl and forms the basis for most protein sequence databases used in proteomics. There is a huge focus on discovering and validating novel CCDS as well as removing spurious protein coding genes from the reference annotations to finalise the complete complement of reference proteins. This work has involved the implementation of strict criteria based on high quality experimental evidence as well as orthogonal sequence based characteristics to refine a genes protein coding potential and to confirm the structure of novel protein coding transcripts. One recent approach has been the use of PhyloCSF to identify the evolutionary signatures of protein-coding regions using multi-species genome alignments and machine learning. These regions provide potential novel conserved protein-coding sequences that can be searched using mass spectrometry data and has led to the largest annotation of new proteins in the human genome in recent years. This work approach will be important for the annotation of personal genomes and the use of proteomic experiments to help provide translational information about personal protein sequences and allelic specific expression.

922

Discovery of salivary biomarkers for oral cancer by lectin affinity capture coupled with iTRAQ-based quantitative glycoproteomics approach

Chia-Chun Wu[1], Wei-Fan Chiang[2], Jau-Song Yu[3]

1. Chang Gung University, Taoyuan, Linkou, Taiwan

2. Oral and Maxillofacial Surgery, Chi-Mei Medical Center, Tainan, Liouying, Taiwan

3. Liver Research Center, Chang Gung Memorial Hospital, Taoyuan, Linkou, Taiwan

Oral squamous cell carcinoma (OSCC) is the fifth most common cancer in Taiwan and the fourth in males. Oral cancer is associated with high death rate mainly because >1/2 patients were diagnosed at advanced stages that lead to poor prognosis. At present, there are still no suitable biomarkers for detecting early-stage oral cancer. Aberrant glycosylation is known highly associated with cancers. Saliva is an easy-to-obtain body fluid derived from oral cavity and directly contacts oral cancer nidus, representing an excellent resource for identifying novel, noninvasive biomarkers for detecting OSCC. Here, we systemically compared the difference of glycoproteins in saliva samples from healthy controls (HC), subjects with oral potentially malignant disorders (OPMD) and OSCC patients using lectin-based glycoprotein/glycopeptide enrichment coupled with iTRAQ-based quantitative approach. From this analysis, 673 proteins (353 glycoproteins) containing 2366 peptides (834 glycopeptides) were identified and quantified; among them, 15 glycopeptides (derived from 14 proteins) showed significantly higher expression in OSCC saliva and were selected as targets for further verification. We then established multiple reaction monitoring (MRM)-MS assays for these 15 glycopeptides and quantified their levels in 90 individual saliva samples (30 HC, 30 OPMD and 30 OSCC). The results showed that eight glycopeptide targets displayed significantly elevated levels in OSCC compared with HC or OPMD groups. Collectively, we have identified several novel biomarkers from saliva via glycoproteomic and MRM-MS analysis which may have potential to be further validated for their clinical use in oral cancer detection.

923

A rapid and highly selective method for N-glycopeptide enrichment by bacterial cellulose

Mengxi Wu[1], Weiqian Cao[1], Xinwen Zhou[1], Siyuan Kong[1], Pengyuan Yang[1]

1. Fudan University, Shanghai, China

Glycosylation is considered as one of the most frequent post translational modifications, which has great influence on many biological processes. Glycopeptide enrichment with high efficiency is critical in glycosylation research. However, most of the current enrichment methods of N-glycosylated peptides are low efficient, expensive and labor-consuming. Herein, we present a cost-friendly, biodegradable and easily obtainable material, bacterial cellulose, to enrich intact N-glycopeptides from different samples. The enrichment method based on this material turned out to be highly selective, easy to operate, and can maintain the intact glycopeptides. With this materials, intact glycopeptides can be easily obtained through incubation of tryptic digested peptide mixtures with a small piece of bacterial cellulose for merely 10 minutes. With this method, we successfully identified 97 intact N-glycopeptides from human IgG proteins, and 650 intact N-glycopeptides from human serum samples. In addition, the enrichment specificity of this method is more than 80%. Overall, the enrichment method developed in this study shows good performance compared with other existed methods, and a great potential in further application.

924

Comprehensive proteomic analysis of kidney biopsy tissues for prediction of prognosis and response to therapy in IgA nephropathy

Qingqing Wu[1], Rong Zeng[1]

1. Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, SHANGHAI, China

BACKGROUND: IgA nephropathy (IgAN) is the most prevalent type of primary glomerular disease (PGD) worldwide. Currently, prognosis of PGD is predicted depends on estimated glomerular filtration rate (eGFR) and the degree of proteinuria. However, these physiological parameters may be varied and not sensitive enough to predict prognosis and treatment outcomes. Therefore, it is necessary to identify novel biomarkers to guide therapeutic decisions. METHODS AND RESULTS: Using isobaric tags for relative and absolute quantification (iTRAQ) labeling coupled two-dimensional liquid chromatography-tandem mass spectrometry (2D-LC-MS/MS), we quantified a total of 581 proteins in all renal biopsy specimens of 54 IgAN patients who were given, on a randomized basis, either traditional Chinese medicine (TCM) or combined corticosteroids and TCM (CT) therapy. 6 modules of co-expressed networks were revealed by using weighted gene co-expression network analysis (WGCNA). Of these, 4 modules correlated with baseline eGFR (1 negatively, and 3 positively) or change in eGFR between baseline and 12-month follow-up (3 negatively, and 1 positively). Multiple linear regression models were applied to assess effects of individual proteins on treatment outcomes in patients with IgAN. After adjustment for potential confounders, including baseline eGFR, age, and gender, expression levels of 32 proteins in kidney were significantly ($p<0.05$) associated with change in eGFR between baseline and 12-month follow-up in CT treatment group (16 positive, 16 negative). CONCLUSIONS: We constructed the renal protein co-expression networks for IgAN patients which can be correlated with clinicopathological features and prognosis. The proteins associated with treatment outcomes of CT therapy in IgAN patients may serve as potential markers to guide medical decision making.

925

Differential oxidation of protein tyrosine phosphatases in metabolic diseases

Julia Bauza-Martinez[1], Giuseppe Sammarco[1], Esteban N Gurzov[2], Wei Wu[1]

1. Utrecht University, Utrecht, UTRECHT, Netherlands

2. ULB Center for Diabetes Research, Universite Libre de Bruxelles, Brussels, Belgium

Protein tyrosine phosphatases (PTPs) are a class of enzymes that attenuate phospho-tyrosine signaling. In various metabolic diseases, sustained dysregulation of insulin and leptin phosphorylation signaling have been reported, which may be corrected by PTP activity. While metabolic diseases are generally associated with an oxidative tissue environment that could oxidise (and thus inactivate) PTPs, the redox status of individual PTPs in disease settings are still largely unknown. Using a highly sensitive and robust AP-MS workflow, we have previously detected PTP oxidations (oxPTPs) in live zebrafish within 40 min of tail-fin amputation, and validated that the specific oxidation of PTPN11 in this time window is critical to initiate tissue regeneration. Due to high catalytic site sequence conservation between different PTPs and across organisms, collective enrichment and analysis of oxidized PTP active sites is feasible by AP-MS.

In this work, we streamlined the strategy to support oxPTP profiling from animal tissue and extremely limited needle-biopsies from patients, in the contexts of obesity and diabetes, as well as fatty liver and associated steatotic-inflammatory diseases. We show here that (i) high fat diet in mice induced specific PTP oxidations in a manner independent of liver tissue PTP expression, (ii) a unique oxPTP subset correlates with curative responsiveness in diabetic reversal by a novel combination therapy, and that (iii) distinct oxPTP profiles were detected in fatty/fibrotic livers compared to healthy human livers. The unifying oxPTP hits in these datasets may provide leads to understand the origin and bases for predisposition and co-morbidities across these metabolic conditions, and hint at novel mechanistic routes that prevent high BMI from progressing into metabolic disease.

926

Glycopeptidomics Analysis of a Cell Line Model Revealing Potential Marker Molecules for the Early Diagnosis of Gastric MALT Lymphoma

Di Xiao[1], Huifang Zhang[1], Fanliang Meng[1], Jianzhong Zhang[1]

1. National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

Gastric mucosa-associated lymphoma (GML) is a mature B cell tumor that accounts for 2-8% of all gastric cancers. *Helicobacter pylori* (*H. pylori*) infection is the main cause of GML, but the pathogenesis is not well understood. The genetic polymorphism of *H. pylori* is associated with the clinical outcome of infection. The clinical manifestation of GML are not specific, so GML commonly escapes diagnosis or is misdiagnosed, leading to excessive treatment. At present, there are no molecular markers for early GML diagnosis. Glycopeptidomics analyses of host cell lines (a BCG823 cell line, C823) and C823 cells infected by *H. pylori* isolated from patients with GML (GMALT823), gastritis (GAT823), gastric ulcer (GAU823) and gastric cancer (GAC823) were carried out to clarify the host reaction mechanism against GML and identify potential molecular criteria for the early diagnosis of GML. In total, 33 samples were analyzed and approximately 2000 proteins, 200 glycoproteins and 500 glycopeptides were detected in each sample. O-glycans were the dominant glycoforms in GMALT823 cells but not GAT823, GAU823 and GAC823 cells. Four specific glycoforms in GMALT823 cells and 2 specific glycoforms in C823 and GMALT823 cells were identified. Eight glycopeptides with specific glycosylation sites and the glycoforms of 7 glycoproteins were found in GMALT823 cells; of these glycopeptides, 6 specific glycopeptides have a high affinity for T cell epitopes and participate in cellular immunity, and 3 glycopeptides have conformational B cell epitopes and mediate liquid immunity. In this study, the relationship between the predominant glycoforms of host cells and the development of host disease was determined. Seven glycoproteins, 8

glycosylation sites and 9 glycoforms might be closely related to the formation of GML, which provides new insight into the pathogenic mechanisms of *H. pylori* infection and suggests molecular indicators for the early diagnosis of GML.

927

Multilayered N-glycoproteomes reveals generally impaired N-glycosylation promoting Alzheimer's disease

Juanjuan Xie[1], Pan Fang[1], Yang Zhang[1], Pengyuan Yang[1, 2], Huali Shen[1]

1. Institutes of Biomedical Sciences of Shanghai Medical School, Fudan University, Shanghai, China

2. Department of Chemistry, Fudan University, Shanghai, China

Alzheimer's disease (AD) is one of the most common neurodegenerative diseases that currently lacks clear pathogenesis and effective treatment.[1-4]. Protein glycosylation is ubiquitous in brain tissue and is closely related to important functions of the nervous system such as cognition and memory. Site-specific analysis of N-glycoproteome, which is technically challenging, can advance our understanding of the glycoproteins' role in AD. To circumvent the challenges, our lab developed a large scale and precision identification method for site-specific N-glycopeptides named pGlyco 2.0[5]. In this study, we profiled the multilayered variations in proteins, N-glycosites, N-glycans, and in particular site-specific N-glycopeptides in the APP/PS1 and wild type mouse brain through combining pGlyco 2.0 with other quantitative N-glycoproteomic strategies. The comprehensive brain N-glycoproteome landscape was constructed, and rich details of the heterogeneous site-specific protein N-glycosylations were exhibited. In particular, we have found that the overall down-regulation of the N-glycosylation modification in AD, and proved that down-regulation of glycosylation modification causes hyper-excitation and apoptosis of neuronal cells, as well as mis-localization and dysfunction of important proteins such as NCAM-1. In general, our work offered a panoramic view of the N-glycoproteomes in Alzheimer's disease and revealed that generally impaired N-glycosylation promotes Alzheimer's disease progression. Meanwhile, this study demonstrates the important role that general regulation of glycosylation modifications may have crucial effects on biological and pathological systems of AD and will broad the way for AD therapy.

928

Changes in the Insolublome with Aging and Alzheimer's Disease models

Xueshu Xie[1], Manish Chamoli[1], Dipa Bhaumik[1], Kathleen Dumas[1], Renuka Sivapatham[1], Suzanne Angeli[1], Anja Holtz[1], Julie Andersen[1], Gordon J. Lithgow[1], Birgit Schilling[1]

1. Buck Institute, Novato, CA, United States

Studies have shown that protein insolubility increases with age. The proteins that become insoluble with age are enriched for lifespan determining functions. Knocking down some of the genes encoding the insoluble proteins has shown extension of lifespan in nematodes. Age-related Diseases, e.g. Alzheimer's disease (AD) are known to be associated with protein aggregation, but it is not clear how normal aging processes and age-related disease processes are related. In this study, we investigated the insolublome (1% SDS insoluble proteins) of *C. elegans* in aging model (young vs. old for N2 and Ju775 strains) and disease model (human Ab expressing worms vs. control). We aim to decipher the significant changes and the possible correlation of insolublome in aging and age-related disease. Our preliminary results indicate normal aging as an important driver, and relevant observations to reveal the molecular mechanism of AD pathology.

929

Distinguishing pancreatic cancer from benign diseases and healthy individuals by mass spectrometry-based metabolomic pipeline

Yueting Xiong[1], Xiaohui Liu[1, 2], Pengyuan Yang[1]

1. Institutes of Biomedical Sciences, Fudan University, Shanghai, China

2. The Fifth People's Hospital of Shanghai, Fudan University, Shanghai, China

Pancreatic cancer (PC) is the third leading cause of cancer death because of its unobvious clinical symptoms in the early stage, making the urgency to discover particular metabolites as potential diagnostic markers for early diagnosis of PC. The study aims to identify differential metabolites distinguishing PC, benign diseases (BD) and healthy individuals (normal control, NC) using liquid chromatography-mass spectrometry (LC-MS).We combined both reversed-phase (C18) and hydrophilic interaction (HILIC) liquid chromatography separation and a high-resolution quadrupole time-of-flight mass spectrometer detection together which was operated in negative (ESI–) and positive (ESI+) electrospray ionization mode separately, in order to comprehensively analyze serum metabolites obtained from 20 patients with PC, 10 patients with BD and 10 NC. Differential metabolites were selected by univariate (Student's t test) and multivariate (orthogonal partial least squares-discriminant analysis (OPLS-DA)) statistics. Bile acids, amino acids and choline-containing glycerophospholipids were identified as differential metabolites for discriminating PC, BD and NC. These altered metabolites corresponding to 3 most important pathways, including primary bile acid biosynthesis, biosynthesis of amino acids, and glycerophospholipid metabolism. The newly found metabolites by using LC-MS-based metabolomics may be potential biomarkers for the early diagnosis of PC.

1. Kamisawa T, Wood L D, Itoi T, et al. The Lancet, 2016, 388(10039): 73-85
2. Liu X, Zheng W, Wang W, et al. A new panel of pancreatic cancer biomarkers discovered using a mass spectrometry-based pipeline. British Journal of Cancer, 2017, 117(12):1846

93

Measuring protein functional states in central carbon metabolism by limited proteolysis coupled to mass spectrometry

Thomas Hauser[1], Paola Picotti[1]

1. ETH Zürich, Zürich, ZüRICH, Switzerland

Publish consent withheld

1. (1) Gerosa L, Haverkorn van Rijsewijk BR, Christodoulou D, Kochanowski K, Schmidt TS, Noor E, Sauer U (2015) Pseudo-transition Analysis Identifies the Key Regulators of Dynamic Metabolic Adaptations from Steady-State Data. Cell Syst 1: 270-82
2. (2) Feng Y, De Franceschi G, Kahraman A, Soste M, Melnik A, Boersema PJ, de Laureto PP, Nikolaev Y, Oliveira AP, Picotti P (2014) Global analysis of protein structural changes in complex proteomes. Nat Biotechnol 32: 1036-44

930

Urine Proteomics of Japanese Child and Adult Healthy Volunteers by SWATH-MS analysis

Bo Xu[1], Yoshitoshi Hirao[1], HIdehiko Fujinaka[2], Keiko Yamamoto[1], Shuichiro Shimada[1], Amr Elguoshy[1], Tadashi Yamamoto[1, 3]

1. Biofluid Biomarker Center (BB-C), Niigata University, Niigata, Japan

2. Pediatric, National Hospital Organization Niigata Hospital, Niigata, Japan

3. Shinrakuen Hospital, Niigata, Japan

Introduction

Urine is an important specimen for biomarker discovery and health monitoring due to easy and noninvasive accesses. As a part of human urine proteomics, children's samples are also important. Since kidney functions develop fast in young ages, we focus on comparative urinary quantitative proteomics of children and adults to clear age, gender or individual variance.

Methods

16 urine samples of 4 to 13-year-old children and 50 urine samples of 20's to 70's healthy volunteers were collected. Pre-centrifuged urine samples were stored at -20C until protein preparation. Purified tryptic peptides each was analyzed in IDA and DIA methods alternately. Protein Identification (ID) done by against SWISSPROT human database. SWATH-MS data processing was set as 6 peptides / protein, 6 transitions / peptide with at least 95% confidence threshold. Valve of each protein normalized by summed intensity value of all proteins in ppm level. Quality and quantitative proteomics between 2 groups were compared in gender and ages respectively.

Results

Protein concentration of urine was variable in all ages and no relationship with gender. In current, ~1,200 protein groups in children and ~2,000 protein groups in adults were identified in total. Some proteins with high expression observed in urine of children and some in girls also show a similar trend in adult females. As high expression proteins in boy's samples, most of them are binding proteins and catalytic activity proteins, 1/3 of them are related to cellular biological process. It may show the high level of such process in young male at the age of 5-13.

Conclusions

Urine protein information of young age may be one of important references to support more precision identification of urine biomarker for disease detection and health monitoring for children in future.

931

Characterization of Protein Profile and Lysine Methylation in EZH2 Wild-Type and Mutant Hematologic Cell Lines

junyu xu[1], bingbing hao[1], minjia tan[1]

1. Shanghai Institute of Materia Medica, Pudong, SHANGHAI, China

Publish consent withheld

1. 1. Murn, J. & Shi, Y. The winding path of protein methylation research: milestones and new frontiers. Nature Reviews Molecular Cell Biology 18, 517 (2017).
2. 2.Kim, K.H. & Roberts, C.W. Targeting EZH2 in cancer. Nature Medicine 22, 128-134 (2016).
3. 3. Huang, X. & Roberts, C.W. Targeting Epigenetic Crosstalk as a Therapeutic Strategy for EZH2-Aberrant Solid Tumors. Cell 175,186-199 (2018).

932

Surfaceome of Exosomes Secreted from the ColorectalCancer Cell Line SW480: Peripheral and Integral MembraneProteins Analyzed by Proteolysis and TX114

Rong Xu[1], David Greening[1], Maoshan Chen[1], Alin Ray[1], Hong Ji[1], Nobuhiro Takahashi[2], Richard Simpson[1, 2]

1. La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC, Australia

2. Global Innovation Research Organization, Tokyo University of Agriculture and Technology, Tokyo, Japan

Exosomes are important bidirectional cell–cell communicators in normal and pathological physiology. Although exosomal surface membrane proteins (surfaceome) enable target cell recognition and are an attractive source of disease marker, they are poorly understood. Here, a comprehensive surfaceome analysis of exosomes secreted by the colorectal cancer cell line SW480 is described. Sodium carbonate extraction/Triton X-114 phase separation and mild proteolysis (proteinase K, PK) of intact exosomes is used in combination with label-free quantitative mass spectrometry to identify 1025 exosomal proteins of which 208 are predicted to be integral membrane proteins (IMPs) according to TOPCONS and GRAVY scores. Interrogation of UniProt database-annotated proteins reveals 124 predicted peripherally-associated membrane proteins (PMPs). Surprisingly, 108 RNA-binding proteins (RBPs)/RNA nucleoproteins (RNPs) are found in the carbonate/Triton X-114 insoluble fraction. Mild PK treatment of SW480-GFP labeled exosomes reveal 58 proteolytically cleaved IMPs and 14 exoplasmic PMPs (e.g., CLU/GANAB/LGALS3BP). Interestingly, 18 RBPs/RNPs (e.g., EIF3L/RPL6) appear bound to the outer exosome surface since they are sensitive to PK proteolysis. The finding that outer surface-localized miRNA Let-7a-5p is RNase A–resistant, but degraded by a combination of RNase A/PK treatment suggests exosomal miRNA species also reside on the outer surface of exosomes bound to RBPs/RNPs.

933

Proteomic analysis of low numbers of Hela cells using a microfluidic device

Peng Xue[1, 2], Xiaobao cao[3], Ludovic gillet[1], Sofia Theodorou[1], Federico Uilana[1], Chen Li[1, 4], Andrew deMello[3], Rudolf Aebersold[1]

1. Institute of Molecular Systems Biology, ETH, Zurich, zurcih, Switzerland

2. Institute of Biophysics , Chinese Academy of Sciences, Beijing, China

3. Department of Chemistry and Applied Biosciences , ETH, Zurich, zurcih, Switzerland

4. Biomedicine Discovery Institute, Monash University, Melbourne, Victoria, Australia

Single-cell analysis enables the characterization of highly heterogeneous cellular heterogeneity and the complexity of cellular interactions by the collection of genomic, transcriptomic, proteomics and metabolomics information. While genomic and transcriptomic single-cell based approaches are well-established, the proteomic profiling of single-cells is however still debuting as a novel research area. Capturing proteomic information from individual cells is technically challenging in terms of sample handling and preparation, as it is limited by the sensitivity of the mass spectrometry acquisition level and data analysis. Here we present a novel PDMS(Polydimethylsiloxane) microfluidic device for the proteomic analysis of small cell populations. The developed device enables the capturing of few cells ranging from 1 to 200 cells in a multiplexed way to allow paralleled sample preparations. Within the device, cells are subsequently lysed, and the released proteins are then reduced, alkylated and finally digested. All these steps can be performed in a reduced nanoscale volume of a total of 30 nl. The minimum surface of the device dramatically improves the recovery of the peptides during the sample processing due to the peptide nonspecific adsorption surface binding. The digestion efficiency in the chip demonstrated to be equal or even better than the in-solution bulk (1mL) or low volume (10µL) digestion in the tubes. Additionally, we optimized the data-independent acquisition (DIA) sensitivity by increasing the ion injection time to 246ms and setting the resolution to 120,000 at the expense of the mass range covered. With these settings, we were able to reproducibly measure ~500 proteins from 250 pg of peptides (~1 HeLa cell). In conclusion, by combining this microfluidics device to a sensitive and robust quantitative technique based on a data-independent acquisition mass spectrometry method, we have measured reproducibly 2000 protein groups based on 10,000 peptides extracted from a total of 50 Hela cells consistently.

934

Short gradient time LC-MS for clinical application to analysis of multiple samples

Keiko Yamamoto[1], Yoshitoshi Hirao[1], Bo Xu[1], Amr El Guoshy[1], Shuichiro Shimada[1], Shigeru Miyazaki[2], Tadashi Yamamoto[1, 2]

1. Niigata University Biofluid Biomarker Center, Niigata, NIIGATA, Japan

2. Shinrakuen Hospital, Niigata, NIIGATA, Japan

LC-MS has been used for comprehensive identification of proteins in tissues or biofluids by LC separation of peptides for a long gradient time (~2 hours). On the other hand, recent development of LC and MS may propose an application of LC-MS for measurement of biomarkers in clinical samples at clinical laboratories in hospitals. For the clinical application high-throughput LC-MS measurement should be required.

To make it possible, we analyzed proteomes of 30 urine samples in 24 hours by using a new high-throughput LC (Evosep One) and a new efficient MS (Bruker tims TOF pro) with a 48 min measurement method. Our standard measurement LC-MS (nanoElute-timsTOFpro, Bruker) system was also used in a short gradient time method such as 30 min, 68 min, or 120 min.

~ 2,000 proteins were identified by our standard measurement using the nanoElute-timsTOFpro with 120 min gradient time for a 200 ng protein urine sample measurement. With shortening the gradient time to 68 min or 30 min, ~1,100 or ~900 proteins were identified by using a nanoElute LC system. With a new LC, Evosep One, ~1,000 urine proteins were identified with the 48 min measurement time for consecutive three days. These LC-MS measurements may propose an comprehensive analysis of clinical samples for understanding patients or diseases.

935

Proteomics of FFPE kidney biopsy tissues of chronic kidney disease (CKD) for precision medicine

Tadashi Yamamoto[2, 1], Keiko Yamamoto[1], Bo Xu[1], Amr El Guoshy[1], Shuichiro Shimada[1], Yoshitoshi Hirao[1]

1. Biofluid Biomarker Center (BBC), Niigata University, Niigata, Japan

2. Shinrakuen Hospital, Niigata, Japan

CKD is a disease clinically defined as the patients have kidney function loss and insidiously reached to the end-stage kidney failure, in which they need to be treated with dialysis therapy or renal transplantation. Since causes of CKD are various, CKD is pathologically classified into several disease types including diabetes, hypertension, chronic glomerulonephritis, polycystic kidney disease and so on by histological examination of kidney biopsy. However, in spite of the pathological diagnosis, specific therapies for each type have not been developed as the pathological mechanisms of CKD or each CKD type have not been elucidated yet.

To understand the pathological mechanisms of each CKD type more precisely, we aimed to analyze CKD kidney biopsy specimens by proteomics. Firstly we developed a method to obtain peptides of enough amounts (~1

936

Integrated transcriptomic and proteomic analyses reveal novel immune regulators in pneumonia

Chia-Yu Yang[1, 2, 3], Chia-Yin Chin[3], Chih-Ching Wu[3, 4], Kuo-An Wu[5]

1. Department of Microbiology and Immunology, College of Medicine, Chang Gung University, Taoyuan, Taiwan

2. Department of Otolaryngology-Head and Neck Surgery, Chang Gung Memorial Hospital, Taoyuan, Taiwan

3. Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

4. Department of Medical Biotechnology and Laboratory Science, College of Medicine, Chang Gung University, Taoyuan, Taiwan

5. Department of Internal Medicine, Taoyuan Armed Forces General Hospital, Taoyuan, Taiwan

Bacterial pneumonia is an infectious disease with sever inflammatory condition of the lung. Parapneumonic effusion (PPE) is an accumulation of exudative pleural fluid associated with the pulmonary infection. PPEs are present in 20% to 40% of hospitalized patients with pneumonia and progressed from uncomplicated parapneumonic effusions (UPPE), complicated parapneumonic effusions (CPPE) to empyema. Pneumonia patients with late stage PPE have elevated mortality and a poor prognosis. To study the pathogenesis of PPE progression and identify the useful biomarkers for PPE diagnosis, we performed the transcriptomic and proteomic analyses to identify the biological pathways and networks in pneumonia patients with different stages of effusions. Comparative analysis of PPE transcriptome dataset allowed us to identify 1924 genes as differentially expressed genes; these included 331 upregulated and 1593 downregulated genes in CPPE compared to UPPE. The 331 up-regulated genes were enriched in pathways associated with respiratory burst, canonical glycolysis, and regulation of cell proliferation using gene set enrichment analysis. We also analyzed the secretome in pleural effusion using iTRAQ-based proteomic technology. About 760 proteins were identified in the pleural effusion and 50 proteins were defined as upregulated proteins in CPPE. The enrichment biological processes revealed that the up-regulated proteins were significantly involved in the innate immune response, defense response, and proteolysis. Integration of these two datasets validated several novel immune regulators as secreted proteins in pleural effusion. One of these secreted proteins was highly expressed in neutrophils and can be induced in *Streptococcus pneumonia* infected dHL60 cells. Also, the overexpression of this candidate protein was validated in CPPE patients using ELISA assay, western blot, and immunohistochemical staining. The area under curve of this candidate protein was 0.75 for distinguishing CPPE from UPPE. In conclusion, we identified one novel protein involved in the defense response to bacteria and were significantly associated with CPPE.

937

Discovery of putative drug resistance markers of MCF-7/ADR cells using GPSeeker-enabled quantitative structural N-glycoproteomics

Hailun Yang[1], Yue Wang[1], Kaijie Xiao[1], Zhixin TIAN[1]

1. Tongji University, Shanghai, China

Background: Multidrug resistance (MDR), a major hurdle in the cancer treatment, has been closely associated with aberrant regulation of N-Glycosylation; identification of the related differentially expressed N-glycoproteins is essential for understanding the underlying molecular mechanisms and developing efficient drugs and therapies.

Methodologies: Differentially expressed N-glycosylation in MCF-7/ADR cells was characterized with our recently developed GPSeeker-enable quantitative structural N-glycoproteomics pipeline. With trypsin digestion, zicHILIC enrichment, isotopic diethyl

labeling, 1:1 mixture of the labeled intact N-glycopeptides from MCF-7/ADR and MCF-7 cells was analyzed using C18-RPLC-nanoESI-MS/MS (HCD with stepped NCEs) on a Q Exactive Orbitrap MS; three technical replicates were acquired and the raw datasets were searched by intact N-glycopeptide search engine GPSeeker; intact N-glycopeptide IDs were further quantified with the abundance of the precursor ions in the MS spectra using GPSeekerQuan; final differentially expressed intact N-glycopeptides (DEGPs) were obtained with the criteria of observation of at least twice among the three technical replicates, ≥1.5 fold-change and $p<0.05$.

Findings: intact N-glycopeptides in MCF-7/ADR and MCF-7 cells each with the comprehensive structural information of both the peptide backbone (amino acid sequence, N-glycosite) and the N-glycan moiety (monosaccharide composition, sequence, linkages) were identified; unique N-glycan structures were characterized with structure-specific fragment ions including cross-ring A/X ions. Putative drug resistance N-glycosylation markers were found at DEGPs.

Conclusions: putative drug resistance N-glycosylation markers were discovered in the high-throughput omics fashion with GPSeeker-enable quantitative structural N-glycoproteomics pipeline at the intact N-glycopeptide level.

Figure 1. Intact N-glycopeptide SLSNSTAR-N2H8F0S0 on the N-glycosite N120 of N-glycoprotein Serpin H1 (SERPH_HUMAN, P50454) were identified from spectrum 3788 of TR1. From top to bottom, the isotopic envelopes of the paired precursor ions, the annotated MS/MS spectrum and the graphical fragmentation map with the matched fragment ions.

938

Proteogenomic analysis of cetuximab-resistant clonal populations derived from colorectal cancer cells

SEOJIN YANG[1], Hyeryeon Jung[1], Yu Ri Seo[1], Ye-Lim Park[2, 1], Hwang-Phill Kim[2, 1], Eugene C Yi*[1]

1. Department of Molecular Medicine& Biophamaceutical Science, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Korea

2. Cancer Research Institute, Seoul National University, Seoul, Korea

Acquired resistance to cetuximab in wild-type KRAS colorectal cancer (CRC) has become a major clinical challenge, while mutation events in KRAS are well recognized as frequent drivers of acquired resistance to cetuximab in CRC. Hence, elucidating accurate molecular mechanisms conferring resistance to cetuximab in wild-type KRAS CRC would provide new treatment options in CRC patients. In-vitro cell model system can mask molecular signatures of individual cells with intracellular heterogeneity, which limits defining distinct molecular mechanisms. Hence, an approach at the relatively-homogeneous cellular level would allow study of specific aspect of resistance mechanisms in detail. Here we established five CRC homogenous clonal populations through single-cell cloning and expansion of cetuximab-resistant human colorectal cancer cells (NCI-H508, mix cell line), which were generated by prolonged incubation with cetuximab. Utilizing Tandem Mass Tag (TMT) isobaric labeling, we performed a quantitative global and phosphoproteome analysis of the five different resistant clonal populations including the parental and the mix cell lines by a serial proteome sample enrichment process. We then integrated proteomic data with previously acquired transcriptomic data and systematically identified functional proteins along with key biological pathways of each resistant clonal population. In this work, we present wild-type KRAS-centric specific biological pathways associated with cetuximab-resistant CRC. The integrative approach combining multiple datasets collected from the homogenous cell populations enables us to accurately capture key biological pathways related to anti-cancer drug resistance.

939

Prediction of tandem mass spectra and retention time of peptide by deep learning

Yi Yang[1], Liang Qiao[1], Xiaohui Liu[1], Chengpin Shen[2], Yu Lin[3]

1. Fudan University, Shanghai, SHANGHAI, China

2. Shanghai Omicsolution Co., Ltd, Shanghai, China

3. The Australian National University, Canberra, Australia

Liquid chromatography (LC)–tandem mass spectrometry (MS/MS) has been the most widely used tool for proteomics studies over the past decades. Developing more accurate MS/MS spectrum and retention time (RT) prediction methods is one of the most effective way to improve the confidence of peptide identification.

In this work, we propose a deep learning-based method for accurate MS/MS spectrum and RT prediction using a hybrid model that combines convolutional neural network and bi-directional long-term and short-term memory network. The model takes a peptide sequence as an input, and outputs relative intensities of b/y product ions at each possible fragmentation site including neutral loss of ammonia or water, as well as normalized RT (iRT [1]) of the peptide.

We trained and validated the model with three LC-MS/MS data sets of two organisms acquired in two laboratories, i.e. two data sets of HeLa cells and a data set of mouse cerebellum [2,3]. The median dot products between the predicted and experimental b/y peak intensities (0.87–0.94) and the Pearson correlation coefficients of predicted and experimental iRT (> 0.98) were comparable to experimental repeats with each data set and for cross-organism and cross-lab validation. For benchmark purpose, we further compared the performance of our model on peptide fragmentation and RT prediction to existing tools. Our model outperformed most of the existing methods and can be adapted to both collision-induced dissociation and higher energy collisional dissociation fragmentation.

In contrast to traditional machine learning methods requiring manual design of appropriate features, the deep neural networks can learn different representations of objects automatically, and thus are more capable of handling the complexity of peptide fragmentation and retention. Our method will benefit assays development for selected reaction monitoring or parallel reaction monitoring, as well as data-independent acquisition proteomics [4].

1. Escher, C. et al. Proteomics 12, 1111-1121 (2012).
2. Bruderer, R. et al. Mol. Cell. Proteomics 16, 2296-2309 (2017).

3. Kelstrup, C. D. et al. J. Proteome Res. 17, 727-738 (2018).
4. Qiao, L. et al. HUPO 2019 Abstract 63848.

94

Integrated TAILS terminomics, shotgun, and transcriptomics analysis of macrophage polarization and activation

Nestor Solis[1], Antoine Dufour[2], Melanie Y White[3], Reinhild Kappelhoff[1], Anders Kristensen[1], Leonard J Foster[1], Stuart J Cordwell[3], Christopher M Overall[1]

1. University of British Columbia, Vancouver, BRITISH COLUMBIA, Canada

2. University of Calgary, Calgary, Alberta, Canada

3. University of Sydney, Sydney, NSW, Australia

Background

Macrophages are critical immune cells in the activation and timely resolution of inflammation. Classically-activated macrophages (M1) perform hallmark roles in inflammation including mobilization to sites of injury and phagocytosis of foreign bodies. We used Terminal Amino Isotopic Labeling of Substrates (TAILS) N terminomics to identify targets of intracellular cysteine and serine proteases that are required for macrophage polarization.

Methodologies

THP-1 monocytes were polarized with phorbol myristate acetate (PMA) to M0 macrophages and further differentiated with interferon gamma (IFNγ) or interleukin-4 (IL-4) to M1 or M2-type macrophages, respectively. We profiled macrophages against 720 human proteases, homologs and inhibitors using the protease/inhibitor CLIP-CHIP cDNA microarray. Macrophages were grown in triplex SILAC media, lysates digested, fractionated and analysed by a Thermo LTQ-Velos Orbitrap. SILAC-labeled M1 macrophages were treated with E64 (broad cysteine protease inhibitor) or AEBSF (broad serine protease inhibitor) and their N-terminomes analysed by TAILS using trypsin and lysargiNase digestion on a Thermo Q-Exactive HF.

Findings

CLIP-CHIP cDNA microarray quantified 429 protease and homologs as being expressed in M0, M1 and M2 macrophages, with 14 genes upregulated in M1 macrophages, of interest Cathepsin-B, -L, -S, Caspase-1 and TNFAIP3. Proteomic SILAC analysis quantified 1421 proteins showing that interferon response regulators were the most abundant protein changes in M1 macrophages as well as CATL, LAP3, PSMB9 being M1 protease effectors. Serine and cysteine protease inhibition abolished cell migration whereas cysteine protease inhibition alone abrogated endocytosis. SILAC TAILS terminomics identified 19,363 N-terminal peptides that were otherwise undetected by preTAILS (shotgun) analysis. SILAC abundance of 460 N-dimethylated termini were significantly changed >2σ by inhibition. These clustered to networks related to actin-, tubulin- and phagosome dynamics and calcium signaling.

Concluding statement

TAILS positional proteomics quantified 19,363 N-termini and transcriptomics revealing essential cysteine proteases in the cytoskeletal remodeling that is crucial for endocytosis in M1 macrophages.

940

Comparative proteomic analysis reveals cellular response to ionizing radiation exposure in bone marrow of mice

Jeonghun 1 Yeom[1], Jiyoung 1 Yu[1], Hee Sung 1 Ahn[1], Hwangkyo 1 Jeong[1], Je-Won 1 Ryu[1], Sang-wook 1 Lee[2], Kyunggon 1 Kim[1]

1. Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea

2. Department of Radiation Oncology, Asan Medical Center, University of Ulsan College of Medicine, SEOUL, Republic of Korea

Radiation exposure such as radiation therapy inevitably causes side effects in normal tissues and cells. These side effects are typically inflammatory reactions due to abnormal bone marrow responses. In order to understand the molecular mechanism of these side effects, we tried to identify the differences between Lipopolysaccharide(LPS)-induced sepsis model mice and radiation exposed mice. Sepsis model mice except normal mice and radiation exposed mice were also analyzed by considering the conditions for 4 hours and 24 hours after treatment to observe the change with time. For proteomics analysis, mouse bone marrow of each condition was extracted and mass analysis was performed by obtaining protein from the extracted bone marrow. A TMT-labeling quantification method was introduced for comparative analysis of each sample. A total of 4341 proteins were excavated through three repeated analyzes. Among them, 2,675 proteins containing quantitative information in each sample were used for comparative analysis. As a result, it was confirmed that the sample was exposed to radiation rather than the sample of other conditions, and the change of the proteome of the sample after 24 hours appeared. In particular, we observed changes in NRF2-mediated oxidative stress response. These results will help to understand the mechanism of the inflammatory response caused by radiation exposure.

941

Peptide end sequence information in HCD spectra for protein identification

Akiyasu C. Yoshizawa[1], Tsuyoshi Tabata[1], Naoyuki Sugiyama[1], Yasushi Ishihama[1]

1. Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, KYOTO, Japan

Background
Since Nielsen et al. (J Proteome Res 2002, 1, 63) reported the usefulness of peptide “end sequencing” using MALDI Q-TOFMS, product ion information generated by tandem mass spectrometry from both peptide termini have been widely used to facilitate protein identification. For example, the N-terminal amino acid determined by the presence of a1 and/or b1 ions can reduce the search space in database search to 1/19. Likewise, y1 ion can determine the C-terminal amino acid, reducing the search space down to 1/2 for tryptic peptides. Here, we investigated the utility of product ions from the peptide end sequence, obtained by LC/ESI/Q-orbitrap instruments with HCD fragmentation.

Methods
One data file from a public dataset (PXD005159) was used as a sample dataset, and 44,047 spectra were searched by Mascot against Swiss-Prot human all sequence database. The annotated peak list was obtained and the peptide end sequence information as well as immonium (iminium) ions were investigated.

Results
Out of the 26,321 total annotated spectra, immonium ions, a2 ions, y1 ions and y2 ions were observed in 17,894, 18,661, 10,134 and 20,059 spectra, respectively; y1 ions were observed in 86% of all spectra. These tendencies were the same as those observed by Nielsen et al.

Conclusions
HCD spectra by LC/ESI/Q-orbitrap contain rich information on peptide end sequence, which can be utilized to facilitate peptide/protein identification. In this poster, the results for the large-scale datasets will be described.

942

Identification of isoaspartylated peptides by electron-transfer/higher-energy collision dissociation

Clifford Young[1], Frank Kjeldsen[2], Michael L Nielsen[1]

1. The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

2. Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

Isoaspartic acid is a post-translational modification that occurs as a result of deamidation or isomerisation reactions. Although numerous MS methodologies can characterise these subtle protein modifications, isoaspartylation sites are commonly identified by electron-transfer dissociation (ETD) due to the production of c+57 and z˙-57 diagnostic ions that pinpoint the site of modification. However, ETD is largely inefficient for doubly protonated peptides, which is unfortunately the predominant charge state of protonated tryptic peptides after electrospray ionization. We performed electron-transfer dissociation with supplemental collisional activation or electron-transfer/higher-energy collision dissociation (EThcD) on four doubly charged monoisoaspartylated peptides to see if the overall sequence coverage and/or isoaspartic acid discrimination was improved when compared to ETD alone. Although all three methods produced z˙-57 ions that confidently located the isoaspartic acid, only EThcD produced MS/MS spectra rich in additional y ions that allowed the peptide sequences to be extensively characterised. The use of EThcD permits the identification of isoaspartylated peptides when their physicochemical properties may be unsuitable for ETD.

943

MSFragger fast and sensitive peptide identification in diverse proteomic datasets

Fengchao Yu[1], Guo Ci Teo[1], Andy T Kong[1], Felipe V Leprevost[1], Dmitriy M Avtonomov[1], Hui-Yin Chang[1], Daniel Geiszler[1], Sarah E Haynes[1], Alexey I Nesvizhskii[1]

1. University of Michigan, Ann Arbor, MICHIGAN, United States

MSFragger is an MS/MS database search tool that uses a fragment ion indexing method to achieve very fast search speeds. We present comparisons showing that MSFragger is the fastest search engine among the five competitors supporting open (mass-tolerant) searching (MODa, PIPI, MSFragger, pFind3, and TagGraph). It also has the highest sensitivity and the lowest error rate. Since its publication in 2017, MSFragger has been gaining wide adoption by many users and in a variety of applications. Our group has continued to improve MSFragger for robustness and versatility. Among these advances, there are a number of key features that significantly improve its performance. These include shifted ions searching, mass calibration, parameter optimization, split database search, and additional ion series (e.g. *a*, *c*, *x*, and *z* ions). The shifted ions searching takes the effects of unknown modifications into account in scoring, which increases the sensitivity of identifying modified peptides. The mass calibration feature calibrates both precursor and fragment masses, which allows narrower tolerances in searching and reduces the number of false positives (especially in open searches). The parameter optimization feature tunes parameters according to spectral properties and calibrated masses. Owing to fast run time, MSFragger can try different combinations of the parameters to maximize the search results. The split database search feature allows using MSFragger to search very large databases, and to perform nonspecific searches (e.g. for HLA peptides) on computers with limited memory (RAM). We also demonstrate that MSFragger can analyze timsTOF PASEF data with higher sensitivity and significantly shorter run time compared to MaxQuant and PEAKS Studio X. Additional ongoing developments include support for reading raw files (for Thermo Fisher and Bruker

instruments), and improved analysis of glycopeptides. MSFragger can be run as a command-line tool, using FragPipe GUI, or as a ProteinDiscover Node (https://msfragger.nesvilab.org/).

944

BRAF protein immunoprecipitation, elution, and digestion from cell extract using a microfluidic mixer for mutant BRAF protein quantification by mass spectrometry

Yen-Heng Lin[1], Chia-Chun Wu[2], Jau-Song Yu[2]

1. Department of Electronic Engineering, Chang Gung University, Taoyuan, Taiwan

2. Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

This study utilized a microfluidic mixer for the sample pretreatment of cell extracts for target protein quantification by mass spectrometers, including protein immunoprecipitation and protein enzymatic digestion. The time of sample pretreatment was reduced and thus the throughput of quantitative mutant proteins was increased by using the proposed method. Whole cell lysates of the cancer cell line HT-29 with gene mutations were used as the sample. The target protein BRAF (v-raf murine sarcoma viral oncogene homolog B1) was immunoprecipitated using magnetic beads in a pneumatic micromixer. Purified protein was then eluted and digested by trypsin in another two micromixers to yield peptide fragments in the solution. Using stable isotope-labeled standard as the internal control, wild-type and mutant BRAF V600E proteins were quantified using mass spectrometry, which could be used for cancer screening. Compared with conventional methods in which protein immunoprecipitation lasts overnight, the micromixer procedure takes only 1 h, likely improving the throughput of mutant BRAF protein quantification by mass spectrometry.

945

Immunoproteomics approach to discover immunogenic SFTS virus antigen for vaccine and diagnostic kit development

Jiyoung Yu[1], Eunsil Kim[2], Hee-Sung Ahn[1], Jeonghun Yeom[1], Yumi Oh[1], Hwangkyo Jeong[1], Yujin Jo[1], Keun Hwa Lee[3], Yangsoo Kim[2], Kyunggon Kim[4]

1. Dept. of Convergence Medicine, Asan Medical Center, Songpa-gu, SEOUL, South Korea

2. Dept. of Infectious Disease, Asan Medical Center, Seoul, South Korea

3. Dept. of Microbiology and Immonolgy, Jeju National University, Jeju, South Korea

4. Dept. of Convergence Medicine, Asan Medical Center/University of Ulsan, Seoul, South Korea

Severe fever with thrombocytopenia syndrome (SFTS) is an emerging infectious disease in Asia area from 2013. The major clinical symptoms of SFTS are fever, vomiting, diarrhea, multiple organ failure, thrombocytopenia , leucopenia and elevated liver enzyme levels, showing its fatality rates ranging from 12% to as high as 30%. SFTS virus is a phlebovirus in the family of *Bunyaviridae*, and consist of 3 gene segments, large (L), medium (M) and small (S). And 6 proteins have been identified—an RNA dependent RNA polymerase (RdRp), a glycoprotein precursor (M), a glycoprotein N (Gn), a glycoprotein C (Gc), a nuclear protein (NP) and a non structural protein (NSs). In this study, using immunoproteomics approach coupled with high resolution LC-MS platform, several highly immunogenic SFTSv antigens were discovered. Those antigens are expected to be applied in vaccine development and rapid detection kit development. Also, immunoproteomic approach shows its possibility as a useful tool for antigen discovery.

946

LIN28A interactome analysis illustrates its diverse gene regulatory functions

Nam-Kyung Yu[1], Daniel B McClatchy[1], Sarah Romero[2], Alysson Muotri[2], John R Yates[1]

1. The Scripps Research Institute, San Diego, CA, United States

2. University of California San Diego, San Diego, California, United States

Publish consent withheld

947

In-depth serum proteomics reveals biomarkers of psoriasis severity and response to traditional Chinese medicine

Meng Xu[1], Jingwen Deng[2], Kaikun Xu[3], Tiansheng Zhu[4], Ling Han[2], Yuhong Yan[2], Danni Yao[2], Hao Deng[2], Dan Wang[3], Yaoting Sun[4], Cheng Chang[3], Xiaomei Zhang[3], Jiayu Dai[3], Liang Yue[4], Qiushi Zhang[4], Xue Cai[4], Yi Zhu[4], Hu Duan[3], Yuan Liu[3], Dong Li[3], Yunping Zhu[3], Timothy R.D.J. Radstake[5], Deepak M.W. Balak[6], Danke Xu[1], Tiannan Guo[4], Chuanjian Lu[2], Xiaobo Yu[3]

1. State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, China

2. Guangdong Provincial Hospital of Chinese Medicine, Guangdong Provincial Academy of Chinese Medical Sciences, Guangdong Provincial Key Laboratory of Clinical Research on Traditional Chinese Medicine Syndrome, Guangzhou, China

3. State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences, Beijing Institute of Lifeomics, Changping District,, BEIJING, China

4. School of Life Sciences, Westlake University; Institute for Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, China

5. Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands.

6. Department of Dermatology and Allergology, University Medical Center Utrecht, Utrecht, The Netherlands.

Background: Serum and plasma contain abundant biological information that reflect the body's physiological and pathological conditions and is therefore a valuable sample type for disease biomarkers. However, comprehensive profiling of the serological proteome is challenging due to the wide range of protein concentrations in serum.

Methodology: To address this challenge, we developed a novel in-depth serum proteomics platform capable of analyzing the serum proteome across ~10 orders or magnitude by combining data obtained from Data Independent Acquisition Mass Spectrometry (DIA-MS) and customizable antibody microarrays.

Results: Using psoriasis as a proof-of-concept disease model, we screened 50 serum proteomes from healthy controls and psoriasis patients before and after treatment with traditional Chinese medicine (YinXieLing) on our in-depth serum proteomics platform. We identified 106 differentially-expressed proteins in psoriasis patients involved in psoriasis-relevant biological processes, such as blood coagulation, inflammation, apoptosis and angiogenesis signaling pathways. In addition, unbiased clustering and principle component analysis revealed 58 proteins discriminating healthy volunteers from psoriasis patients and 12 proteins distinguishing responders from non-responders to YinXieLing. To further demonstrate the clinical utility of our platform, we performed correlation analyses between serum proteomes and psoriasis activity and found a positive association between the psoriasis area and severity index (PASI) score with three serum proteins (PI3, CCL22, IL-12B).

Conclusion: Taken together, these results demonstrate the clinical utility of our in-depth serum proteomics platform to identify specific diagnostic and predictive biomarkers of psoriasis and other immune-mediated diseases.

948

Fully Automated Protein Absolute Quantification Platform for Simultaneous Analysis of Multiple Tumor Markers in Human Plasma

Huiming Yuan[1], Yuanbo Chen[1], Lihua Zhang[1], Yukui Zhang[1]

1. Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, LIAONING, China

The proteins in the circulatory system reflect the individual's physiological and disease states. However, many protein biomarkers used for clinical diagnosis are present at very low concentration in body fluids and the detection is very difficult. In this study, we developed a fully automated protein abosolute quantification platform for simultaneous analysis of multiple tumor markers in human plasma, by which alpha-fetoprotein (AFP) , prostate-specific antigen (PSA), carcino-embryonic antigen (CEA) and mucin-1(MUC-1) were first simultaneously enriched from human plasma by aptamers immobilized capillary column with polymer@graphene oxide @polyethylenimine@Au (PM@GO@PEI@Au) modified microspheres as the matrix, followed by on-line protein denaturation, reduction, desalting and tryptic digestion by a fully automated sample treatment (FAST) device, finally the resulting peptides were analyzed by pararell reaction monitoring on a LTQ-orbitrap velos mass spectrometry. Compared to traditional ELISA assay widely used in routine clinical practice, such a platform exhibited significant advantages such as short analysis time (2 h vs. 12 h), low limit of detection (for AFP : 5 fg/μL vs. 46 fg/μL, and for PSA : 50 fg/μL vs. 69 fg/μL), and ease of automation. Furthermore, our developed platform was also applied in the absolute quantification of tumor markers from clinical plasma samples, and the results were comparable to those obtained by clinical immunoassay. All the results demonstrated that our developed fully automated platform should provide a promising tool for achieving high sensitivity, high accuracy, and high throughput detection of tumor protein markers in the routine physical examination and clinical disease diagnosis.

949

Multistage Native MS Enables Direct Identification of Unknown Ligands Bound to Protein Assemblies

Joseph F Gault[1], Idir Liko[2], Denis Shutin[1], Rosa Viner[3], Romain Huguet[3], Helene Cardasis[3], John EP Syka[3], Jesse D Canterbury[3], Phillip M Remes[3], Graeme McAlister[3], Hsin-Yung Yen[2], Vlad Zabrouskov[3], Michael Landreh[4], Christopher Mullen[3], Carol Robinson[1]

1. Oxford University, Oxford, UK

2. Omass Therapeutics , Oxford, UK

3. ThermoFisher Scientific, San Jose, CA, United States

4. Karolinska Institutet, Stockholm, Sweden

All protein complexes interact with ligands during their life cycle. Identifying endogenous ligands is necessary to understand the molecular details of the complex cellular processes, but current methodologies almost universally rely on ligand extraction and separation, and thus sever the link between binding partners. This makes multiple interactions difficult to capture and unknown orphan ligands extremely challenging to identify. In this work, we combined native and omics-based mass spectrometry, to showcase a new method to identify unknown ligands directly bound to human proteins and therapeutic targets. We used multistage fragmentation to progressively dissect protein-ligand assemblies, which culminates in ligand identification via database searching. We showed that for mixtures of proteoforms and/or ligands, it is not possible to identify endogenous ligands bound to proteins either by mass measurement alone or by application of other high-resolution structural techniques such as cryoEM. To realise our workflow, we used an Orbitrap Eclipse mass spectrometer which enables transmission of high *m/z* ions, MSn isolation, fragmentation and simultaneous manipulation of high and low *m/z* species. Precursor ions up to *m/z* 8,000 can be routinely coupled with HCD/CID and UVPD to perform multiple stages of MS. We assessed the performance of our Multistage-nMS workflow using a variety of model proteins systems, bound to an increasingly complex cohort of endogenous ligands. We showed that performing MSn experiments in both negative and positive polarity was necessary to capture the full complement of bound species, including different families of lipids. Finally, we apply and extend our Multistage-nMS approach to identify previously unknown ligands bound to protein assemblies. The Orbitrap Eclipse capabilities offer tremendous potential to reveal how proteins

assemblies are regulated in vivo, discover previously unknown endogenous ligands or signalling pathways, and to develop new avenues for novel therapeutics.

95

The degrading business: Measurement of proteome turnover in intact animals

Robert J Beynon[1], Dean E Hammond[1], Deborah M Simpson[1], Mark Prescott[1], John Waters[1], Jane L Hurst[1], Edward Lau[2]

1. Institute of Integrative Biology, University of Liverpool, Liverpool, Merseyside, United Kingdom

2. Cardiovascular Institute, Stanford University, Stanford, CA, United States of America

A steady state proteome is changing rapidly, as proteins are synthesised and degraded at rates that balance each other. Changes in protein abundance are caused by an imbalance in these opposing processes. Thus, a full definition of the proteome must include an understanding of the rate at which any protein is replaced. In a steady state, this can only be attained by monitoring the behaviour of a tracer, and at the proteome level, this means stable isotope precursors. These can be simple (e.g. deuterated water or 15N NH_4Cl) or more complex (stable isotope labelled amino acids). The labelling strategies, isotope incorporation patterns and treatment of the kinetic data vary with the labelling protocol. For systematic comparison of two labelling methods we have taken two cohorts of identical, inbred C57BL/6J mice and measured protein turnover rates in four tissues, using either $[^2H_2]O$ (heavy water, 'HW', supplied in drinking water) or $[^{13}C_6]$ lysine (amino acid, 'AA', provided as a supplement to normal laboratory chow). We have then compared the measurement of turnover using the two methods, asking 'do we obtain the same values?'.

Tryptic digests were analysed by LC-MS/MS and data were processed using Crux, and label incorporation using open source software RIAna, (written by EL, see associated poster). The major difference between the two labelling protocols lies in the speed at which the precursor equilibrates within each tissue, and this delay can lead to distortion in the determination of the rate of turnover, particularly for high turnover proteins. However, the precursor pool labelling trajectory is measurable using GC-MS of plasma for HW, and analysis of di-lysine peptides for AA. Correction for precursor pool labelling leads to rate constants that are very similar between the two methods. Thus, either approach can yield proteome-wide turnover measurements with high accuracy and precision.

950

Differential Plasma Proteomic Analysis of ß-thalassemia patients in response to Hydroxyurea treatment

Muhammad Zohaib[1], Saqib H Ansari[2], Tahir Shamsi[2], Roman A Zubarev[3], Shamshad Zarina[1]

1. National Center for Proteomics, University of Karachi, Karachi, SINDH, Pakistan

2. National Institute of Blood Diseases and Bone Marrow Transplantation , Karachi, Sindh, Pakistan

3. Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

ß-thalassemia is the most common genetic disorder caused by mutations in β-globin gene resulting in absence/reduced production of Adult Hemoglobin (HbA). Regular blood transfusions and effective iron chelation is necessary for survival of β-thalassemia patients. An emerging therapeutic approach to handle β thalassaemia is production of fetal hemoglobin (HbF). Hydroxyurea (HU), a potent HbF inducing agent, has shown to be an effective drug for β-thalassemia management with variable response from transfusion independence to null outcome. Diversified response of β-thalassemia patients to HU therapy remains ambiguous for hematologist. Clinical proteomics has revolutionized the study of differential protein expression associated with disease and offers a unique technique to monitor the response to treatment. In current study, we focused on comparative analysis of plasma proteome in pre- and post- HU-treated β-thalassemia patients, as well as responders and non-responders to HU treatment. Plasma was collected from β-thalassemia patients before and after 6 months of HU treatment, and the treated group were sub-categorized on the basis of response to HU. Label-free LC-MS/MS analysis was performed on nano-HPLC coupled Q-Exactive Orbitrap Plus mass spectrometer. Quantitative data and statistical analysis was carried out using MaxQuant and Perseus respectively. Proteomics analysis revealed identification of 400 proteins in all groups; among them the expression of twenty eight proteins were significantly different in pre- vs post- HU treated groups, with, twenty six proteins being differentially expressed in responder vs non-responder groups. Transferrin receptor protein 1 (TfR) was the most down-regulated and hemopexin and haptoglobin were the most up-regulated proteins in plasma after HU treatment whereas proteins that differed significantly among responder and non-responder were carbonic anhydrase 1, hemoglobin subunit gamma-1 and peroxiredoxin-2.

951

Structural characterizations of intact monoclonal antibodies by native MS

Angela Criscuolo[1, 2, 3], Tabiwang N. Arrey[1], Eugen Damoc[1], Thomas Möhring[1], Catharina Crone[1], Martin Zeller[1], Markus Kellmann[1]

1. Thermo Fisher Scientific (Bremen) GmbH, Bremen, HB, Germany

2. Institute of Bioanalytical Chemistry, Faculty of Chemistry and Mineralogy, University of Leipzig, Leipzig, Germany

3. Center for Biotechnology and Biomedicine, University of Leipzig, Leipzig, Germany

In the production of therapeutic monoclonal antibodies (mAbs) intact mass spectrometry analysis has become a key step. In particular, the characterization of PTMs is a fundamental regulatory requirement as these modifications may affect quality, efficacy and safety of these biopharmaceuticals. These MS analyses can be performed under denaturing or native conditions. The main advantage of native MS is to avoid acid and organic solvents. The protein, which then can maintain a three dimensional structure, will accept less charges during the ionization process with an improvement in terms of spectral resolution. Most

commonly, these analyses are performed by size exclusion chromatography (SEC) coupled with MS. Alternatively, charge variant analysis by ion exchange chromatography (IEC) has shown is potential to reveal mAb heterogeneity.

Preliminary experiments were performed on Trastuzumab by size exclusion chromatography coupled to a modified quadrupole Orbitrap mass spectrometer. The protein was eluting in a sharp peak, about 20 sec at the baseline. The average spectrum over the elution time shows a nicely distributed envelope with charge state between 21 and 30. Zooming inside a single charge state revealed five major glycoforms baseline separated. Then, the datasets were analyzed using ReSpect algorithm in Thermo Scientific™ BioPharmaFinder™ 3.0 software. The deconvoluted observed masses matched very well to the average theoretical masses for the known amino acid sequence with various combinations of glycoforms (delta ppm < 7).

SEC experiments were compared with IEC, particularly cation exchange. Experiments conducted by Bailey et al. [1], Trappe et al. [2] and Fuessl et al. [3] report this new tool as a promising technique to complete structural elucidation of therapeutic monoclonal antibodies. This technique adds another dimension of separation, the retention time which is influenced by the chemical nature of the PTM. Consequently mAbs variants characterized by minimal mass difference may be now identified.

952

PhosMap: a comprehensive R package for analysing quantitative phosphoproteomics data

Dongdong Zhan[1, 2], Tianyang Zhao[3], Nairen Zheng[1], Mingwei Liu[1], Yi Wang[1], Jun Qin[1]

1. National Center for Protein Sciences (Beijing), Beijing, NOT US OR CANADA, China

2. East China Normal University, Shanghai, China

3. Beihang University, Beijing, China

Protein phosphorylation is a post-translational modification (PTM) that plays an important role in diverse cellular processes. Liquid chromatography-mass spectrometry (LC-MS)-based quantitative phosphoproteomics has allowed for that measurement of thousands of phosphorylation sites from biological specimens. However, there is a lack of bioinformatics tools to systematically analyse the complex data. Most of the current tools only provide a single function, which cannot be used to perform phosphoproteomics routine analysis effectively. Here, we describe PhosMap, a modular R package used to pre-process and analyse phosphoproteomics data. PhosMap supports data quality control at the phosphorylation site and peptide levels, allowing for mapping of phosphorylation sites to the corresponding protein sequence and normalization of phosphorylation based on protein abundance from proteomics data. The cleaned data can be analysed in four analysis modules: (1) clustering and differential expression analysis, (2) time course analysis, (3) kinase activity prediction and (4) phosphorylation motif enrichment analysis. In addition, a user-friendly and compatible visualization module is embedded in PhosMap for more intuitive data visualization. The PhosMap package is compatible with R 3.6 and above on Windows, Mac OS X and Linux and freely-available on GitHub: https://github.com/ecnuzdd/PhosMap.

953

Quantitative proteomics, a missing part of proteomic research

jiandi Zhang[1]

1. Quanticision Diagnostics, Inc, RTP, NC, United States

The proteomic research at present is crippled in certain sense, as it focuses on analyzing thousands of proteins over a few samples. An equally important part of proteomic research where a few proteins are examined over thousands of samples (quantitative proteomics), is missing. We introduced here Quantitative Dot Blot (QDB) analysis, an immunoassaywhich can achieve absolute quantitation of a specific protein over thousands of samples in high throughput format. The application of this technique in clinical diagnosis has achieved unprecedented success, where the subtyping of breast cancer patients, which relies traditionally on genetic analysis, is realized with the 3D distribution of three key protein biomarkers of Estrogen receptor (ER), Progesterone Receptor (PR) and Her2 as absolute and continuous variables. Continuing research on quantitative proteomics should have significant impact on the practical side of precision medicine, and it would also complement the current proteomic research to present a whole picture of proteomics in the future.

954

The new strategy for co-immunoprecipitation procedure to improve protein-protein interaction study

Keren Zhang[1], Yuxing Zhang [1], Yan Ren[1], Siqi Liu[1]

1. Clinical Laboratory of BGI Health, BGI, Shenzhen, Beishan Industrial Zone 11th Building, Yantian District, Guangdong, China

During co-immunoprecipitation (Co-IP), the antibodies that can specifically recognize the target proteins, termed as IP-antibodies, are necessary because they have special binding capacity distinct from WB-antibodies that binds to denatured proteins. To seize the transient and weak interactions in a complex, cross-linking with small molecules to some residues in proteins is often adopted, which enables fixation of the interacted structure between target proteins and their bound partners. According to traditional approach, a protein extraction is treated by a cross-linking reagent, then the IP-antibody is added to capture the cross-linked target protein and its interaction proteins. In such experiments, a key issue is how to gain a functional IP-antibody. Because of time- and time-dependence in IP-antibody preparation, Co-IP often encounters a technique barrier lack of proper antibodies. Herein, we raise a question whether WB-antibodies could be used in Co-IP study.

The MS-cleavable linker DSSO was taken to stabilize the interaction complex. The treated protein complexes were denatured by 8M urea to facilitate epitope exposure followed by incubation with WB-antibody. In A549 cell lysate, 22 GAPDH linked proteins were identified with WB-antibody batch, in which most of them are known GAPDH binding proteins. A total of 17 GAPDH linked proteins were found with IP-antibody batch, 5 of them overlapped with the binding components detected by WB-antibody.

Surprisingly, number of binding proteins captured by IP-antibody had not reported yet as the interacted proteins with GAPDH. Moreover, one of interacted proteins enriched in WB-antibody batch possessed the transport function, which was consistent with GAPDH functions involved in the vesicle transport.

Thus the preliminary results supported our hypothesis that WB-antibody could be used for identification of protein components upon Co-IP. The strategy might be useful for identifying the binding proteins with low-affinity and at low-abundance.

955

20S Proteasome Complex Structure Conformation and Dynamics Study by Hydrogen Deuterium Exchange Mass Spectrometry

terry zhang[1], Shaunak Rava[2], Vlad Sarpe[3], Rosa Viner[1], Albert Konijnenberg[4], David Schriemer[2], Andrease Huhmer[1]

1. ThermoFisher Scientific, San Jose, CA, United States

2. Department of Chemistry, Univerisity of Calgary, Calgary, Alberta, Canada

3. Department of Biochemistry, Univerisity of Calgary, Calgary, Alberta, Canada

4. MSD, ThermoFisher Scientific, Eindhoven, The Netherlands

Proteasome is a large protein complex responsible for protein degradation in mammalian cells. It is important to determine the protein complex structure for better understanding of its catalytic function. There are multiple analytical methods could be used to access the protein structure [1]. In recent years HDX-MS has emerged as a powerful tool for studying protein, protein complex's conformation and conformation dynamics. In this presentation, HDX-MS was used to study the structure conformation of 20S proteasome complex in presence/absence of SDS.

Rabbit 20S proteasome protein complex and protein complex with 0.1% SDS samples were labeled with deuteration buffer and incubated for multiple time points. The deuterated samples were then quenched and digested online with a pepsin column in a fully automated manner. The digested peptides were injected into a reverse phase column with a short gradient. MS analysis was performed with Thermo Scientific™ Orbitrap Lumos™ mass spectrometer. Peptides mapping and PTM analysis were performed with Thermo Scientific™ BioPharma Finder 3.1. HDX experimental data were analyzed by BioPharma Finder 3.1, Mass Spec Studio (Dr. Schriemer, University of Calgary).

The 20S proteasome complex contains 28 subunits arranged into four stacked rings: seven alpha non-catalytic subunits and seven beta-subunits [2]. All seven alpha and seven beta-subunits were identified with sequence coverage range from around 79% to 93%. The identified peptides were used to calculate the deuterate incorporation of the proteasome with/without SDS in two protein states. The deuterium incorporation varied indicating the proteins conformed with different solvent associability. Overall, the presence of SDS caused most of the protein regions to have more of deuterium incorporation. The protection factors were calculated by BioPharma Finder 3.1 which reviewed the protein's solvent accessibility on amino acid level. Structural information obtained by HDX-MS for 20S proteasome complex dynamics was validated using structure obtained through electron microscopy.

1. Greber, B.J. et al, Science, 348, 303-308, 2015
2. Van de Waterbeemd, M. et al, Nat.Methods, 14, 283-286, 2017

956

Early changes in the urine proteome in a rat liver tumor model

Yameng Zhang[1], Yufei Gao[2], Youhe Gao[1]

1. Department of Biochemistry and Molecular Biology,College of Life Sciences, Beijing Normal University, Beijing, China

2. College of Information Science and Technology, Beijing Normal University, Beijing, China

Urine, as a potential biomarker source among the body fluids, can accumulate many changes in the body due to the lack of a mechanism to maintain a homeostatic state. Previous studies have demonstrated that proteomic technology can find many potential biomarkers to reflect different diseases in the urine. This study aims to detect early changes in the urinary proteome in a rat liver tumor model. The tumor model was established with the Walker-256 carcinosarcoma cell line (W256). Compared to before the injection, ninety-five differential proteins were significantly changed in the experimental rats. At day 3, twelve proteins were identified in the absence of pathological changes, and four of them were altered at all four time-points (B2MG, VCAM1, HA11, and LG3BP). Seven had previously been associated with liver cancer. At day 5, fifty-two differential proteins were identified. At day 7 and day 11, there was a significant decrease in the body weight of the rats, and tumor tissue was observed in the liver. Fifty-two and forty differential proteins were changed significantly at day 7 and day 11, respectively. Of the proteins that were identified at these three time-points, and twenty-four were reported to be associated with liver cancer. Comparing the differential urinary proteins and biological processes of liver tumor model with those in different models of W256 grown in other organs, specific differential protein patterns were found among the four models, which indicates that the differential urinary proteins can reflect the differences when the same tumor cell grown in different organs. This study demonstrated that (1) the rat liver tumor model caused early changes in urinary proteins may give new insight into the early diagnosis of liver cancer; (2) the same tumor cell grown in different organs can be reflected in differential urinary proteins.

957

Large Scale Preoteome Study on Peritoneal Dialysis Effluent Exosomes

Yan Tong[1], Yi Zhang[2], Jialing Li[2], Huiping Zhang[2], Hongqiang Ruan[2], Feng Ding[1], Yingli Liu[1]

1. nephrology, Shanghai Nine People's Hospital, Shanghai Jiao Tong Univrsity School of Medicine, Shanghai, China

2. APT, Minhang District, Shanghai, CHINA, China

Peritoneal dialysis (PD) is well recognized as an effective renal replacement therapy to remove metabolic waste in blood and eliminate excessive liquid in dialysis patients with end-stage renal disease (ESRD). The efficiency of PD depends on the structural and functional integrity of the peritoneum. The assessment of the functionality of the peritoneum is usually conducted by investigating the transport of small solutes and fluid. Many studies have shown an association between high transport status and poor patients or technique survival. As a so called "liquid biopsy", esosomes are attractive sources of biomarker in PD effluent (PDE). However, their proteome has not been widely studied. Here we optimized the isolation by differential centrifugation, and applied the SP3 digestion and data independent acquisition (DIA) label free quantitative technology for PDE exosomes proteome study. Through the large scale proteome study, over 2500 proteins have been chareacterized and quantified in PDE exosomes, including many marked proteins with known roles in peritoneal pathophysiology, such as AQP1 (Aquaporin1) and NHE (sodium/hydrogen exchanger). These two proteins play an important role in the process of transmembrane water transport, which can affect the ultrafiltration and prognosis of PD patients. Bioinformatic analysis shoe lots of candidates between high transport and low transport with ESRD. This study suggests lots of potiential biomarkers for peritoneal dialysis and function mechanism for high transport of peritoneum.

958

A combination approach to explore the missing proteins at low abundance with higher hydrophobicity and lower molecular weight

yuanliang zhang[1], zhilong lin[1], yan ren[1], siqi liu[1]

1. BGI, Shenzhen, NOT US OR CANADA, China

In the list of missing proteins (MPs) released in 2019, the proteins with higher hydrophobicity and/or with lower molecular weight occupy at the two top rankings. With MPs being shrunk, a question is naturally raised how to identify a MP class that exhibits low abundance in cells or tissues but possesses high hydrophobicity and/or low molecular weight. A further inquiry is which factor is decisive in technology for discovery of such MPs. Herein, we proposed a combination strategy that was likely as a partial solution for identification of those MPs. The membrane fractions from four cancer cell lines were isolated by ultracentrifuge and the correspondent proteins were equalized in their abundance by ProteoMiner kit. After tryptic digestion of the treated proteins, the resulted peptides were re-extracted by high concentration organic solvent and the extracted peptides were delivered to a mass spectrometer of Orbitrap Fusion™ Lumos™ Tribrid™ for acquiring MS/MS data with high resolution. Based upon the HPP guidelines v 2.10, a total of 16 MPs with 61 unique peptides,in which a MP was defined 2 non-nested unique peptides with ≥9 amino acids at least, were identified in the four cancer cell lines, 2, 8, 2 and 7 MPs in HeLa, HCT116, SNU-1 and HepG2 cells, respectively. The two non-nested unique peptides responding to the individual MPs were further verified through parallel reaction monitoring based upon thechemically synthesized peptides as the standards of retaining time and MS/MS overlay. The combination strategy was thus proven effective in MP discovery.

959

A urinary metabolomics study of colorectal cancer diagnosis and metastasis markers

Jun Li[1], Xiaoyan Liu[2], Yan Sun[1], Jing Li[2], Yan Zhao[1], Xiang Liu[2], Yulin Sun[1], Zhengguang Guo[2], Wei Sun[2], Xiaohang Zhao[1]

1. State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

2. Core Facility of Instruments, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Colorectal carcinoma (CRC) is one of the most common cancers in the world, and approximately 35%-55% of patients with CRC will develop hepatic metastases while their disease. Currently, there are no accurate predictive markers for metastatic CRC. Urine, a kind of promising biofluids, is characterized by its ease of collection, richness in metabolites and its ability to reflect imbalances of all biochemical pathways within the body. To identify stable and reproducible biomarkers for CRC metastasis in noninvasively collected urine, a comprehensive urinary metabonomics analysis of CRC was conducted using the liquid chromatography/high-resolution mass spectrometry (LC-HRMS) strategy. In which, a total 205 urine samples from CRC patients and healthy controls including CRC without metastasis (CRC-NM, n=30), CRC with lymph nodes metastasis (CRC-LNM, n=57) and CRC with liver metastasis (CRC-LM, n=58), as well as healthy controls (HC, n=60) were analyzed at the first stage. Compared to the healthy controls, 40, 50 and 68 kinds of differentiated urinary metabolites were identified in CRC-NM, CRC-LNM and CRC-LM groups. We initially established a urinary metabolomic panel of CRC metastasis containing 4 kinds of urinary metabolites which could distinguish CRC without metastasis and healthy controls both in the training set (Area Under the ROC Curve, AUC=0.938) and in the validation set (AUC=0.898), respectively. Simultaneously, some metabolites could also distinguish CRC-LNM or CRC-LM and HC in both training (AUC=0.882 or 0.937) and validation (AUC=0.818 or 0.882) sets, respectively. The major metabolic pathways of CRC-NM, CRC-LNM and CRC-LM were found including linolenic acid, caffeine and tyrosine metabolism. We explored the urine metabolome by different MS-based techniques and the preliminary findings provided potentially diagnostic urinary biomarkers. The data needs to be validated in large samples sets in the future.

1. Siegel RL, Miller KD, Fedewa SA, et al. Colorectal cancer statistics, 2017. CA Cancer J Clin. 2017; 67(3):177-193.

2. Khamis MM, Adamko DJ, El-Aneed A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. Mass Spectrom Rev. 2017; 36(2):115-134.
3. Mathé EA, Patterson AD, Haznadar M, et al. Noninvasive urinary metabolomic profiling identifies diagnostic and prognostic markers in lung cancer. Cancer Res. 2014 Jun; 74(12):3259-70.
4. Deng L, Ismond K, Liu Z, et al. Urinary Metabolomics to Identify a Unique Biomarker Panel for Detecting Colorectal Cancer: A Multicentre Study. Cancer Epidemiol Biomarkers Prev. 2019 May 31. pii: cebp.1291.2019. doi: 10.1158/1055-9965.EPI-18-1291.

960

An integrated MS data processing strategy for fast identification, in-depth and reproducible quantification of protein O-glycosylation in large cohorts of human urine samples

Xinyuan Zhao[1], Weijie Qin[1], Xiaohong Qian[1]

1. National Center for Protein Sciences Beijing, State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Lifeomics, Changping, Beijing, China

Protein O-glycosylation has long been recognized to be closely associated with many diseases, particularly with tumor proliferation, invasion and metastasis. The ability to efficiently profile the variation of O-glycosylation in large-scale clinical samples provides an important approach for the development of biomarkers for cancer diagnosis and for therapeutic response evaluation. Therefore, mass spectrometry (MS)-based techniques for high throughput, in-depth and reliable elucidation of protein O-glycosylation in large clinical cohorts are in high demand. However, the wide existence of serine and threonine residues in the proteome and the tens of mammalian O-glycan types lead to extremely large searching space composed of millions of theoretical combinations of peptides and O-glycans for intact O-glycopeptide database searching. As a result, exceptionally long time is required for database searching which is a major obstacle in O-glycoproteome studies of large clinical cohorts. Furthermore, due to the low abundance and poor ionization of intact O-glycopeptides and the stochastic nature of data-dependent MS2 acquisition, substantially elevated missing data levels are inevitable as the sample number increases, which undermines the quantitative comparison across samples. Therefore, we report a new MS data processing strategy that integrates glycoform-specific database searching, reference library-based MS1 feature matching and MS2 identification propagation for fast identification, in-depth and reproducible label-free quantification of O-glycosylation of human urinary proteins. This strategy increases the database searching speeds by up to 20-fold and leads to a 30-40% enhanced intact O-glycopeptide quantification in individual samples with an obviously improved reproducibility. In total, we obtained quantitative information for 1068 intact O-glycopeptides across 36 healthy human urine samples with a 30-40% reduction in the amount of missing data. This is currently the largest dataset of urinary O-glycoproteome and demonstrates the application potential of this new strategy in large-scale clinical investigations.

961

A circulating exosomes-based novel screening tool for colorectal cancer revealed by Shotgun and Data Independent Acquisition mass spectrometry

Xi Zheng[1], Kailun Xu[1], Yingkuan Shao[1], Biting Zhou[1], Ting Chen[1], Yanqin Huang[1], Fei Wen[1], Weiting Ge[1], Jian Wang[1, 2], Shaojun Yu[2], Liang Yue[3, 4], Xue Cai[3, 4], Qiushi Zhang[3, 4], Guan Ruan[3, 4], Tiansheng Zhu[3, 4], Yi Zhu[3, 4], Tiannan Guo[3, 4], Shu Zheng[1]

1. Cancer Institute(Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education), the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

2. Department of surgical oncology, the Second Affiliated Hospital, Zhejiang University School of Medicine , Hangzhou, Zhejiang, China

3. School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

4. Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China

Background: Early screening for colorectal cancer (CRC) is essential in order to improve prognosis. Liquid biopsies are increasingly being considered for diagnosing cancer due to the low invasiveness and reproducibility. In addition, circulating exosomes (crExos) expressing tumor-specific proteins are potential biomarkers for various cancer types. Here, we developed an optimized DIA-MS-based diagnostic method in liquid biopsies, which can be potentially adapted as a rapid and non-invasive screening tool to detect early stages of CRC.

Methodologies: Exosomes were isolated from the culture supernatants of human CRC cells lines, and the plasma of CRC patients at different tumor stages, by prolonged ultracentrifugation combined with sucrose density gradient centrifugation. Tumor-specific exosomal proteins were identified using Tandem Mass Tag (TMT)-based shotgun proteomics and phosphoproteomics. The results were verified on a second independent cohort and a mouse tumor-bearing model using Western blotting, and the candidate biomarkers were further validated in a third CRC cohort by Data Independent Acquisition (DIA)-mass spectrometry (MS). Finally, the DIA-MS methodology was optimized to permit high-throughput detection of exosome biomarkers in another independent cohort of CRC patients and healthy controls.

Findings: High levels of total and phosphorylated exosomal fibronectin 1 (FN1), haptoglobin (HP), S100A9 and fibrinogen α chain (FGA) were significantly associated with tumor progression, and FGA was the most dominant biomarker. Analysis of the human CRC cell lines and the mouse model indicated that FGA+ crExos were likely released by the CRC cells. Furthermore, optimized DIA-MS validated that FGA+ crExos could distinguish colon adenoma and CRC patients from healthy individuals with higher specificity and sensitivity than the conventional tumor biomarkers.

Conclusion: DIA-MS detected proteome of crExos is a potential non-invasive screening tool to identify early stage of CRC.

962

Coupling subcellular fractionation and basic reverse phase chromatography to achieve improved proteomic coverage

Rebekah Ziegman[1], Keshava K. Datta[1], Harsha Gowda[1, 2, 3]

1. QIMR, Herston, QLD, Australia

2. School of Medicine, University of Queensland, Herston, QLD, Australia

3. School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia

One of the main goals of global proteomic profiling studies is to achieve sufficient depth to identify and/or quantify all the proteins expressed in a cell or tissue type. LC-MS/MS analysis of tryptic digests from whole cell lysates often shows better coverage for high abundant proteins compared to low abundant proteins. Pre-fractionation methods using chromatography have been widely used to achieve greater depth in proteomics. However, these strategies often show greater coverage for most cytosolic proteins that are more abundant than nuclear proteins including transcription factors that are relatively less abundant. We reasoned that coupling a simple ultracentrifuge independent subcellular fractionation strategy with bRPLC based fractionation will provide better coverage of both cytosolic and nuclear proteins.

A simple method utilizing osmotic lysis and centrifugation was performed on human melanoma cell lines in order to separate them into nuclear and cytoplasmic fractions. These, along with whole cell lysates of the same cell lines were digested using trypsin and fractioned by bRPLC. 10% of each fraction was used for global proteome profiling and the remaining 90% was used for TiO2 based phosphopeptide enrichment. All the samples were subjected to LC-MS/MS analysis on Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer. MS/MS search was carried out using SEQUEST search algorithm in Proteome Discoverer software platform.

Coupling subcellular separation and bRPLC fractionation provided better proteome and phosphoproteome coverage for both cytosolic and nuclear proteins. The strategy also enabled identification of more peptides per protein increasing the confidence in these identifications. Better coverage was observed for low abundant proteins including transcription factors that are often underrepresented in global proteomic datasets. This study demonstrates a relatively easy strategy that can be employed for achieving better proteome and phosphoproteome coverage in global proteome profiling studies.

963

Investigating Pierce's Disease Tolerance mechanism in Grape using Shotgun Proteomics

Sedigheh Shokri[1, 2], Ramesh Katam[2], Varshini Sridhar[2], Joseph Bundy[3], Alireze Babaei[1]

1. Department of Horticulture Sciences, College of Agriculture, Tarbiat Modares University, Tehran, Iran

2. Department of Biological Sciences, Florida A&M University, Tallahassee, FL, USA

3. College of Medicine, Florida State University, Tallahassee, Fl, United States

Pierce's disease (PD), caused by bacterium *Xylella fastidiosa*, seriously hampers the cultivation of *Vitis vinifera* (bunch grapes) worldwide. The bacterium clogs xylem vessels and forms a biofilm, resulting in the wilting of the plant. *Vitis* species, such as, Florida hybrid bunch (FH) and muscadine grape (*Vitis rotundifolia)* are cultivated in United States, and are known for their tolerance to PD. The overall goals of this study are to determine the signal sequences associated with xylem and sap for the delivery of therapeutic proteins to control *Xylella fastidiosa.* The objectives of this research project are: 1) to compare the proteome profiles of xylem tissue and xylem sap from PD tolerant and -susceptible grapevine cultivars, and 2) to determine the role of proteins in the tissue and sap associated with PD tolerance mechanism. In this study, we used Bunch, FH, and Muscadine grape cultivars to characterize differentially expressed and unique proteins using LC MS/MS spectrometry searched against *Vitis* database. A total of 2519 and 402 proteins were identified in xylem and sap respectively, of which 151 proteins were common to both tissues. The cluster dendrogram analysis of the sap proteome showed that all of the *Vitis* species are bifolious. Florida hybrid bunch and muscadines are more closely related to each other than to bunch grape. Functional analysis and gene ontology revealed that proteins involved in carbohydrate metabolic process are more abundant in bunch grape, while FH and muscadine grape have more defense related proteins. Proteins involved in the defense and peroxidase activity are abundantly present in xylem and sap of FH and muscadine, and these proteins are relatively in reduced levels in bunch xylem and sap. Together, our findings highlight the possible roles of the identified unique proteins that confer PD tolerance to Florida hybrid bunch and muscadine cultivars.

Identifying high-priority proteins across the human diseasome using semantic similarity

Edward Lau[1], Vidya Venkatraman[2], Jennifer E Van Eyk[2], Maggie Pui Yu Lam[3]

1. Stanford Cardiovascular Institute, Stanford University, Stanford, CA, United States

2. Advanced Clinical Biosystems Research Institute, Department of Medicine and The Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, United States

3. Department of Medicine, Division of Cardiology; Department of Biochemistry and Molecular Genetics, University of Colorado, Anschutz Medical Campus, Aurora, CO, United States

Identifying the genes and proteins associated with a biological process or disease is a central goal of the biomedical research enterprise. However, relatively few systematic approaches are available that provide objective evaluation of the genes or proteins known to be important to a research topic, and hence researchers often rely on subjective evaluation of domain experts and laborious manual literature review. Computational bibliometric analysis, in conjunction with text mining and data curation, attempts to automate this process and return prioritized proteins in any given research topic.

To this end, we developed a method to identify and rank protein−topic relationships by calculating the semantic similarity between a protein and a query term in the biomedical literature while adjusting for the impact and immediacy of associated research articles. We term the calculated metric the weighted copublication distance (WCD) and show that it compares well to related approaches in predicting benchmark protein lists in multiple biological processes. We used WCD to extract prioritized " popular proteins" across multiple cell types, subanatomical regions, and standardized vocabularies containing over 20,000 human disease terms. The collection of protein−disease associations across the resulting human "diseasome" supports data analytical workflows to perform reverse protein-to-disease queries and functional annotation of experimental protein lists. We envision that the described improvement to the popular proteins strategy will be useful for annotating protein lists and guiding method development efforts as well as generating new hypotheses on understudied disease proteins using bibliometric information.

1. Lau E, Venkatraman V, Thomas CT, Wu JC, Van Eyk J, Lam MP. Identifying high-priority proteins across the human diseasome using semantic similarity. J Proteome Res. 2018 Dec 7;17(12):4267-4278.