1

Integrative omics at the core of discoveries in virology

Ileana Cristea[1]

1. Princeton University, NJ, United States

The coexistence and coevolution of hosts with pathogens is intrinsic to our ecosystem. Pathogenic infections induce an array of changes in the hosts that are tightly linked to the progression of infection and establishment of disease. At the cellular level, this is reflected in alterations in host cell composition, organization, and ability to communicate with other cells. Thus, changes in the host proteome, metabolome, lipidome, and secretome have started to be recognized as signatures of infectious or disease states. These alterations function to either induce host defenses that counteract the infection or promote pathogen replication for spread of infection. Consequently, the discovery and characterization of these signature changes are essential for both understanding the biology of infection and identifying novel targets for therapeutic interventions. This presentation will highlight the value of advanced mass spectrometry-based proteomics for defining the dynamics of proteome organization and understanding mechanisms of cellular defense during viral infections. Examples will be given from our studies of spatial-temporal remodeling of subcellular organelles during infection. We will show how the integration of proteomics with lipidomics, live cell microscopy, mathematical modeling, and genetic knockouts has allowed us to discover a novel function for peroxisomes in the assembly of infectious particles. Additionally, the contribution of localization-dependent posttranslational modifications will be discussed, with a focus on the role of protein acetylation in cellular host defense against infection.

2

High-sensitivity phosphoproteomics reveals pathogenic signalling changes in diabetic islets

Sean J Humphrey[1, 2, 3], **Francesca Sacco**[4], **David James**[2, 3], **Matthias Mann**[1]

1. Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Munich, Germany

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

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

4. Department of Biology, University of Rome Tor Vergata, Rome, RM, Italy

The progressive decline of pancreatic beta cell function is key to the pathogenesis of type-2 diabetes. Protein phosphorylation represents an important mechanism controlling glucose-stimulated insulin secretion from the beta cell. Mass spectrometry (MS)-based phosphoproteomics now enables the global measurement of protein phosphorylation in an unbiased manner, making this an attractive means of studying pathogenic changes underlying disease. However, owing to technological limitations, studies involving minuscule amounts of protein have largely remained out of reach for deep phosphoproteome analysis. We recently made major improvements to our “EasyPhos” phosphoproteomics workflow, which collectively have improved the sensitivity of the method several-fold, and have also improved throughput, reduced sample handling and processing time, and minimized technical variability. Coupled with recent benchtop Orbitrap mass spectrometers employing enhanced ion optics, these methods now enable deep and reproducible phosphoproteomics studies from limited starting materials. Here, we applied these technologies to perform in-depth characterization of changes in signalling networks of pancreatic cells from various models, including islets isolated from diabetic mice, as well as from healthy humans. To obtain a comprehensive view of signalling network rewiring we combined our phosphoproteomic analysis with a characterization of proteome changes. Our analysis revealed drastic remodelling of the proteome and phosphoproteome, including numerous proteins involved in the control of insulin secretion, glucose uptake, and metabolism. This study highlighted a novel signalling axis contributing to beta cell failure, that we demonstrated is amenable to therapeutic intervention.

3

Large scale crosslinking mass spectrometry to map the yeast nuclear interactome

Tara K Bartolec[1], **Daniela-Lee Smith**[1], **Marc R 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 novel cross linkers with specialised mass-spectrometry (MS) methods to generate high-throughput and rich protein-protein interaction networks from complex samples. Application of XL-MS to the quantitative study of dynamic interactomes (for example across time, or between wild-type and gene knockout systems) will reveal new insights into cellular signalling and function. However, the large dynamic range of protein expression levels within eukaryotic cells presents a challenge in traditional proteomic, and furthermore XL-MS, studies; low abundance yet functionally important nuclear proteins are generally underrepresented. Here we will present the first large-scale yeast nuclear XL-MS protein interactome. Nuclei from actively dividing wild-type yeast were isolated by enzymatic cell wall digestion, mechanical disruption and sucrose density gradient centrifugation. Intact nuclei were then crosslinked using the cleavable crosslinker DSSO. Offline strong cation exchange gradients were optimised for crosslinked peptide enrichment and separation, alongside complementary online reverse phase liquid chromatography optimisation. A highly interconnected protein interactome was generated containing 2,091 crosslinks (identified at a 5% FDR) and representing 1,148 proteins. 1,625 crosslinks involved an interaction where at least one partner was annotated as nuclear, and 922 with both interactors as nuclear (representing 823 and 431 proteins, respectively). Over one third of all crosslinks could be matched to a known BIOGRID physical interaction, and clusters of known complexes including the RNA polymerase, chromatin structure remodelling (RSC) and preribosome complexes emerged within the network. This high depth dataset represents a significant foundation towards the broader goal of proteome-scale, quantitative XL-MS studies of the dynamic yeast nuclear interactome.

4

Multi-omic profiling of the liver in a rat model of type 2 diabetes

Desmond K Li[1, 2], Lauren E Smith[1, 2], Yen Chin Koay[2, 3], Holly McEwan[2, 4], Anthony Don[2, 4], John F O'Sullivan[2, 3], Stuart J Cordwell[5, 1, 2, 6], Melanie Y White[5, 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. Heart Research Institute, Sydney, NSW, Australia

4. ACRF Centenary Cancer Research Centre, Sydney, NSW, Australia

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

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

Altered glucose metabolism via 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, via elevated 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. To achieve this we performed a multi-omic analysis including proteomics, lipidomics and metabolomics in a rat model of T2D combining the effects of high fat diet feeding (calorie overload) and streptozotocin (elevated BGL). To quantify alterations in protein abundance, samples were isobarically tagged prior to mass spectrometry (MS). Discovery lipidomics was achieved with relative quantitation by comparison with synthetic standards. Targeted metabolomics was performed using multiple reaction monitoring, in the presence of deuterated metabolite standards. We quantified close to 7,000 proteins, 300 lipid species and 100 metabolites in the course of this study. Proteomics revealed increased levels of proteins regulating phospholipid biosynthesis and fatty acid metabolism. A concurrent decrease in proteins involved in steroid biosynthesis was observed. Lipid analysis show increased sphingomyelin levels and decreased levels of phosphatidylcholines in T2D, both of which are components of cell membranes and can play a role in metabolic and apoptotic signalling. Elevated levels of branched chain amino acids as well as changes in metabolites indicative of altered energy and amino acid metabolism were detected by metabolomics. The current study has identified changes in protein, lipid composition and metabolite levels indicative of dysregulated energy utilisation and molecular adaptations that contribute to the pathogenesis of T2D.

5

In-depth lipidomic profiling of the Australian imaging biomarker and lifestyle flagship study of aging

Kevin Huynh[1], Wei Ling Florence Lim[2, 3], Corey Giles[1], Kaushala S Jayawardana[1], Pratishtha Chatterjee[3, 4, 5], Natalie A Mellett[1], Ian Martins[2, 3], Simon M Laws[2, 6, 7], Christopher C Rowe[8], Ashley Bush[8], Victor L Villemagne[9, 10], David Ames[11], Colin L Masters[11], Brian G Drew[12, 1], Ralph N Martins[2, 3, 4, 5, 13, 14], Peter J Meikle[12, 1]

1. Baker Heart & Diabetes Institute, Melbourne, VIC, Australia

2. Cooperative research Centre (CRC) for Mental Health, Australia

3. School of Medical and Health Sciences, Edith Cowan University, Joondalup, , Western Australia, Australia

4. Department of Biomedical Sciences, Macquarie University, North Ryde, New South Wales, Australia

5. KaRa Institute of Neurological Disease, Sydney, Macquarie Park, New South Wales, Australia

6. Collaborative Genomics Group, School of Medical and Health Sciences, Edith Cowan University, Perth, Western Australia, Australia

7. School of Pharmacy and Biomedical Sciences, Faculty of Health Sciences, Curtin Health Innovation Research Institute, Curtin University, Western Australia, Australia

8. The Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Victoria, Australia

9. Department of Nuclear Medicine and Centre for PET, Austin Health, Heidelberg,, Victoria, Australia

10. Department of Medicine, Austin Health, The University of Melbourne, Heidelberg, Victoria, Australia

11. National Ageing Research Institute, Parkville, Victoria, Australia

12. Central Clinical School, Monash University, Melbourne, Victoria, Australia

13. School of Psychiatry and Clinical Neurosciences, The University of Western Australia, Perth, Australia

14. Australian Alzheimer's Research Foundation, Nedlands, Western Australia, Australia

Publish consent withheld

6

Development of integrative Protein and Lipid Organelle Profiling (iPLOP) method for high throughput organelle analysis

Harley R Robinson[1], Jeffrey Molendijk[1], Ahmed Mohamed[1], Alexandre Cristino[2], Michelle M Hill[1, 2]

1. QIMRBerghofer, Herston, QLD, Australia

2. UQ Diamantina Institute , Brisbane, QLD, Australia

Compartmentalization of constituents into distinct membrane bound organelles is an essential element of mammalian cells. Lipid metabolism and homeostasis are critical for organelle maintenance in health, and abnormalities in cellular lipids underlie

numerous chronic diseases. While methods of spatial proteomics are relatively well established, high throughput organellar lipidomics methods are largely unexplored. This study aims to develop a high-throughput method that enables proteo-lipidomic profiling of organelles, termed integrative Protein and Lipid Organelle Profiling (iPLOP). Starting on the same premise as PCP and hyperLOPIT methods previously reported for spatial proteomics, organelle identities are inferred from known marker proteins with a single known organelle localization. As most lipid species localize to several organelles at varying abundances, lipidomics aimed to quantify the lipid species at each organelle. 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. Computational analysis use correlation profiles and network classifiers to infer organelle proteome and lipidome. This presentation will report the development and proof of concept data for iPLOP. 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.

7

Defining the *Campylobacter jejuni* interactome by cross-linking mass spectrometry (XL-MS)

Ashleigh L Dale[1, 2], Dylan Harney[1, 2], Isobel Tennison[1, 2], Joel Cain[1, 2], Mark Larance[1, 2], Stuart Cordwell[1, 2, 3, 4]

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

2. Charles Perkins Center, Sydney, NSW, Australia

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

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 using DSSO 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 623 proteins were identified to partake in 826 unique and significant PPIs governed by 2,457 unique Lys-Lys residue contacts. The XL-MS approach successfully covered 38.4% of the predicted proteome of *C. jejuni* and 47.9% 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*.

8

In depth definition of the immunopeptidome in haematological malignancies and its potential for immunotherapy

Kirti Pandey[1], Nicole A Mifsud[1], Sri R Ramarathinam[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

Introduction – Haematological malignancy (HM) is an umbrella term used to refer to neoplasms originating from haematopoietic and lymphoid tissues. These clinically manifest as leukaemia, including Acute Myeloid Leukaemia (AML), which is characterised by an increase in blasts cells in the bone marrow. Despite recent successes in cancer immunotherapy involving checkpoint inhibitors and CAR T cell therapy in cancers including Hodgkin lymphoma and Melanoma, the success of these therapies is limited in HM. This makes it imperative to identify new targets for HM, which can be used in T-cell based vaccine therapies. The aim of this study was to define the HLA ligandome associated with AML and thereby identifying leukaemia-specific peptide antigens that include peptides derived from tumour-specific antigens and those bearing novel post translational modifications (PTM) involving phosphorylation and methylation. *Methods*- 1 X 10^9 cells of human AML cell line, THP-1 (HLA-A*02:01, -B*15:11, -C*03:03) were lysed and native HLA-peptide class I complexes purified by immunoaffinity chromatography and analysed using liquid chromatography and tandem mass spectrometry (LC-MS/MS). The experiment was performed using biological triplicates. *Results* – A total of 19,624 HLA-A*02:01 bound peptides along with 17,993 combined HLA-B*15:11 and HLA-C*03:03 bound peptides were identified from the triplicate dataset. Of note, many of the peptides were derived from pathways known to be dysregulated in AML, including Ras and Notch pathways. A number of tumour-specific or associated antigen derived peptides, including peptides from known oncogenes such as Retinoblastoma, A-raf and B-raf were found (n=6). Moreover, PTM peptides including phosphopeptides from upregulated or dysregulated signalling cascades (n=292) and methylated peptides (n=52) were also identified. *Conclusion* - This study identified naturally presented AML-specific and associated peptides, which will aid in development of improved diagnostic and patient stratification tools and novel peptide based immunotherapeutic approaches.

9

Lung cancer proteome biomarkers and in vitro diagnostics-multivariate index assay

Je-Yoel Cho[1]

1. Seoul National University, 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. 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. These 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. 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 based-on more sensitive assay will further improve the detection and diagnosis of lung cancers.

10

Spliced HLA peptides, a novel source for immunotherapy

Pouya Faridi[1], Patricia T. Illing[1], Chen Li[1], Sri H. Ramarathinam[1], Julian P. Vivian[1], Nicole A. Mifsud[1], Rochelle Ayala[1], Jiangning Song[1], Linden J. Gearing[2], Nicola Ternette[3], Jamie Rossjohn[1], Katherine Woods [4], Ralf Schittenhelm [1], Andreas Behren [4], Nathan P. Croft[1], Anthony W. Purcell[1]

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

2. Centre for Innate Immunity and Infectious Diseases, Hudson Institute of Medical Research and Department of Molecular and Translational Science, School of Clinical Science, Monash University, Melbourne, Victoria, Australia

3. Jenner Institute, University of Oxford, Oxford , UK

4. Olivia Newton-John Cancer Research Institute, Melbourne, Australia

Human leukocyte antigen class I (HLA-I) molecules sample the proteome following the degradation of intracellular proteins via the proteasome and other proteolytic mechanisms. These complexes of HLA-I and peptide (pHLA) are then recognized by T cells and the nature of the bound peptide ligand is the key driver of adaptive immunity. We have developed a sophisticated data-driven workflow for the identification of distal and proximal cis-spliced as well as trans-spliced (where distinct proteins contribute peptide segments) peptide antigens. We have shown for 17 different monoallelic cell lines, that up to 40% of pHLA are generated through post-translational splicing mechanisms. We also found that spliced peptides display canonical HLA-binding sequence features (1). To understand the biological relevance of spliced peptides, we have interrogated the immunopeptidome of HLA-I molecules immunoaffinity purified from influenza-infected cells using LC-MS/MS and the novel bioinformatic pipeline. We have found around 42% of influenza derived p-HLAs (83 out of 198 peptides) are spliced peptides. We have also applied our approach to p-HLA derived from melanoma cell lines and identified 38766 peptides of which 6% were cis-spliced and 21% trans-spliced. 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. Immunogenicity studies of a subset of the MAA shown that multiple of the tested spliced peptides were highly immunogenic, compared to a smaller fraction of linear peptides. Our results highlight the complexity and diversity of HLA peptide antigen presentation and identify a mechanism whereby the available peptidome is diversified to enhance immunity. 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.

1. P. Faridi et al., A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. Science Immunology 3, (2018).

11

1,566 prostate cancer proteomes generated by PCT-SWATH mass spectrometry reveal insights into patient survival and time to recurrence

Rebecca C Poulos[1], Rohan Shah[1], Tiansheng Zhu[2], Guobo Chen[3], Jelena Ljubicic[4], Peter Hains[1], Natasha Lucas[1], Yi Zhu[2, 5], Rutishauser Dorothea[4], Rui Sun[2], Hannes Roest[5], George Rosenberger[5], Janis Neumann[6], Konstantina Charmpi[6], Matteo Manica[7], Marija Buljan[5], Wenguang Shao[5], Guan Ruan[2], Niles Rupp[4], Daniel Schirmacher[5], Pedrioli Patrick[5], Maria Martinez[7], Andreas Beyer[6, 8], Asim Anees[1], Roger Reddel[1], Phillip Robinson[1], Peter Wild[4, 9], Tiannan Guo[2, 5], Ruedi Aebersold[5, 10], Qing Zhong[1, 4]

1. Children's Medical Research Institute, University of Sydney, Westmead, NSW, Australia

2. Westlake University, Hangzhou, China
3. People's Hospital of Hangzhou Medical College, Hangzhou, China
4. University Hospital Zurich, Zurich, Switzerland
5. ETH Zurich, Zurich, Switzerland
6. CECAD, University of Cologne, Cologne, Germany
7. IBM, Zurich, Switzerland
8. CMMC, University of Cologne, Cologne, Germany
9. University Hospital Frankfurt, Frankfurt am Main, Germany
10. University of Zurich, Zurich, Switzerland

People diagnosed with prostate cancer at an intermediate Gleason score can harbour aggressive or non-aggressive disease. These disease subtypes cannot always be accurately distinguished, and this can result in under- or over-treatment of some patients. In this study, we applied machine learning methods to proteomic data obtained from mass spectrometry to assess prostate cancer prognosis. We performed quantitative proteomic analyses of 1,566 prostate tissue samples from 290 patients obtained from the Prostate Cancer Outcomes Cohort (ProCOC). We processed small tissue samples via pressure-cycling technology (PCT) and acquired proteomic data by Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) mass spectrometry. Tumour tissues were classified by two pathologists, alongside matched benign tissue samples for each patient. We divided the cohort into 31 batches, each containing two control samples, allowing us to evaluate technical reproducibility. Each sample was analysed in technical duplicate at ProCan (Children's Medical Research Institute) to ensure high reproducibility. Altogether, we quantified 2,800 SwissProt proteins (FDR<1%) in all samples, with an average missing value of 30%. We quantified 73% of proteins in over 50% of samples. We corrected for batch effects and imputed missing values with technical and biological replicates. Using this dataset, we applied machine learning methods to find protein signatures to identify tumour and normal samples, and to assess Gleason scores. Using Random Forest, we could separate tumour and normal samples with an area under the curve of 0.92. We applied a random survival forest model to tumour samples and identified the top 100 proteins that predict survival, using time to recurrence and censoring information. To our knowledge, this is the largest SWATH-based proteomic dataset generated to-date in cancer (>1,500 proteomes). Our study demonstrates the feasibility of SWATH mass spectrometry for the proteomic analysis of prostate cancer prognosis.

12

Proteomic profiling of stage I – IV colorectal carcinoma specimen for early onset biomarker detection

Christoph Krisp[1], Sönke Harder[1], Susanne Burdak-Rothkamm[2], Gerrit Wolters[3], Jakob Izbicki[3], Hartmut Schlüter[1]
1. Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Hamburg Eppendorf, Hamburg, Germany
2. Institute of Molecular and Cytopathology , University Medical Ceter Hamburg Eppendorf, Hamburg, Germany
3. General, Visceral and Thoracic Surgery Department and Clinic, Medical Glycobiology Group, University Medical Center Haburg Eppendorf, Hamburg, Germany

Introduction - Colorectal cancer (CRC) is one of the most common cancer types in the world. It predominantly develops from benign polyps and is frequently caused by lifestyle choices. Early detection and treatment of the tumour results in 5-year survival rats of greater 95% however decreases dramatically to about 10% for stage IV. The availability of reliable biomarkers for CRC is sparse. Here we want to use mass spectrometry approach to profile fresh frozen CRC specimen from stage I-IV for early onset biomarker detection. Methods- Seventy human fresh frozen CRC specimen, with at least 10 patients per group, were digested with trypsin and analysed by fixed window DIA 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, exosome enrichment by centrifugation from serum samples of the same patients (10 per stage) and 10 serum samples from healthy individuals was performed and analysed by data dependent acquisition on a Fusion mass spectrometer. Results - Basic RP chromatography resulted in a spectral library including ~4600 proteins. Data extraction with Skyline revealed 1900 proteins quantifiable across all 70 specimen. Statistical analysis was performed and stage-specific marker proteins were observed. Especially for Stage I compared to the other stages, significant differences were observed. For later stages (III & IV) metastatic markers were significantly increased. For the exosome samples, label free MS1 quantification was performed and ~1400 proteins were quantified. Similar to the tissue specimen, stage-specific protein profiles were observed. With increasing stage, the number of detectable proteins associated to cancerous alteration increased which may be used for blood-based colon cancer diagnosis. Furthermore, exosomes from patients with cancer drastically varied from the healthy individuals. Conclusion - Mass spectrometric characterisation of fresh frozen CRC tissue specimen and serum exosomes identified stage-specific marker proteins with potential application in early cancer onset detection.

13

Soybean proteomics: Application to elucidation of flooding- tolerant mechanisms

Setsuko Komatsu[1]
1. Fukui University of Technology, Fukui, Japan

Food shortages are one of the most serious global problems in this century and it is important to increase the food production. Soybean is one of the important crop in the world, which is a source for protein, vegetable oil, and phytochemicals. On the other hand, global climate changes influence the magnitude and frequency of hydrological fluctuations and cause unfavorable environment for plant growth and development. Soybean is sensitive to flooding stress, which markedly reduces its growth. To identify the mechanism of flooding tolerance at initial stage in early-stage soybean, proteomic, transcriptomic, and metabolomic techniques were used. Flooding tolerant mutant line and abscisic acid-treated soybean, which exhibited flooding tolerant phenotype, were used as materials. Early-stage soybeans were treated during initial stage of flooding stress and roots were

collected for proteomic as well as metabolomic and transcriptomic analyses. Data were analyzed using functional categorization, cluster separation, and *in silico* protein-protein interaction. Furthermore, commonly changed metabolites, proteins, and genes between mutant and abscisic acid-treated soybeans were considered as flooding-tolerance related candidate factors. Finally, omics results were integrated to analyze the flooding tolerant mechanism in soybean and confirmed using biochemical and biological techniques. These results suggest that flooding tolerance at initial stage in early-stage soybean might be through protecting newly synthesized proteins and enhancing activities of antioxidative enzymes to remove reactive oxygen species. Furthermore, an integrated approach of proteomics and 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.

14

Proteomics for vaccine discovery: development of two potential whole-protein vaccines with efficacy against *Pseudomonas aeruginosa*

William Klare[1], Paula Niewold[2, 3], 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 defined 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 and subsequently screened for protection against pulmonary *P. aeruginosa* infection in mice, as well as defining immune correlates. We demonstrated high levels of protective efficacy for two out of three protein vaccines assessed. Presented is a workflow for identification and optimisation of vaccine candidates that is scalable, effective and widely applicable to other culturable bacterial species.

15

Identification and quantitation of wheat and barley α-amylase trypsin inhibitors, the triggers of non-coeliac gluten sensitivity

Utpal Bose[1], Keren Byrne[1], Malcolm Blundell[2], Crispin Howitt[2], Michelle Colgrave[1]

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

2. Agriculture and Food, CSIRO, Canberra, ACT, Australia

Some components, primarily proteins, from cereal grains can trigger a number of clinical symptoms which can affect health and quality of life. The α-amylase trypsin inhibitors (ATIs) present in the cereal grains are believed to be elicitors for non-coeliac gluten sensitivity (NCGS) in some individuals, but the composition and abundance of ATIs in major cereal cultivars such as wheat and barley are not available. Here, discovery proteomics data revealed the presence of 24 and 21 ATIs from wheat cultivars Chara and Magenta, respectively; whilst 12 and 14 ATIs, respectively detected from barley cultivars Hindmarsh and Commander. Next, a recently developed two-step ATI-specific extraction protocol followed by targeted LC-MRM-MS-based experiment was employed to monitor the ATI quantities across 15 commercial wheat cultivars, the 8 parental lines from a multi-parent advanced generation intercross (MAGIC) wheat population and 12 commercial barley cultivars. The results showed that the dimeric (0.19+0.53) and CM3 ATIs comprised ~45% of the total measured ATIs, whilst monomeric (0.28) ATIs comprised 14%, CM1 (7%), CM2 (14%), CM16 (14%) and CMx comprised (7%) within wheat and MAGIC cultivars. Notably, the measured ATI content was decreased (~50%) in the wheat cultivars Janz, Sunvale, Diamond Bird and Longreach Scout. Compared to the MAGIC parent lines the ATI content was decreased by ~40% in a reference wheat cultivar Chara. In barley, the ATI peptide level variation (biological) ranged between 5-35% across the 12 commercial lines. The barley cultivars Oxford and Yagan contain higher levels (~135% relative to the average barley ATI content), whereas cultivars Fleet and Bomi contained the lowest levels (~61% and 80% respectively). This approach employing detection and relative quantitation of ATIs would be applicable to identify or develop wheat/barley cultivars with a significantly lowered level of ATIs that might be suitable for people with NCGS.

16

SWATH-MS reveals functional differences between glycogenin 1 and 2 in yeast metabolism

Xinle Tan[1, 2], Ben Schulz[1]

1. school of chemistry and molecular bioscience, University of Queensland, Brisbane, QLD, Australia

2. Center for nutrition and food science, University of Queensland, bris

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.

17

RNA binding proteins mediated post transcriptional regulation in protist stage transitions

Balu Balan[1, 2], Samantha Emery-Corbin[1], Waruni Abeysekera[1], Gordon K Smyth[1], Andrew Webb[1], Staffan Svard[3], Aaron Jex[1, 2]

1. Walter and Eliza Hall Institute, Parkville, VICTORIA, Australia

2. Faculty of Veterinary and Agricultural Science, University of Melbourne , Melbourne, Victoria, Australia

3. Cell and Molecular Biology, University of Uppsala, Uppsala, Stockholm, Sweden

Protists are single-celled eukaryotes that reside in vastly different ecological niches, have complex life cycles and diverse phylogenetic origins. Many are parasitic and of significant global health relevance. The capacity of protists to rapidly adapt to major changes in their environment is critical to their survival. In most instances, they undertake these adaptations without cell division and must massively reconfigure their transcriptional and expressional behaviours. Among parasitic protists, these changes are essential for establishing infection and must occur within a few minutes of invasion, such that they appear pre-programmed. Pre-programming ('cell-fating') has been explored heavily in eukaryotic cells, including in developmental biology and stem cell differentiation. Post-transcriptional regulation (PTR), particularly mediated through translational repression, a phenomenon through which transcripts are held in stasis for later translation, is a key regulator of cell fate. PTR mechanisms, including in cell-fating, metabolism and other key functions, are heavily effected through a suite of RNA-binding proteins (RBPs). These have been shown, particularly in *Plasmodium ('malaria parasite')*, to have importance during it's initial infection. Noting this, RBPs are poorly characterized in most parasitic protists, beyond the behavior of a small number of RBPs in *Plasmodium* and *Toxoplasma*. We have bioinformatically mined and curated the 'RBPome' of 8 major parasitic protists – including three major diarrhoeal parasites, *Cryptosporidium parvum, Giardia duodenalis, Entamoeba histolytica*, the primary two human malaria parasites,*Plasmodium falciparum*and *Plasmodium vivax*, the causative agent of African sleeping sickness, *Trypanosoma brucei*, the sexually transmitted*Trichomonas vaginalis,* and*Toxoplasma gondii*. We have employed a multi-omics approach to explore behaviour of these RBPs over each parasite life-cycle and with respect to major ecological transitions during cyst formation and infection. This analysis found the global down-regulation of RBPs for most species at each major ecological change and specific up-regulation of a small suite of conserved RBPs. We explore the role of these up-regulated RBPs in translational repression and discuss their potential in setting the fate of infective cells as an essential and conserved mechanism under-pinning parasitic protozoan infection biology.

18

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 one of the most widespread post-translational modifications in the eukaryotic cell. It is known to regulate 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, by our lab and others, and in particular in yeast and human. We discovered two yeast enzymes, Efm3 (1) and Efm7 (2), as well as two human enzymes, eEF1A-KMT1 (2) and eEF1A-KMT3 (3), that each methylate specific residues in translation elongation factors. To assist the characterisation of these methyltransferases we developed Methyltransferase Motif Analysis by Mass Spectrometry (MT-MAMS) (4). 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. This has led to the realisation that eEF1A is targeted by more independent methyltransferases than any other protein in eukaryotes (5). Through SILAC-based proteomics, we found that eEF1A methylation is subtle in function, suggesting it may have the role of fine-tuning 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. Zhang, L., Hamey, J. J., Hart-Smith, G., Erce, M. A., & Wilkins, M. R. (2014). Elongation factor methyltransferase 3–A novel eukaryotic lysine methyltransferase. Biochemical and Biophysical Research Communications, 451(2), 229-234.
2. 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.

3. 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.
4. 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.
5. Hamey, J. J., & Wilkins, M. R. (2018). Methylation of Elongation Factor 1A: Where, Who, and Why? Trends in Biochemical Sciences, 43(3), 211-223

19

Interactions between small molecules recapitulate the exercise phosphoproteome and regulate protein secretion in vitro

Elise J Needham[1], Sean J Humphrey[1], Kristen C Cooke[1], Benjamin L Parker[1], David E James[1, 2]

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

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

Exercise promotes health through adaptive metabolic and mechanical remodelling, mediated by a network of kinases in response to homeostatic stress. Identifying the specific signalling changes that drive beneficial adaptations in exercise would provide novel drug targets for a range of diseases. Multiple stress- and metabolic-sensing pathways are acutely activated during exercise, however it is unclear whether they interact. We hypothesised that combining exercise-like treatments could recapitulate signalling interactions *in vitro*. Since interactions may contribute to beneficial effects of exercise, interrogating these effects rather than simple linear pathways is crucial to understanding exercise response. Comprehensive screening of exercise-like treatments in rat L6 myotubes ranked isoproterenol and thapsigargin as the most similar to the *in vivo* human skeletal muscle acute exercise phosphoproteome. Measuring global phosphoproteomes of L6 myotubes stimulated with thapsigargin, isoproterenol and their combination quantified >20,000 Class I phosphopeptides, of which 25% were regulated in at least one treatment. Strikingly, the combination of isoproterenol and thapsigargin uniquely regulated 962 phosphosites that were not regulated in either treatment alone. These unique sites were highly correlated to those regulated in exercised human biopsies suggesting that interactions recapitulate *in vivo* exercise signal transduction. Exercise regulates the secretion of factors to produce health benefits throughout the body. However, the underlying mechanisms regulating secretion are poorly understood. To investigate this, we measured the secretome of exercise-like treatments to delineate the pathways controlling protein secretion. Remarkably, the combined exercise-like treatments uniquely modulated the secretion of 80 proteins. Most notably, the secretion of fibrosis and thrombosis-stimulating proteins such as Serpine1 were downregulated by these treatments. Regulated phosphosites appear to drive this phenotype, including multiple phosphosites on the Serpine1 mRNA degrading protein Serbp1. This work provides a resource to link exercise-related signalling changes to protein secretion and other exercise-related phenotypes.

20

Distribution of isomerized and racemized amyloid β isoforms in the Human Brain using ion-mobility mass spectrometry

Soumya Mukherjee[1], Blaine Roberts[1]

1. The Florey Institute of Neuroscience and Mental Health, Carlton, ACT, Australia

Extracellular amyloid plaques and intracellular neurofibrillary tangles are the pathological hallmarks of Alzheimer's Disease (AD). It takes on average 19 years for amyloidβ (Aβ) peptides to deposit as insoluble plaques from onset to clinical dementia symptoms in AD. Such long-lived proteins and peptides without degradation and clearance can undergo further post-translational modifications (PTM). Several biochemical and analytical approaches have estimated very high degree of isomerization and racemization of Asp and Ser residues in Aβ purified from the insoluble plaques, along with sequential loss of the N-terminal amino acids. In this study we have characterized the most common isomerization and racemization of the Asp-1 and Asp-7 residues of the Aβ peptides present in AD brain based on both their chromatographic resolution as well as their collisional cross section (CCS) using high resolution ion mobility (IM) Q-TOF mass spectrometer (Agilent 6560). Using stable isotope labeled peptides we have also quantified the amount of these isomers/racemers in the different fractionated biochemical pools of the frontal cortex grey matter of human AD and control brains. Distribution of these isomerized and racemized peptides change from lower levels in the soluble/cytosol to higher levels in the insoluble/aggregated debris in AD brain, also indicating loss in the biochemical exchange of the pool of Aβ with the progression of the disease. These findings have implications in Aβ neurotoxicity, oligomerization, structures of amyloid fibrils present in the AD brain as well establishing CSF/blood-based biomarkers.

1. Masters, C. L.et.al. Proc. Natl. Acad. Sci. U. S. A. 1985, 82, 4245.
2. Bohme, L.et. al. Biol. Chem. 2008, 389 (8), 1055-66.
3. Lyons, B.et.al. Anal. Chem. 2016, 88 (5), 2675-84.
4. Roberts, B. R. et.al. Brain 2017, 140 (5), 1486-1498.

21

The reduced protein methylation network in the early-branching protozoan parasite, *Giardia duodenalis*

Samantha J Emery-Corbin[1], Swapnil Tichkule[1], Balu Balan[2, 1], Brendan RE Ansell[1], Louise Baker[1], Daniel Vuong[3], Ernest Lacey[4, 3], Staffan G Svärd[5], Aaron R Jex[2, 1]

1. Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

2. Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Melbourne, Victoria, Australia

3. Microbial Screening Technologies Pty. Ltd., Smithfield, NSW, Australia

4. Macquarie University, North Ryde, NSW, Australia

5. Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

Giardia duodenalis causes ~300 million cases of gastroenteritis (giardiasis) annually, with its unique cell biology shaped by early-branching origins in the eukaryotic phylogeny. In particular, protein methylation enzymes in *Giardia* are minimised relative to higher eukaryotes, with no annotated demethylase or protein arginine methyltransferase (PRMT) domain-containing proteins, and only six SET-domain-containing proteins in support of reduced lysine methylation (K-Me) machinery. However, despite a lack of methylation data, there is sufficient evidence for K-Me regulation in surface antigen switching, differentiation and drug resistance.

We have comprehensively demonstrated a conserved, functionally essential K-Me network in *Giardia*. Using *in silico* structural modelling, domain homology and inhibitor auto-docking, we have consolidated known Class V SET-domain methyltransferases, and annotated new Class I seven-beta-strand lysine methyltransferases. We have detected no PRMTs as per previous studies, which we further verified by no detectable arginine methylation via immunoblotting, and the lack of PRMT inhibitor activity in chemical screens. Immunoaffinity enrichment (IAP) of lysine methylation in the infective (trophozoite) and transmission (cyst) life-stages has identified 524 methylation sites on 322 proteins. Cytoskeletal proteins are significantly enriched, and we detected methylated RNA helicases and ribosomal/ribonucleoproteins in support of a role of K-Me in gene regulation, as well as on histone H2 and H3 variants. Additional mass spectrometry of histone-enriched fractions has allowed us to detect over 50 acetylation, methylation and phosphorylation sites on *Giardia* histone variants for the first time, including conserved, canonical H3 methyl marks. Indeed, only inhibitors of histone lysine methyltransferases (HKMTs) have detectable activity in *Giardia*, in particular inhibitors for H3 methyl marks. Indeed, inhibitor exposure assays during *in vitro* log-phase and encysting cultures, complimented by quantitative proteomics of over 2000 proteins, provides first links between gene regulation, phenotype and inhibitor pharmacology, and confirms K-Me plays a dynamic, essential role in *Giardia* biology.

22

Phosphoproteomic analysis incorporating titanium dioxide - and phosphotyrosine superbinder SH2 domain-affinity purification unveils unique signaling mechanism of the tumour suppressor protein kinase CHK in colorectal cancer cells

Heung-Chin Cheng[1], Andy J.-Y. Low[1], Ching-Seng Ang[1], Nicholas Williamson[1], Amardeep Dhillon[2], Jet Phey Lim[2], John Mariadason[3]

1. University of Melbourne, Parkville, VIC, Australia

2. Faculty of Medicine, Deakin University, Waun Pond, Vic, Australia

3. Oncogenic transcription laboratory, Olivia Newton-John Cancer Research Institute, Heidelberg, Vic, Australia

The Csk-homologous kinase CHK is a potential colorectal cancer (CRC) tumour suppressor of which the expression is down-regulated by epigenetic silencing. It is well known that CHK exerts its tumour suppressive action in part by phosphorylating and inhibiting the Src-family protein tyrosine kinases (SFKs), which are over-activated in CRC cells. Being a protein kinase, CHK can potentially exert its tumour suppressive action by phosphorylating other non-SFK protein substrates. In this study, the tumour suppressive mechanism of CHK is investigated by incorporating two phosphopeptide enrichment approaches to analyse the phosphoproteome of the transduced CRC DLD1 cells capable of expressing recombinant CHK upon induction by doxycycline. The first approach, capable of identifying and quantifying the abundance of mainly phospho-serine- and phosphothreonine-containing peptides, employs a label-free proteomic method coupled with enrichment of phosphopeptides by titanium dioxide-affinity purification. To determine the changes of phospho-tyrosine (pY)-containing proteome of the transduced DLD1 cells in response to induced expression of recombinant CHK, we adapted the use the mutant SH2 domains (referred to as SH2 pY-superbinders) with significantly enhanced affinities for pY-containing proteins and peptides to enrich pY-containing peptides for proteomic analysis. In the second approach, perturbation of the pY proteome was accessed by the isotopic dimethyl labelled phosphoproteomic method coupled with the use of SH2 pY-superbinders. The combined results from both approaches revealed for the first time activation of the mitogen-associated protein (MAP) kinase signaling pathway upon CHK induction in CRC cells. Furthermore, the second approach identified several cellular proteins such as the membrane-bound myelin protein zero-like protein 1 (PZR) as potential non-SFK substrates of CHK. Further bioinformatic analysis suggests that its enhanced phosphorylation mediates activation of the MAP kinase signaling pathway induced by expression of recombinant CHK. The analysis also suggests that the CHK/PZR/MAP kinase pathway mediates the effect of CHK to cause cell senescence of CRC cells.

1. Advani G, Lim YC, Catimel B, Lio DSS, Ng NLY, Chüeh AC, Tran M, Anasir MI, Verkade H, Zhu HJ, Turk BE, Smithgall TE, Ang CS, Griffin M, Cheng HC (2017) "Csk-homologous kinase (Chk) is an efficient inhibitor of Src-family kinases but a poor catalyst of phosphorylation of their C-terminal regulatory tyrosine." Cell Commun Signal. 15(1):29. doi: 10.1186/s12964-017-0186-x.

23

Spatial proteomics with translational imaging mass spectrometry

Ron M.A. Heeren[1]

1. The Maastricht MulitModal Molecular imaging insitute, Maastricht University, Maastricht, Netherlands

A comprehensive understanding of molecular patterns of health and disease is needed to pave the way for personalized medicine and tissue regeneration. One barrier to predictive, personalized medicine is the lack of a comprehensive molecular understanding at the tissue level. As we grasp the astonishing complexity of biological systems (whether single cells or whole organisms), it becomes more and more evident that within this complexity lies the information needed to provide insight in the origin, progression and treatment of various diseases. The best way to capture disease complexity is to chart and connect multilevel molecular information within a tissue using mass spectrometry and data algorithms. It is the realm of big molecular data for disease classification. Charting this territory through the generation of molecular maps from cells and tissue has become reality through the clinical implementation of imaging mass spectrometry complemented with high throughput peptidomics and proteomics approaches. We have demonstrated how new MS based chemical microscopes target biomedical tissue analysis in various diseases as well as other chemically complex surfaces. In concert they elucidate the way in which local environments can influence molecular signaling pathways on various scales. State-of-the-Art molecular imaging with mass spectrometry now enables high resolution tissue screening that provides direct insight into tissue metabolism. Applications have penetrated various research domains from drug metabolism to visualize molecular signaling pathways in cancer. This lecture will highlight innovations in translational imaging mass spectrometry and showcase how mass spectrometry based multimodal molecular imaging can be used to reveal the phenotypes in complex, heterogeneous biological systems from cells to organs.

24

Combining chemical biology tools with metabolomics to identify small molecule targets of peroxide antimalarials

Carlo Giannangelo[1], Dovile Anderson[1], Ghizal Siddiqui[1], Jonathan Vennerstrom[2], Susan Charman[1], Darren Creek[1]

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

2. University of Nebraska, Lincoln, Nebraska, USA

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 treatment currently relies on peroxide antimalarials. New synthetic peroxides (OZs) are now in clinical trials and early clinical usage, but their mechanism(s) of action remain poorly defined. It is proposed that iron-mediated peroxide cleavage generates free radicals that alkylate a range of targets within the parasite. The aims of this study were to use metabolomics to reveal the small molecule targets of peroxide antimalarials, and to develop a Click Chemistry-based pull-down approach to enhance the sensitivity of their detection.

The untargeted metabolomics analysis of OZ-treated parasites, based on methanol extraction and HILIC LC-MS, demonstrated depletion of a subset of small peptides that indicated disruption of haemoglobin metabolism, but did not reveal any alkylated metabolites. A second untargeted metabolomics analysis using acidic acetone extraction and reversed-phase LC-MS revealed several novel metabolite features in the treated parasites, which were identified as covalent adducts of OZ-derived free radicals with haem, in addition to several novel degradation products arising from this OZ-haem adduct.

A novel chemical biology method was then developed to improve the sensitivity of detection for alkylated metabolites. Azide-modified OZ analogues were incubated with parasites and enriched with photocleavable Click Chemistry beads, followed by metabolomics analysis with high resolution LC-MS. Our novel enrichment approach successfully pulled-down azide-OZ analogues and their metabolites from a parasite extract using copper-free Click Chemistry. Photocleavage with UV irradiation then allowed release of the 'clicked' compounds for detection by LC-MS and identification by the IDEOM software with minor modifications. In conclusion, we have developed a novel approach for the identification of drug-derived metabolites and demonstrated that haem is the major small molecule target of OZ-derived radicals in the malaria parasite.

25

The cannabis metabolome

Simone Rochfort[1], Aaron Elkins[1], Christian Krill[1], German Spangenberg[1]

1. Agriculture Victoria Research, Bundoora, VIC, Australia

Cannabis is an herbaceous flowering plant of the *Cannabis* genus (Rosale) that has been used for its fibre and medicinal properties for thousands of years. In recent decades medicinal cannabis has become legal in several jurisdictions and the possibility of legalisation is being explored in many more. This legalisation has opened the field of medicinal cannabis research. The biochemistry of cannabis is rich and varied including phytocannabinoids, terpenes and phenolics. Each of these metabolite classes contains individual compounds with biological activity. This chemical diversity and the interaction between molecules may underpin the 'entourage effect' that is believed to contribute to the medical efficacy of cannabis. In order to fully explore this biochemistry we have undertaken both targeted and untargeted metabolomic and volatolomic analysis of diverse cannabis strains. 70 diverse strains have been characterised using liquid chromatography mass spectrometry (LCMS), nuclear magnetic resonance (NMR) and gas chromatography mass spectrometry (GCMS). The chemotaxonomic relationship between the strains will be discussed. In combination with genomics these analyses offer the potential for 'designer strains' to be developed for particular medical applications.

26

Rapid Evaporative Ionisation Mass Spectrometry: a new tool for rapid screening of detailed composition data in food

Alastair B Ross[1], Sara Stead[2], Paul Middlewood[1], Stefan Clerens[1]

1. AgResearch, Lincoln, N/A, New Zealand

2. Waters Corporation, Wilmslow, N/A, United Kingdom

Rapid Evaporative Ionisation Mass Spectrometry (REIMS) is a relatively new sample introduction method for mass spectrometry, based on direct detection of metabolites in vaporised solid or liquid samples. Data acquisition typically takes a few seconds, meaning that potential sample throughput can be very high. We have assessed the capabilities of REIMS with a quadrupole-high resolution mass spectrometer for food analysis across a wide range of food matrices including fruit, dairy products and meat. Our work with REIMS found that we could differentiate between food products from different production systems and between similar foods from different geographic origins, based on the REIMS fingerprint. REIMS is a promising new approach for using mass spectrometry-based metabolomics to do deep compositional analysis or 'foodomics' on a wide range of food matrices.

27

Utilising metabolic profiling and inflammation markers as diagnostic tools

Robert Trengove[1], Ben Hunter[1], Monique Ryan[2], Joel Gummer[1], Hayley Abbiss[1], Naomi Trengove[2], Garth Maker[1]

1. Murdoch University, Perth, WA, Australia

2. Health Sciences, The University of Notre Dame Australia , Perth, WA, Australia

The knowledge of the peptide Hepcidin-25 has evolved cysteine rich antimicrobial peptide to master iron regulator. We have observed the move from measurement in urine to measurement in blood and translation from MALDI to ESI with high resolution MS systems. The additional Hepcidin isoforms (Hepcidin-20, Hepcidin-22 and Hepcidin-24), in conjunction with Hepcidin-25, have been extensively evaluated in correlation with a range of diseases. The utilisation of Hepcidin measurements in early childhood necessitates the measurements with minimal blood volumes and we aim for 50uL as the target serum volume for measurement. The correlation with metabolite profiles and Hepcidin isoform changes is required in order to utilise Hepcidin as part of the diagnostic panel for monitoring the onset of sepsis and development of rapid Hepcidin profiling methodologies.

28

Uncovering the role of brain-derived lipid exosomes in Alzheimer's disease

Gavin Reid[1, 2], Huaqi (Kate) Su[1, 3], Kevin J. Barnham[3, 4], Laura J. Vella2[1, 2]

1. Biochemistry and Molecular Biology, University of Melbourne, Melbourne, VIC, Australia

2. School of Chemistry, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Melbourne, VIC, Australia

3. Florey Institute of Neuroscience and Mental Health, Melbourne, VIC, Australia

4. Department of Pathology, University of Melbourne, Melbourne, VIC, Australia

Alzheimer's disease (AD) is the most common form of dementia. The brain is highly enriched in lipids, and the disruption of lipid homeostasis has been reported for decades to be associated with AD pathogenesis. Recent studies have suggested that small extracellular vesicles (EV's), namely exosomes, that are released from brain tissue into the periphery (e.g., CSF or blood), have biomarker potential for AD diagnosis and as targets for therapeutic treatment. Exosomes are comprised of constitutive molecules and cargo, including proteins, RNA and lipids, which act as key players in cell-to-cell communication and that can be characterised to provide a snapshot of parental cell homeostasis. Here, we have employed a quantitative liquid chromatography-tandem mass spectrometry (LC-MS/MS) based lipidome analysis workflow to characterize, for the first time, the lipidome compositions of exosomes released by post-mortem frontal cortex brain tissue from a series of AD patients versus healthy controls. The results from this study reveal selective enrichment and remodelling of multiple exosome lipid classes and subclasses including phosphatidylethanoamine, sphingomyelin, lactosylceramide and ganglioside lipid species, that are known to play key roles in the regulation of physiological processes relevant to AD pathogenesis. These results therefore 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.

29

Lipidr: targeted lipidomics analysis workflow in R

Ahmed Mohamed[1], Jeffrey Molendijk[2, 1], Thi-My-Tham Nguyen[2], Johanna Barclay[3], Federico Torta[4], Markus Wenk[4, 5], Mark Hodson[6], Michelle M Hill[2, 1]

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

2. The University of Queensland Diamantina Institute, Woolloongabba, QLD, Australia

3. Mater Research Institute, The University of Queensland, Translational Research Institute, Woolloongabba, QLD, Australia

4. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Medical Dr, Singapore

5. Department of Biological Sciences, National University of Singapore, Singapore

6. Victor Chang Research Institute, Sydney, NSW, Australia

Mass spectrometry (MS)-based lipidomics is an emerging field, which enables simultaneous measurement of numerous lipid classes. In particular, multiple reaction monitoring (MRM) MS assays measure targeted lipids with high quantitative precision and reproducibility. Several methods have been reported for lipidomics analysis, which target different lipid molecules depending on the biological applications. While spectrum inspection and peak integration are often performed in Skyline or other vendor software, strikingly, software tools are still lacking for downstream analysis of targeted lipidomics, hampering data interpretation. Moreover, although molecular information such as chain length and unsaturation is readily obtained in MS experiments, it is not utilized in current downstream analysis workflows. Here, we present *lipidr*, an easy-to-use R package implementing a complete workflow for downstream analysis of lipidomics data. *lipidr* parses results exported from Skyline directly into R, allowing integration into current analysis frameworks. *lipidr* allows data inspection, normalization, univariate and multivariate analysis, displaying informative visualizations. We also implemented a novel Lipid Set Enrichment Analysis (LSEA), harnessing molecular information such as lipid class, chain length and unsaturation. We demonstrate the use of *lipidr*, along with in-house developed MRM assays, to analyse serum lipids from mice fed an experimental high-fat diet. A companion step-by-step guide is provided with *lipidr* allowing users effectively utilize the package, even with limited programming experience.

30

The developmental lipidome and proteome of *Haemonchus contortus* reflect unique adaptations in the transition to parasitism

Tao Wang[1], Shuai Nie[2], Guangxu Ma[1], Ching-Seng Ang[2], Pasi K. Korhonen[1], Rong Xu[3], Anson V. Koehler[1], Richard J. Simpson[3], David W. Greening[3], Gavin E. Reid[4, 5], Nicholas A. Williamson[2], Robin B. Gasser[1]

1. Department of Veterinary Biosciences, Melbourne Veterinary School, The University of Melbourne, Parkville, VIC, Australia

2. Melbourne Mass Spectrometry and Proteomics Facility, Bio21 Institute, The University of Melbourne, Parkville, VIC, Australia

3. Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, VIC, Australia

4. School of Chemistry, The University of Melbourne, Parkville, Victoria, Australia

5. Department of Biochemistry and Molecular Biology, The University of Melbourne, Parkville, Victoria, Australia

Publish consent withheld

31

Hooked on fat: nutrient uptake in chronic lymphocytic leukaemia

Lauren A Thurgood[1], Lara Escane[1], Karen Lower[1], Bryone Kuss[1]

1. Flinders University, Bedford Park, SA, Australia

Dysregulation of cancer cell bioenergetics is one of the hallmarks of cancer. The Warburg effect is one such documented change. However, glucose metabolism is not universally increased in cancer cells. Uptake of ^{18}F-FDG in chronic lymphocytic leukemia (CLL), the most common incurable malignancy of B-cells, fails as a marker of proliferation and whilst the underlying reason is poorly understood it suggests that CLL cells utilize energy sources other than glucose to proliferate. Using a combination of genetic, proteomic and lipidomic analyses, complemented with microscopy and nutrient uptake assays the preferred metabolic pathways of CLL cells have been identified. In a variety of nutrient deprived cell lines we measured the uptake of fluorescently labelled short, medium and long-chain fatty acids and the glucose analog 2-NBDG by flow cytometry. The two CLL lines, MEC1 and MEC2 have a preference for long chain fatty acid uptake (LCFA), over short and medium chain and low uptake of 2-NBDG. We found an increase in the LCFA uptake receptor, CD36, in primary CLL samples and found that the expression of GLUT1 and GLUT3 were similar in CLL samples compared to healthy B-cells.

The data taken from our proteomic studies found an upregulation of proteins involved in lipogenesis in CLL peripheral blood lymphocytes, which are often quiescent and an increase in β-oxidation proteins in the proliferative compartment of the lymph node. These findings together with our morphological examination of CLL cells using transmission electron microscopy and confocal microscopy suggests that peripheral CLL cells scavenge lipids, which are stored in lipid droplets and are protected from degradation by a high expression of PLIN proteins. These cells circulate back to the lymph nodes to proliferate, the lipid droplets are degraded, likely by lipophagy and neutral lipolysis which frees fatty acids for β-oxidation. Our results begin to unravel CLL bioenergetics and dysregulation of cellular metabolism that occurs in this disease. We are now investigating whether the manipulation of these pathways, particularly lipophagy, may represent a novel therapeutic approach in CLL.

32

Surfaceome of exosomes secreted from the colorectal cancer cell line SW480

Rong Xu[1, 2], David Greening[1, 2], Maoshan Chen[1], Alin Rai[1], Hong Ji[1], Nobuhiro Takahashi[3], Richard Simpson[1, 2, 3]

1. La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC, Australia

2. Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Australia

3. Global Innovation Research Organization, Tokyo University of Agriculture and Technology, Tokyo, Japan

Publish consent withheld

33

Metabolic reprogramming in cancer

Kristin K Brown[1, 2]

1. Peter MacCallum Cancer Centre, Melbourne, VICTORIA, Australia

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

Metabolic reprogramming is a hallmark of cancer that contributes to malignant transformation and tumour progression. In recent years, there has been growing interest in developing strategies to exploit the metabolic vulnerabilities of cancer cells for therapeutic gain. However, our ability to do this is dependent on a thorough understanding of the ways in which cancer cell metabolism is influenced by cell-intrinsic and cell-extrinsic factors.

Signalling networks downstream of oncogenes regulate cancer cell metabolism. Our recent studies have focused on the oncogenic transcriptional co-activator YAP. Aberrant activation of YAP is widespread in human cancers yet, there is little knowledge regarding mechanisms by which YAP drives tumourigenesis. We find that YAP overexpression induces *de novo* lipogenesis *in vitro* and *in vivo* via transcriptional upregulation of a critical effector of the oncogenic phosphoinositide 3-kinase (PI3K) pathway. Importantly, inhibition of key enzymes in the *de novo* lipogenesis pathway blocks the uncontrolled proliferation associated with YAP-driven transformation. Our data reveal a mechanism of crosstalk between two important oncogenic signalling pathways and reveal a metabolic vulnerability that can be targeted to disrupt oncogenic YAP activity.

A variety of factors in the tumour microenvironment also have a major impact on cancer cell metabolism. Our studies have focused on characterising metabolic reprogramming events triggered upon chemotherapy exposure. Using *in vitro* and *in vivo* metabolomic profiling, we find that chemotherapy exposure induces an increase in the abundance of pyrimidine nucleotides as a result of increased flux through the *de novo* pyrimidine synthesis pathway. We find that pharmacological inhibition of *de novo* pyrimidine synthesis sensitizes cancer cells to genotoxic chemotherapy agents by exacerbating DNA damage. Our studies provide pre-clinical evidence to demonstrate that adaptive reprograming of *de novo* pyrimidine synthesis represents a metabolic vulnerability that can be exploited to improve the anti-cancer activity of genotoxic chemotherapy agents for the treatment of TNBC.

Together these studies highlight the importance of understanding how intrinsic and extrinsic factors are integrated to influence cancer cell metabolism.

34

Multi-omics analysis of esophageal adenocarcinoma reveals impaired lipid antioxidant mechanisms during disease progression

Jeffrey Molendijk[1, 2], **Thomas Stoll**[1, 2], **Federico Torta**[3], **Markus Wenk**[3, 4], **Thomas Hennessy** [1, 5], **Mark Hodson**[6], **Michelle M Hill**[1, 2]

1. The University of Queensland Diamantina Institute, The University of Queensland, Brisbane, QLD, Australia

2. Precision and Systems Biomedicine Group, QIMR Berghofer, Brisbane, QLD, Australia

3. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

4. Department of Biological Sciences, National University of Singapore, Singapore

5. Life Sciences & Applied Markets Group, Agilent Technologies, Australia

6. Victor Chang Cardiac Research Institute, Sydney, NSW, Australia

The incidence of esophageal adenocarcinoma (EAC) has increased in recent years, whilst the five-year survival rate remains low at ~15%. EAC is associated with altered lipid metabolism, with obesity as a major risk factor. Conversely, cholesterol-lowering statin drugs are protective and attenuate growth and malignant potential of EAC cells. MALDI and spectroscopy studies report changes in the lipid profiles of EAC tissue compared to healthy squamous epithelium however, the specific lipid metabolism pathways altered in EAC and its pre-cancerous condition called Barrett's esophagus (BE) remain unclear. To study these pathways, we conducted mass spectrometry (MS) -based proteomics and lipidomics experiments on EAC, BE and healthy esophageal biopsies. Lipid and proteins were extracted from cell pellets using a biphasic MTBE/methanol method. Proteomic profiling was performed on a QE+ MS (Thermo). Untargeted discovery lipidomics experiments were performed on a 1290 Infinity II/6550 Q-TOF LC/MS system (Agilent) whilst targeted MRM experiments were performed on a 6490 Triple Quadrupole MS system. Combined analyses of proteomic and lipidomic alterations were performed to identify lipid metabolic pathways differentially expressed in EAC progression. Our results show that changes in the metabolism of the antioxidant plasmenyl lipids are associated with advanced disease stages. This is in agreement with the known mechanism of EAC development, as a result of chronic insults to the esophagus by acids, digestive enzymes and other sources of oxidative stress. These findings provide new insights in the roles of lipid metabolism in the development and prevention of BE and EAC.

35

From phenotypic screens to mode of action: A metabolomics approach to guide the development of novel anti-trypanosomal Drugs

Anubhav Srivastava[1], **Darren Creek**[1]

1. Drug Delivery, Disposition and Dynamics, Monash Institute of Pharmaceutical Sciences, Monash University, Melbourne, VIC, Australia

Access to large phenotypic screens has enabled the discovery of novel anti-infective compounds. Translating these compounds into new drugs faces a number of challenges. Finding the mode-of-action (MoA) can help in focusing the efforts to develop the most promising leads. Using a medium-throughput method, we performed untargeted metabolomics analyses of *Trypanosoma*

brucei, a kinetoplastid parasite which causes a neglected tropical disease known as sleeping sickness. This allowed us to investigate the mode of action of over 20 potent trypanocidal compounds from the Pathogen Box (available from Medicines for Malaria Venture). Over 500 metabolite features were identified in this study and mapped to the predicted metabolic network. This approach identified the metabolic pathways targeted by the most potent compounds which spanned nucleotide metabolism, lipid metabolism, co-factor synthesis and redox metabolism. Lysophospholipid metabolism, glutamate metabolism, purine salvage and S-adenosyl-methionine and polyamine pathways were found to be specifically hit. Interestingly, multivariate analyses showed that compounds with similar chemical structures did not always have similar biochemical activities. This study showed that a simple metabolomics assay can rapidly reveal the MoA of newly discovered anti-infective compounds. This information can be used for prioritising compounds in the optimization pipeline and help in designing combination therapies that target discrete pathways to overcome emerging drug-resistance.

36

A meta-analysis of metabolites associated with BMI in > 30,000 individuals from the COMETs consortium

Jessica A Lasky-Su[1], Rachel S Kelly[1], Steve C Moore[2], Ewy Mathe[3], Temprosa Marinella[4], Zanetti Krista[5]

1. Brigham and Women's Hospital and Harvard Medical School, Newburyport, MA, United States

2. Division of Cancer Epidemiology & Genetics, National Cancer Institute, Behesda, MD, USA

3. Medical Informatics, Ohio State University, Columbus, OH, USA

4. Biostatistics Research Center, George Washington University, Rockville, MD, USA

5. Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

Obesity is a significant public health burden with nearly 40% of the world's population classified as overweight and 13% as obese. Body mass index(BMI) results from a complex interaction between lifestyle, environmental factors and underlying genetic susceptibility. The metabolome represents a dynamic functional readout of the state of a biological system; encompassing both genetic and environmental influences. Consequently, metabolomics is ideally suited to explore the drivers and manifestations of BMI on a mechanistic and metabolic level. We are conducting a meta-analysis within the context of the global COnsortium of METabolomics Studies (COMETS). Participating COMETS cohorts with plasma metabolomic profiling and concurrent measures of BMI. We utilized models that explore the correlation between BMI and metabolite levels, with adjustment for and stratification by potential effect modifiers. These models are run independently within all participating cohorts and then the results are meta-analyzed using random and fixed effects models. Bonferroni is used to adjust for multiple testing and the Cochran's Q test and the Wald test are used to explore between-study and between-strata heterogeneity respectively. A total of 43 cohorts encompassing >98,000 participants have enrolled in the study to date. Preliminary findings using 8 cohorts and 4450 subjects, there were 256 metabolites that could be harmonized across all cohorts. In a baseline model adjusting for gender, age, and race, 137(53.5%) metabolites were significantly associated with BMI after Bonferroni correction. The top hit was the amino acid Glutamtate (meta-analyzed spearman's r=0.33($1.1x10^{-47}$). A number of other metabolites demonstrated significant heterogeneity across the cohorts that was primarily driven by gender and fasting status. An updated analysis of ~30,000 participants retained these initial findings. These analyses confirm the feasibility of large-scale meta-analyses of metabolomics and suggest an important role for glutamate metabolism in BMI, which may be acting through hypothalamic regulation of appetite, insulin sensitivity or dyslipidemia. The results also suggest the metabolome of BMI may differ between genders and that BMI-plasma metabolite relationships are sensitive to fasting status at the time of blood draw.

37

Components of reproducible quantitative mass spectrometry-based research: a statistician's perspective

Olga Vitek[1]

1. Northeastern University, Brookline, MA, United States

This talk presents a statistical perspective on reproducible quantitative mass spectrometry-based proteomics. Statistical components of reproducibility include experimental design, from both biological perspective (which proteins and samples, and how many, do we need to quantify?) and technological perspective (are the assays appropriate for the task? Do the experimental steps run properly?). Statistical components of reproducibility also include data processing (which features should we use to quantify a protein?) and downstream statistical analysis (how to detect changes in protein abundance? Are our conclusions consistent with prior results?). Answer these questions requires the availability of statistical methods, and but also of publicly available data that help understand the advantages and the limitations of the methods. This talk will highlight the contributions of our lab to these components of reproducible research.

38

Toffee: A highly compressed file format for time of flight and orbitrap DIA-MS

David Clarke[1], Akila Seneviratne[1], Brett Tully[1]

1. Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

Data generated by Data Independent Acquisition mass spectrometry (DIA-MS) exists in a proprietary data format that is opaque to open-source software. Typically, a user will convert this data to the open mzML format before utilising a computational proteomics pipeline such as OpenSWATH [1]. However, mzML files tend to consume a large disk space and are organised around slices (or scans) in retention-time space. The former limitation significantly increases the computational hardware and funding required for high throughput proteomics, while the latter limits the manner in which software can efficiently access the data. Toffee is an HDF5 backed file format that is portable, open, and highly efficient with resulting file sizes that are 5% smaller than the original vendor file. It takes advantage of the inherent sparsity of DIA-MS raw data and the physics of the relevant mass analyser, to compress the data while preserving the information content. Toffee also enables fast access of subsections of data which enables trivial exploration and manipulation of the raw data. These benefits come at a cost of a mass accuracy loss of 5-10 ppm; however, initial testing shows that this mass accuracy error has negligible impact on OpenMSToffee (a wrapper around OpenSWATH that enables its use with toffee files) results.

1. Röst, H. L., Rosenberger, G., Navarro, P., Gillet, L., Miladinović, S. M., Schubert, O. T., … Aebersold, R. (2014). OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nature Biotechnology, 32(3), 219–223. https://doi.org/10.1038/nbt.2841

39

LFQ-Analyst: An interactive web-platform to analyse 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 application to comprehensively and globally quantify proteins across biological samples and conditions. Owing to the popularity of the technique, software suites like MaxQuant have been developed to extract, analyse and compare spectral features and report quantitative information of peptide, protein and even post-translationally modified (PTM) sites. There is still a lack of accessible tools for the interpretation and downstream statistical analysis of these complex datasets. Therefore, we developed LFQ-Analyst; a web application to streamline differential expression analysis and visualisation of label-free quantitative proteomics datasets. This tool provides a wealth of user-analytics features including differential expression, dimensionality reduction, clustering and different quality checks in tabular and graphical forms to facilitate exploratory and statistical analysis of quantitative datasets produced from proteomics investigations. LFQ-Analyst was designed to use sophisticated techniques for analysis, yet be accessible to researchers without extensive proteomics experience. Furthermore, users can automatically generate comprehensive analysis reports.

40

Signalling networks in the analysis of proteomic data

Hannah Huckstep[1], Jarrod Sandow[1], Andrew Webb[1], Liam Fearnley[1], Melissa Davis[1]

1. Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

Signalling networks have the potential to provide useful insight into mechanisms driving disease progression. 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 signalling. Nevertheless, the simple identification and quantification of proteins from different conditions is not sufficient to reconstruct the mechanisms underpinning the observed differences. Functional analysis methods have been developed to help with the interpretation of proteomic and phosphoproteomic data, however, these methods suffer from a range of limitations and fail to account for the complexity of cellular signalling networks. Thus, there is a need for tools, methods and frameworks that consider underlying network structures to aid accurate interpretation and reconstruction of the biological mechanisms at play. An important first step is the derivation of the network since most knowledgebases today deal in pathways, which do not properly represent the global flow of information across the entire signalling system. Here we have developed a set of algorithms to extract and interrogate a more-global signalling network from the knowledgebase determined to be the most complete for this purpose. We also demonstrate how phosphoproteomics measurements can be mapped to this network to interpret the functional consequences of the observed changes in protein phosphorylation. This approach will enable a more unbiased and complete analysis to be performed over networks encompassing specific proteins and phosphoproteins of interest.

41

An atlas of protein-protein interactions across mammalian tissues

Leonard Foster[1]

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

Cellular processes arise from the dynamic organization of proteins in networks of physical interactions. Mapping the complete network of biologically relevant protein-protein interactions, the interactome, has therefore been a central objective of high-throughput biology. Yet, because widely used methods for high-throughput interaction discovery rely on heterologous expression or genetically manipulated cell lines, the dynamics of protein interactions across physiological contexts are poorly understood. Here, we use a quantitative proteomic approach combining protein correlation profiling with stable isotope labelling of mammals (PCP-SILAM) to map the interactomes of seven mouse tissues. The resulting maps provide the first proteome-scale survey of interactome dynamics across mammalian tissues, revealing over 27,000 unique interactions with an accuracy comparable to the highest-quality human screens. We identify systematic suppression of cross-talk between the evolutionarily ancient housekeeping interactome and younger, tissue-specific modules. Rewiring of protein interactions across tissues is widespread, and is poorly predicted by gene expression or coexpression. Rewired proteins are tightly regulated by multiple cellular mechanisms and implicated in disease. Our study opens up new avenues to uncover regulatory mechanisms that shape *in vivo* interactome responses to physiological and pathophysiological stimuli in mammalian systems.

42

Pathogenic mutation in the ALS/FTD gene, *CCNF*, causes elevated Lys48-linked ubiquitylation and defective autophagy

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

1. Macquarie University, North Ryde, NSW, Australia

2. University of Wollongong, Wollongong

3. Sydney University, Sydney

We recently identified mutations in the *CCNF* gene as a novel cause of Amyotrophic lateral sclerosis (ALS) and Frontotemporal dementia (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 protein substrates 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 signature features of ALS pathogenesis in vitro and in vivo. We investigated the effect of the cyclin F^{S621G} mutation on Lys48-specific ubiquitylation of protein substrates, and how this mutation alters its E3 ligase activity and stability that contributes to the ubiquitylation of neuronal proteins and causes dysfunction to the proteostasis network. 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 defects in the autophagy pathway specifically resulting in impairment in autophagosomal-lysosome fusion. We also identified a potential mechanism by which cyclin F interacts with and hyperubiquitylates p62, the receptor responsible for transporting ubiquitylated 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 in the autophagic machinery, that are pathological events leading up and/or contributing to ALS and FTD progression.

43

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 underlying mechanisms linking the dietary perturbation and the beneficial phenotypes remain to be uncovered. Here, we have employed proteome analysis of mouse liver, a key fasting-responsive organ, to identify the key protein abundance changes occurring in response to the EODF intervention compared to *ad libitum* fed animals. From >6,000 proteins quantified, more than 250 proteins were significantly altered by the EODF intervention. Among the most up-regulated proteins after EODF was acyl-CoA thioesterase2 (ACOT2), which can accelerate liver fatty-acid oxidation and when up-regulated is known to have beneficial effects on 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 analysis showed that SERPINA1 was also down-regulated ~3-fold in plasma of EODF animals, in addition to >20 other significant protein changes. We subsequently showed that the SERPINA1 protein abundance change in liver is matched by a significant change in SERPINA1 liver mRNA abundance. HNF4A is a known regulator of the SERPINA1 promoter for induction of transcription. We have used CHIP-qPCR analysis to quantify the changes in association between HNF4A and the promoter of either SERPINA1, or a canonical HNF4A-target gene (ABCC6). This analysis showed that the association of HNF4A with both

promoters was reduced >5-fold by the EODF intervention. These data suggest that HNF4A is globally regulated during EODF, although the abundance of HNF4A was not significantly altered in our liver proteome data. Therefore, we hypothesise that HNF4A is regulated by either post-translational modification and/or changes in protein-protein interactions. Immunoprecipitations of HNF4A from the liver tissue are on-going to identify significant differences in HNF4A modifications and interactions.

44

Systems immunology reveals the importance of direct and cross-presentation in driving cytotoxic T lymphocyte responses to distinct epitopes

Ting Wu[1], Jing Guan[1], Andreas Handel[2], David C Tscharke[3], John Sidney[4], Alessandro Sette[4], Linda M Wakim[5], Xavier Y. X. Sng[1], Paul G Thomas[6], Anthony W Purcell[1], Nicole L La Gruta[1], Nathan P Croft[1]

1. Infection and Immunity Program & Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, VIC, Australia

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

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

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

The magnitude of CD8+ T lymphocyte (CTL) 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 antigen presentation. Measurement of these factors therefore allows an assessment of their contribution to – and ultimately a chance to model and predict – immunogenicity. Here, we present systems immunology dissection into influenza A virus and vaccinia virus immune responses in mice, notably including targeted mass spectrometry to quantify the impact between direct- and cross-presentation of peptides to drive CTL responses as well as correlates between cell line measurements and those taken *ex vivo* from infected mice. 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. This study highlights how high quality quantitative proteomics and peptidomics data are pivotal in unravelling 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.

45

Integrated time-series analysis of proteomic and transcriptomic profiles for activated human T-cell function

Harshi Weerakoon[1, 2], Ahmed Mohamed[1], Ailin Lepletier[1], Jeremy Potriquet[1, 3], Yide Wong[1, 3], Stephen Kazakoff[1], Pamela Mukhopadhyay[1], Nic Waddell[1], John J Miles[1, 3], Jason Mulvenna[1, 2], Michelle M Hill[1]

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

2. School of Biomedical Sciences, The University of Queensland, St. Lucia, QLD, Australia

3. Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, QLD, Australia

Human T-cells play a pivotal role in adaptive immunity. The two major T cell subsets, CD4+ and CD8+ T cells are defined by expression of distinct T cell receptor (TCR), and perform different functions. Activation of the TCRs triggers proliferation of T cell clones through a complex process of T-cell activation. Transcriptome analysis has been the method of choice to study T cell activation, however, several studies demonstrate poor correlations between the transcriptome and the corresponding proteome in a variety of cells. In this study, we applied a systems biology approach to assess the concordance of temporal profiles of in vitro activated human CD4+ and CD8+ T-cells proteomic and transcriptomic expressions at protein, pathway and network levels. Shotgun proteomics and RNA-sequencing were performed at five different time points following in vitro activation to characterize early (24 hours) and late (24 hours to 7 days) phases of T-cell activation. Strikingly, at 6 hours, only 10% of differentially expressed genes were concordant with differentially expressed proteins. However, 50% of these changes were observed at later time points, indicative of a time-delay for transcriptome changes to be reflected at proteome level. Pathway analysis for proteomics identified activation of cytokine production and protein synthesis pathways during early phase, while late phase pathways were related to glycolysis and mTOR, PI3K and apoptosis signalling. Interestingly, despite the distinct functions of CD4+ and CD8+ T cells, both proteomic and transcriptomic profiling identified similar expression patterns during activation, suggesting a conserved activation process. Nevertheless, specific protein signatures were identified for T cell subsets including specific cytotoxic proteins in CD8+ T cells. In summary, this integrative multi-omic evaluation of T cell activation provides insight into the temporal relationship between transcriptome and proteome. To our knowledge, this is the first study to compare transcriptome and proteome through a time course in T cells.

46

New insights into HLA-B27 peptide repertoire: Elucidating the connection between *Salmonella* infection and ankylosing spondylitis

Terry C.C Lim Kam Sian[1], Ralf B Schittenhelm[1], Pouya Faridi[1], Anthony W Purcell[1]

1. Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

The association of Human Leukocyte Antigen (HLA) B27 with ankylosing spondylitis (AS) is one of the strongest links between possession of a number of closely related HLA alleles and an autoimmune disease. A prominent hypothesis for this association is the arthritogenic peptide theory, which invokes an argument that molecular mimicry between a foreign- and a self-peptide leads to the breakdown of immune tolerance and ensuing autoimmunity. A potential source of arthritogenic peptides comes from gastrointestinal bacteria, such as during *Salmonella* infection, which commonly precedes development of disease in a subgroup of AS patients. Another potential source of arthritogenic peptides is the newly described class of peptides, termed spliced peptides. These peptides form by fusion of two non-contiguous regions of the same or different proteins. The aim of this study is to investigate how the linear and spliced peptide pools are modulated post infection with *Salmonella* across the 8 most common HLA-B27 allotypes (HLA-B*27:02 – HLA-B*27:09). *Methods:* High resolution mass spectrometry combined with *de novo* sequencing were used to identify linear and spliced peptides across the 8 most common HLA-B27 in mock and *Salmonella* infected antigen presenting cells. *Results:* Our results showed that very low number of linear *Salmonella*-derived peptides are presented by the HLA-B27 allotypes studied. Interestingly, the number of spliced peptides derived from *Salmonella* antigens were higher than linear peptides, with some of the peptides forming human-*Salmonella* hybrids. We also observed that the consensus-binding motifs of the spliced peptides were significantly different to linear peptides, in particular the strong preference for arginine at the P2 anchor position was drastically reduced. *Conclusion:* The increased number and non-canonical motifs of spliced peptides could be potential sources of arthritogenic peptides and therefore could play a role in disease pathogenesis. However, further analyses such peptide binding studies, immunological assays and structural investigations are required to confirm immunogenicity of these peptides.

47

Glycosylation in sparkling wines

Cassandra L Pegg[1], Suchada Niamsuphap[1], Benjamin L Schulz[1]

1. University of Queensland, St Lucia, QLD, Australia

Sparkling wine is a high value product with commensurately high consumer quality demands. The quality of sparkling wines depends on both consumer expectations and the product's actual molecular composition, which is a complex and dynamic mixture of ethanol, carbon dioxide, and organics derived from grapes and yeast. Proteins are a key component of wines, but the details of the wine proteome remain largely unexplored. Here, we used SWATH-MS to study the proteome of sparkling wines. As expected, we identified a moderately complex mixture of proteins from grapes and yeast. Surprisingly, we also identified highly abundant glycopeptide fragments from secreted yeast proteins. Although these are abundant, and likely critical determinants of sparkling wine quality, they have not been previously reported. We present optimised methods for sample preparation and relative quantitative proteomics and glycoproteomics that will be useful in process and product quality control of sparkling wines and related beverages.

48

Proteomics reveals multiple phenotypes associated with *N*-linked glycosylation in *Campylobacter jejuni*

Joel A Cain[2, 1], Ashleigh L Dale[2, 1], Paula Niewold[2, 3], William P Klare[2, 1], Lok Man[2, 1], Melanie Y White[2, 3], Nichollas E Scott[1, 4], Stuart J 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

Campylobacter jejuni is the leading cause of acute gastroenteritis, and is typically acquired *via* consumption of undercooked 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. 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), quantifying 1359 *C. jejuni* proteins. Data independent analysis (DIA-SWATH-MS) was also employed to validate changes to 1080 proteins. Deletion of the *pglB* oligosaccharyltransferase (*ΔpglB*) resulted in a significant change in abundance of 185 proteins, with 137 restored by the reintroduction of *pglB* (*ΔpglB*::*pglB*). Loss of *pglB* was associated with significantly reduced abundances of known *pgl* targets and increased levels of stress-related proteins, including ClpB, GroEL, GroES, GrpE and DnaK. *pglB* mutants demonstrated reduced survival following temperature and osmotic shock, and showed altered biofilm phenotypes compared to wild-type *C. jejuni*. Finally, targeted metabolomics established that loss of *N*-glycosylation resulted in significant alterations to the *C. jejuni* metabolome, including changes in amino acid and carbon source preferences and disruption of chemotactic responses to key growth substrates. These effects could be correlated to protein-level changes within *pgl* mutants towards a host of cellular transporters, transducer-like proteins required for substrate recognition, and respiration-associated proteins necessary for growth under low oxygen conditions. These data indicate a multi-factorial role for *N*-glycosylation in *C. jejuni* physiology.

49

Post-Column Make-up Flow (PCMF) of organic modifiers enhances the sensitivity of capillary-flow PGC-LC-MS/MS-based glycomics

Hannes Hinneburg[1], Falko Schirmeister[2], Terry Nguyen-Khuong[3], Nicolle Packer[1], Erdmann Rapp[5, 4], Morten Thaysen-Andersen[1]

1. Macquarie University, NSW, Australia

2. Max Planck Institute of Colloids and Interfaces, MPIKG, Postdam, Germany

*3. Bioprocessing Technology Institute (BTI), A*STAR, Singapore, Singapore*

4. glyXera GmbH, Magdeburg, Germany

5. Max Planck Institute for Dynamics of Complex Technical Systems, MPI Magdeburg, Magdeburg, Germany

Publish consent withheld

50

Dynamic cellular glycoproteome and proteome during influenza infection

Cassandra L Pegg[1], Kirsty R Short[1], Benjamin L Schulz[1, 2]

1. School of Chemistry and Molecular Biology, University of Queensland, St Lucia, QLD , Australia

2. Centre for Biopharmaceutical Innovation at the Australian Institute for Bioengineering and Nanotechnology, University of Queensland, St Lucia, QLD, 4072

Influenza A viruses are responsible for the annual epidemics that cause severe illness in millions of people worldwide. Seasonal vaccines are administered to prevent infection but the two major antigenic glycoproteins found on the viral surface, hemagglutinin and neuraminidase, for which there are numerous subtypes, are subject to continual antigenic change. Protective immunity is only conferred when there is antigenic similarity between the strains used for vaccine development and circulating influenza isolates. Insights into influenza biology and host responses to viral infection are therefore needed to guide new effective therapeutics and vaccines. Proteomic profiling has highlighted changes in host-cell responses to influenza infection that are strain-specific and related to pathogenicity, however the dynamics of the host and viral glycoproteomes during infection have not been taken into consideration. Glycosylation is an essential regulatory mechanism of protein function and can have a profound influence on both normal and irregular biological processes. The importance of glycosylation in host-pathogen interactions is well established, viral surface proteins have been implicated in protein biosynthesis, attachment and entry, induction of immune responses and evasion of host-immune defences. Influenza must subvert host glycosylation machinery to synthesise the biomolecules required for productive infection, and in doing so, the virus disrupts these host pathways. We studied the glycoproteome and proteome of adenocarcinomic human alveolar basal epithelial cells at time points during influenza infection to try and identify novel pathways that can be exploited as targets for therapeutics. For the glycoproteomic analyses we enriched glycopeptides by HILIC from subcellular fractions and investigated site-specific occupancy and the monosaccharide composition of the attached glycans. For the proteomic analyses, we implemented a SWATH approach to quantify changes in proteins from subcellular fractions. The work presented herein highlights the these mass spectral approaches to monitor the dynamic glycoproteome and proteome during influenza infection.

51

Application of ion mobility and electron capture dissociation (ECD) mass spectrometry to top-down and bottom-up proteoform characterization

Blaine R Roberts[1], Eugene Kapp[1], Soumya Mukherjee[1], Yury V Vasil'ev[2], Valery G Voinov[2], Joesph 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. We have developed 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 with ion mobility separations. Nearly complete sequence coverage is obtained with “native”-folded proteins such as the $5^+$and 6^+charge states of ubiquitin. 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); half of the human proteome is smaller than 30kDa. 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. Labile post-translational modifications were also retained. The copper and zinc cofactors in superoxide dismutase (17 kDa) remained bound to their respective binding sites in ECD fragments. 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.

52

Laser ablation for tissue sampling for proteomics and metabolomics

Hartmut Schlüter[1], Pascal Steffen[2], Nils-Owe Hansen[3], R.J.Dwayne Miller[3], Mark Molloy[2]

1. Mass Spectrometric Proteomics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

2. Bowel Cancer & Biomarker Lab, Kolling Institute, The University of Sydney, St.Leonards, Sydney , New South Wales, Australia

3. Atomically Resolved Dynamics Department, Max-Planck Institute for the Structure and Dynamics of Matter (MPSD), Hamburg, Germany

For every omics approach sample preparation is one of the most critical steps. Especially tissue sampling is very challenging because during the process of homogenization biomolecules are known to be converted by chemical reactions like oxidation as well as enzymatic reactions. The latter are caused by the release of enzymes from their compartments during homogenization. We have recently shown that sampling of tissues for proteomics with a picosecond infrared laser (PIRL) is very soft. Labile molecules like glycoproteins stay intact during laser ablation. Even enzymatic activities of proteins are maintained (1). Furthermore, the total yield of proteins with respect to their total number and total amount is significantly better compared to conventional homogenization. Proteoforms during tissue sampling with PIRL also remain intact in contrast to conventional homogenization (2). Thus, especially for the future of top-down proteomics PIRL is very promising (3). In addition to the application of PIRL for tissue sampling for proteomics, we performed a series of experiments applying infrared laser systems for tissue sampling for lipidomics and metabolomics. We sampled different pig brain areas with PIRL and investigated the condensates of the tissue aerosols with shot-gun lipidomics. By principle component analysis we were able to differentiate the different brain tissues. We compared the yields of metabolites obtained by conventional tissue sampling and by PIRL. The metabolites of both homogenization procedures were analyzed with p180-kit of Biocrates, by which approximately 180 metabolites are quantified via selected reaction monitoring by LC-MS and flow-injection analysis (FIA) by a triple quadrupole mass spectrometer. The results clearly showed that PIRL is more efficient in tissue sampling for proteomics. Furthermore, PIRL is giving access for volatile labile molecules by irradiation of tissues. In summary tissue sampling with PIRL for lipidomics, metabolomics and proteomics is advantageous compared to classical sampling and homogenization methods, giving higher yields of intact biomolecules and thereby allowing a closer and more realistic view to the original composition of biomolecules in intact tissues.

1. Ultrafast extraction of proteins from tissues using desorption by impulsive vibrational excitation. Kwiatkowski M, et al., Miller RJ, Schlüter H. Angew Chem Int Ed Engl. 2015 Jan 2;54(1):285-8.
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., Steffen P, et al., Miller RJD, Schlüter H. J Proteomics. 2016 Feb 16;134:193-202.
3. How many human proteoforms are there? Aebersold R, et al. Kelleher NL, et al., Schlüter H, et al., Zhang B. Nature Chem Biol. 2018 Feb 14;14(3):206-214.

53

Evaluation of a novel LC system that embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics

Nicolai Bache[1], Philipp Geyer[2], Dorte Bekker-Jensen[3], Ole Hoerning[1], Lasse Falkenby[1], Peter Treit[2], Sophia Doll[2], Igor Paron[2], Florian Meier[2], Jesper Olsen[3], Ole Vorm[1], Matthias Mann[2]

1. Evosep Biosystems, Odense, FYN, Denmark

2. Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Martinsried, Bavaria, Germany

3. Novo Nordisk Foundation Center for Protein Research, Proteomics Program, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Seeland, Denmark

Background **-** Mass spectrometry-based proteomics and metabolomics are fast growing and powerful technologies, with the potential to revolutionise health care and precision medicine. However, available separation systems have so far limited throughput and robustness and thereby prevented omic technologies from being fully integrated and routinely applied in clinical settings. Here, we evaluate a conceptually novel liquid chromatography (LC) system that significantly increases robustness and sample throughput while maintaining the sensitivity of current nano-flow LC. Methods **-** The new system, called Evosep One, uses four low-pressure pumps in parallel to elute samples from a disposable and single use trap column while also forming a chromatographic gradient. The sample and gradient are moved into a storage loop that subsequently is switched in-line with a single high-pressure pump and an analytical separation column for separation. Results **-** We have characterised the performance of the new system regarding cross contaminations (<0.07%), retention time shifts and peak width (<3 sec) in over 1500 HeLa runs. The short overhead time of approximately 2 min allows us to efficiently measure 300, 200, 100, 60 or 30 samples per day with corresponding gradient lengths of 3.2, 5.6, 11.5, 21 and 44 minutes, respectively. The performance and applicability in various proteomics LC-MS strategies was evaluated for simple, medium and complex sample types. From fractionated HeLa cell lysates, deep proteomes covering more than 130,000 sequence unique peptides and around 10,000 proteins were rapidly acquired (18 h total instrument time). Using this data as a library for data independent acquisition, we demonstrate the quantitation of 5200 proteins in only 21 min. Conclusion **-** We evaluated and benchmarked how to use the Evosep One in cutting edge LC-MS strategies to significantly increase overall performance and throughput. We also demonstrate how this can be applied to clinical research workflows that require uninterrupted analysis of thousands of crude biological samples as well as routine applications such as doping control and drug screening.

54

Reversed phase x reversed phase on-line comprehensive two-dimensional liquid chromatography of intact proteins with triple quadrupole mass spectrometric detection

Kevin Schug[1], Yehia Z Baghdady[1], Evelyn H Wang[1], Durga D Khanal[1]

1. The University of Texas Arlington, Arlington, TEXAS, United States

The interest in analysis of intact proteins has increased significantly over the past several years, especially as it relates to biomarker discovery and biopharmaceutical development. On-line comprehensive two-dimensional liquid chromatography (LCxLC) has only been applied in a very limited fashion for analysis of intact proteins. We set out to create the first LCxLC method based on pH switching between two successive dimensions of reversed phase chromatography. Reversed phase x reversed phase LCxLC offers significant advantages in terms of availability of column chemistries, good compatibility of solvents between the two dimensions, and good compatibility of eluent mobile phases for mass spectrometric detection. Significant efforts succeeded to develop high pH separation conditions on a Water Acquity BEH wide pore C4 column using a triethylammonium bicarbonate buffer (pH = 10). Selectivity for a series of model proteins was demonstrated to be different and complementary than for a standard low pH reversed phase separation on an AMT Halo protein C4 column using a combined formic acid/trifluoroacetic acid-based mobile phase. LCxLC was accomplished by combining the high pH separation in the first dimension with the low pH separation in the second dimension on a Shimadzu Nexera-e instrument,t interaced with a LCMS-8050 triple quadrupole mass spectrometer. Several proteins in the model mixture, unresolved in the first dimension were successfully resolved in the 1 minute separation in the second dimension. A complex *E. coli* proteome sample was used to demonstrate a high degree of orthogonality and good coverage of the two-dimensional separation space. The effective peak capacity of the LCxLC-MS analysis was 2268, with a peak production rate of 37.8 peaks per minute over the entire 60-minute analytical run.

55

Mass spectrometry based, proteome-wide strategy for discovery and characterization of cellular nucleotide-protein interactions

Yan Ting Lim[1], Nayana Prabhu[2], Lingyun Dai[2], Ka Diam Go[2], Dan Chen[2], Lekshmy Sreekumar[2], Liyan Chen[1], Chris Soon Heng Tan[1], Andreas Larsson[2], Radoslaw Sobota[1], Par Nordund[2, 3, 1]

1. Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A-STAR), Singapore

2. School of Biological Sciences, Nanyang Technological University, Singapore

3. Department of Oncology-Pathology , Karolinska Institutet, Stockholm, Sweden

Metabolite-protein interactions define the output of metabolic pathways and regulate many cellular processes. Although diseases are often characterized by distortions in metabolic processes, efficient means to discover and study such interactions directly in cells have been lacking. A stringent implementation of proteome-wide cellular thermal shift assay (CETSA) to study interactions of human proteins with key nucleotide-based metabolites was developed and applied to 11 key cellular nucleotides, where previously experimentally confirmed protein-nucleotide interactions were well recaptured. Many predicted, but never experimentally confirmed, as well as novel protein-nucleotide interactions were discovered. Interactions included a range of different protein families where nucleotides serve as substrates, products, co-factors or regulators. In cells exposed to thymidine, a limiting precursor for DNA synthesis, both dose- and time-dependence of the intracellular binding events for sequentially generated thymidine metabolites were revealed. Interactions included known cancer targets in deoxyribonucleotide metabolism as well as novel interacting proteins. This stringent CETSA based strategy will be applicable for a wide range of metabolites and will therefore greatly facilitate the discovery and studies of interactions and specificities of the many thousand metabolites in human cells that remain uncharacterized.

56

High-throughput proteomic analysis of FFPE tissue using PCT-SWATH MS

Steven G Williams[1], Natasha Lucas[1], Andrew B Robinson[1], Belinda Serafin[1], Peter G Hains[1], Phil J Robinson[1]

1. The Childrens Medical Research Institute, Westmead, NSW, Australia

Formalin-fixed paraffin-embedded (FFPE) tissue archives present as a potential rich source for proteome data, spanning a vast range of diseases and time. While many methods for handling FFPE samples are known, they are limited by the large number of washing steps and tissue disintegration during preparation, and require up to 12 hours to achieve. These issues limit use of these samples to small cohorts. Currently, there is no published fast and reproducible method suitable for both high-throughput and small-scale sample analysis. Developing an effective method to analyse and collect proteomes from FFPE tissue would allow true high-throughput sample preparation that is robotics-capable to capitalise on the available samples. This aim aligns with ProCan's goal to process 70,000 tumour samples over the next 7 years. Here we describe an effective high-throughput protocol for the rapid processing of FFPE samples, including wax removal and cross-link reversal. The method is based on pressure cycling technology (PCT) in a barocycler instrument. The washing steps are reduced to 3 minutes and changes to the barocycler processing allows compatibility with standardised workflows used for fresh frozen samples. With this method, 96 samples can be taken from tissue sample to MS-ready in 6 hours. We outline the method optimisation and show that PCT is a robust and reproducible method with significant advances over traditional methods used to process archival tissue. By comparing the proteome by SWATH-mass spectrometry of several rat tissue FFPE samples to their equivalent fresh frozen samples, we show the variation between sample types was no different to that between samples of the same type. The results allow authentic high throughput sample preparation of FFPE tissues for the first time, suitable for robotics and industrial-scale proteomics analysis.

57

Universal Solid-Phase Protein Preparation (USP^3) for high-throughput clinical proteomics

Laura F Dagley[2, 1], Giuseppe Infusini[2, 1], Andrew I Webb[2, 1]

1. The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

2. Department of Medical Biology, University of Melbourne, Parkville, VIC, Australia

Selecting a sample preparation strategy for mass spectrometry-based proteomics is critical to the success of quantitative workflows, especially in the context of large clinical cohorts of patient samples where accuracy and reproducibility are of paramount importance. The factors that dictate the overall utility of sample preparation methodologies are balanced between efficiency, sensitivity and robustness versus speed, scalability and flexibility. Here we present a universal, solid-phase protein preparation (USP^3) method which is rapid, robust and scalable, facilitating high-throughput protein sample preparation for bottom-up mass spectrometry (MS) analysis.

We have demonstrated the utility of the USP^3 workflow in a patient cohort of children diagnosed with acute rheumatic fever and rheumatic heart disease, conditions which are highly prevalent in indigenous populations of Oceanian countries, including Australia. We obtained neat plasma and serum samples from children diagnosed with these conditions from ethnically-diverse populations including plasma samples from the Northern Territory of Australia (n=48; 26 ARF, 3 RHD, 7 healthy controls, 12 alternative diagnoses), serum samples from Fiji (n= 38; 9 ARF, 12 RHD, 17 healthy controls) and serum samples from New Caledonia (n=38; 28 ARF, 10 healthy controls). Neat plasma/serum samples were subjected to on-bead digestion (USP^3 method) followed by high-resolution MS analysis on a Bruker Impact II UHR-QTOF instrument. An in-house developed XIC-based feature detection method was then used for MS1-based peptide quantification. To determine the most relevant peptides that can be developed into liquid-biopsy signatures, we employed a Random-forest based machine-learning algorithm. Utilising this machine learning approach, we identified a set of statistically-relevant peptide signatures that distinguish between patients with ARF from all other patient groups. A number of these peptides are unique to the Australian cohort whilst there are others that consistently stratify the patient groups across the diverse genetic ancestries. Efforts are now underway to translate these findings into a rapid clinical diagnostic test using the Noviplex™ Plasma Prep cards, with initial feasibility studies resulting in the identification of ~360 unique proteins from single-shot MS analyses.

58

Developments in the sample preparation of formalin-fixed paraffin-embedded tissues for MALDI imaging

Alice Ly[1], Rémi Longuespée[2], Rita Casadonte[3], Petra Wandernoth[3], Kristina Schwamborn[4], Christine Bollwein[4], Christian Marsching[5], Katharina Kriegsmann[6], Carsten Hopf[5], Wilko Weichert[4], Jörg Kriegsmann[3], Peter Schirmacher[2], Mark Kriegsmann[2], Sören-Oliver Deininger[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany

3. Proteopath GmbH, Trier, Germany

4. Institute of Pathology, Technical University of Munich, Munich, Germany

5. Center for Biomedical Mass Spectrometry and Optical Spectroscopy (CeMOS), Mannheim University of Applied Sciences, Mannheim, Germany

6. Department of Hematology, Oncology and Rheumatology, University Hospital Heidelberg, Heidelberg, Germany

Background & Objective: MALDI Mass Spectrometry Imaging (MALDI-MSI) has been used to address clinically relevant questions, but must be performed on Formalin-Fixed Paraffin-Embedded (FFPE) tissues for widespread clinical use. Due to instrument and sample preparation limitations, the spatial resolution falls short of light microscopy. Additionally, the complex sample preparation has raised questions of reproducibility. We present a sample preparation protocol for MALDI-MSI that can distinguish fine structures in FFPE samples, and is reproducible when carried out at different sites. Methods: All sites used a standard operating procedure (SOP) which also covered instruments. FFPE tissues (3-5 µm; mouse intestine, human ovarian teratoma, tissue microarray (TMA) of tumor entities sampled from three different sites) were prepared for MALDI-MSI. Samples were coated with trypsin using an automated sprayer then incubated in a humid environment. After digestion, alpha-cyano-4-hydroxycinnamic acid was deposited using the same sprayer and the section analyzed with a rapifleX MALDI Tissuetyper. After acquisition, statistical analysis and segmentation feature extraction were conducted. Results: Different anatomical regions of the teratoma could be differentiated based on their mass spectrometric profiles and at high spatial resolution, as these corresponded to anatomical structures. Mouse intestine was used to assess whether operators conducting experiments at different sites can obtain similar results. Twenty measurements from two sites exhibited similar peak statistics; ten measurements conducted over five sites and three time points showed reproducible delineation of the villi from underlying muscle. Measurements of a TMA consisting of different tumor samples were used to investigate if different tissue sampling introduced variation. Statistical analysis of the TMA measurements indicated that tissue biology has a greater influence on the spectra than sample origin. Conclusion: These findings indicate that strict adherence to a SOP produces reproducible MALDI-MSI data from FFPE samples and that sampling site is not a major source of variation.

59

A new procedure of deyolking improves the proteome analysis in zebrafish early embryos

Purushothaman Kathiresan[1], Christopher Presslauer[1], Lim Teck Kwang[2], Steinar D Johansen[1], Qingsong Lin[2], Igor Babiak[1]

1. Faculty of Biosciences and Aquaculture, Nord University, Bodo, Nordland, Norway

2. Department of Biological Sciences, 14 Science Drive 4, National University of Singapore, Singapore, 117543- Singapore

Zebrafish is an excellent and well-recognised organism for investigating vertebrate development. Its proteome is apparently highly similar to that of humans, which makes the zebrafish an attractive model of human diseases. Although genomic and transcriptomic data on zebrafish embryonic development are well established, data on the proteome are scarce, particularly in the cleavage stages. This is mostly due to the high abundance of embryonic yolk proteins, which tend to mask the less abundant cell proteins. In this study, we developed an efficient protocol to reduce the amount of yolk in zebrafish early embryos to improve the shotgun approach-based LC-MS analysis. The analysis of the digested protein samples using the 1D MS/MS shotgun proteomics approach consistently demonstrated that the deyolking procedure resulted in a greater number of total and unique proteins being identified. This protocol resulted in approximately two-fold increase in the number of proteins identified in deyolked samples at cleavage stages. After the maternal-to-zygotic transition, the number of identified proteins increased gradually by 3-4 times compared to non-deyolked samples in both oblong and bud stages. Gene Ontology analysis revealed a high number of functional proteins in the deyolked samples, such as mitochondrial proteins, while the loss of ribosomal proteins was minimal. This deyolking protocol will improve both qualitative and quantitative proteome analyses, and provides an innovative tool in molecular embryogenesis of yolked animals, such as fish, amphibians, reptiles, and birds. The study was financed by the Research Council of Norway (InnControl project #275786).

60

Interactome dynamics; our initial observations with *in vivo* cross-linking

James E Bruce[1], Juan D Chavez[1], Andrew Keller[1], Martin Mathay[1], Jared P. Mohr[1], Xuefei Zhong[1]

1. Department of Genome Sciences, University of Washington, Seattle, WA, US

In vivo chemical cross-linking has evolved to provide new insight on proteins, complexes and supercomplex assemblies that exist in organelle, cells and tissues. Our recent efforts have focused on extending chemical cross-linking capabilities to allow quantitative exploration of the interactome. Dynamic cellular interactome measurements made during changing growth conditions, drug treatment, applied stresses, or other perturbations can increase understanding of cellular response and functional regulation. This presentation will discuss advancements in technologies and informatics capabilities we have pursued to help visualize interactome dynamics. These tools are being applied to the study of interactome dynamics involved in mitochondrial dysfunction in heart failure, chemoresistant cancer cells and pharmacological effects on the interactome, bacterial species/strain interactome comparisons and other areas. Efforts thus far have revealed intra- and inter-molecular changes that are chemoresistant phenotype-specific, drug concentration-dependent and mechanism of action-specific, as well as shared or species-specific interactome features among related bacterial strains. For example, molecular interaction changes among mitochondrial permeability transition pore (mPTP) regulators Cyclophilin D (CypD) and oligomycin sensitivity conferring protein (OSCP) in mitochondria from murine heart failure models or from mitochondria subjected to increased Ca^{2+} stress to cause mPTP activation have been visualized through quantitative mitochondrial cross-linking. Although many challenges yet remain at all levels of interactome dynamic studies, quantitative *in vivo* cross-linking offers unique opportunities for exploring this exciting frontier.

61

Large-scale detection of protein-protein interactions via size-exclusion chromatography, protein correlation profiling and crosslinking mass spectrometry

Daniela-Lee Smith[1], Gene Hart-Smith[1], Marc Wilkins[1]

1. UNSW, Sydney, NSW, Australia

Crosslinking mass spectrometry has the potential to detect thousands of protein-protein interactions in a cell or tissue, thus building interaction networks in a single experiment. However, many technical challenges remain before this can be achieved. One major challenge is how to decomplexify whole cell lysates to maximise crosslink identification, but to do this in a manner which preserves biological relevance. Here we aim to address these considerations through combining size-exclusion chromatography of native protein complexes, protein correlation profiling and crosslinking-mass spectrometry. Wild-type yeast lysate was subjected to offline size-exclusion high-performance liquid chromatography, followed by crosslinking with the mass spectrometry-cleavable DSSO crosslinker. Seventy fractions were then analysed by MS/MS/MS for 180 min per fraction, using LC/MS/MS CID+ETD/MS CID on an Orbitrap Fusion Lumos Tribrid. Crosslinked peptides were identified with XlinkX, using Proteome Discoverer 2.2. Across 70 fractions we detected 2217 crosslinks, representing 1944 protein-protein interactions (FDR 5%). The yeast nucleosome, RNA polymerase, the yeast ribosome, the yeast oligosaccharyltransferase complex and vacuolar ATPase are examples of complexes that were found. The scale of this dataset allow for in-depth benchmarking analyses. Therefore, aspects of crosslink discovery, including score, amino acids involved in links, peptide length and type and fragmentation doublets will be discussed. Overall this experiment represents a significant advance in the detection of crosslinks at a large scale and has strong potential to discover novel protein-protein interactions.

62

A comprehensive protein interactome of the type IX secretion system components in *Pophyromonas gingivalis* revealed using complementary proteomics approach

Dhana Gorasia[1], Paul Veith[1], Dina Chen[1], Jacqueline Heath[1], Christine Seers[1], Michelle Glew[1], Eric Reynolds[1]

1. Melbourne Dental School, The University of Melbourne, Melbourne, Victoria, Australia

The type IX secretion system (T9SS) has been recently discovered and is specific to *Bacteroidetes* species. *Porphyromonas gingivalis,* a keystone pathogen for periodontal disease, utilizes the T9SS to transport many proteins including the gingipain virulence factors across the outer membrane and attach them to the cell surface. The proteins transported by the T9SS have a conserved C-terminal domain (CTD), which is the signal for translocation across the outer membrane via the T9SS. At least 14 proteins, namely; Sov, PorK, PorL, PorM, PorN, PorP, PorQ, PorT, PorU, PorV, PorW, PorZ, PG1058 and PG0534, have been identified as components of the T9SS. However, the overall organisation and the dynamic interactions between these proteins have not been elucidated fully. In this study, we analysed *P. gingivalis* wild-type and *P. gingivalis* T9SS mutants using 1-D Blue Native Page (BN-PAGE) proteomics to identify a comprehensive interactome of the T9SS. The recent advances in mass spectrometry has enabled us to identify needles in a haystack. We show for the first time that the T9SS is organised as four sub-complexes and identify the components in each sub-complex. Sub-complex (i) PorK, PorL, PorM, PorN and PorT, (ii) Sov and PorW, (iii) PorP and PG1058 and (iv) PorU, PorV, PorQ and PorZ. We have validated these interactions using co-immunoprecipitation and chemical cross-linking. Further analysis of the *P. gingivalis* T9SS mutants have shed light on the potential role of the sub-complex PorP-PG1058 as the one that recruits T9SS substrates from the translocation channel and passes them on to PorV. PorV has been identified as a shuttle protein that transports the T9SS substrates to sub-complex (iv) which we have named the attachment complex as it is involved in the attachment of the translocated proteins to the bacterial cell surface. In summary, this is the first report showing a comprehensive interactome of the T9SS.

63

Protein stability in drug discovery and molecular biology

Mikhail Savitski[1]

1. EMBL Heidelberg, Heidelberg, Germany

Determination of proteome-wide thermal stability by the recently developed thermal proteome profiling technology (TPP) can be used to identify drug targets in living cells. This is achieved by combining the cellular thermal shift assay with multiplexed quantitative mass spectrometry. We have recently further developed this technology and substantially increased the sensitivity of detection of thermal stability changes and applied it to study the eukaryotic cell cycle. We observed pervasive variation of protein thermal stability and also of solubility with most changes occurring in mitosis and G1. A number of cellular pathways and components varied in thermal stability but not in abundance, such as cell-cycle factors, polymerases, and chromatin remodelers. We could show that changes in thermal stability reflect enzyme activity, DNA binding, and complex formation in situ. A large cohort of intrinsically disordered and mitotically phosphorylated proteins was stabilized and solubilized in mitosis, suggesting a fundamental remodeling of the biophysical environment of the mitotic cell. Thus protein thermal stability provides complementary information on the state of biological systems. In our latest work we have studied the effects of adenosine triphosphate, ATP, and have revealed novel effects of this metabolite on protein thermal stability and solubility.

64

A search for predictive biomarkers of ovine pre-partum vaginal prolapse

Stuart Brown[1], Gillian Norris[1]

1. Massey University, Palmerston North, NZ, New Zealand

Ovine pre-partum vaginal prolapse (known as bearings in sheep) occurs within a few weeks prior to lambing and unless treated both ewes and unborn lambs will die. Rates of prolapse in New Zealand vary from 0.1 to 5% per annum, varying between season and farms. It is a worldwide problem which also affects the 70 million Australian sheep population. Much research has been undertaken over many years to determine the cause of this condition but no clear etiology has emerged [1]. In this study plasma samples were collected prior to prolapse occurring. An improved method for running sheep plasma on 2D gels was developed resulting in improved spot resolution along with a lower coefficient of variation for spot volume [2]. Using this improved method samples were subjected to 2D DIGE (differential in gel electrophoresis) to determine if there were differences between the protein profiles of ewes that were about to prolapse and control ewes. Results show only a few differences between controls and pre-prolapse samples. One of these was haptoglobin, a major acute phase protein in ruminants, in which some isoforms were upregulated prior to prolapse occurring. The haptoglobin isoform spot volumes were added together to calculate a global haptoglobin response and a haptoglobin assay was then used on the same samples which correlated well with the global haptoglobin gel data. Another finding was that alpha-1B-glycoprotein was down regulated close to prolapse. It also was found that ewes closest to prolapse had increased plasma cortisol concentrations compared to controls. Field observations indicate that ewes can prolapse whilst standing, which should not be possible if they are relaxed, due to negative intra-abdominal pressure [3]. Together these observations raise the question “could prolapse be caused by heightened anxiety?” This question and the biochemical analyses will be discussed.

1. Jackson, R., et al., Epidemiology of vaginal prolapse in mixed-age ewes in New Zealand. New Zealand Veterinary Journal, 2014. 62(6): p. 328-337.
2. Brown, S. and G. Norris, Improved consistency in 2D gel electrophoresis: Sheep plasma as a test case. Electrophoresis, 2017. 38(6): p. 906-913.

3. McLean, J.W. and J.H. Claxton, Vaginal Prolapse in Ewes, Part VII: The measurement and effect of intra-abdominal pressure. The New Zealand Veterinary Journal, 1960. 8(3): p. 51-61.

65

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

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

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. 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 nitrogen species (ROS/RNS). This increase in ROS/RNS overwhelms cellular antioxidant defence mechanisms and can alter protein structure / function via various protein post-translational modifications (PTMs). The most redox active amino acid is Cysteine (Cys) and Cys redox PTMs can be classed as either those that are biologically reversible (e.g. *S*-glutathionylation) or 'irreversible' (sulfinic and sulfonic acid; Cys-SO_2H/SO_3H). Irreversible Cys redox PTM occur with sufficient 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 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. 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.

66

Proteomic heterogeneity of high-grade serous ovarian carinomas

Rohan Shah[1], Maiken M Espersen[3, 2], Steve Williams[1], Srikanth Manda[1], Natasha Lucas[1], Dylan Xavier[1], Sadia Mahboob[1], Andrew Robinson[1], Peter G Hains[1], Brett Tully[1], Roger Reddel[1, 4], Phil J Robinson[1, 4], Rosemary Balleine[1, 3, 4], Qing Zhong[1], Anna deFazio[3, 2, 4]

1. Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

2. Centre for Cancer Research, Westmead Institute for Medical Research, Sydney, NSW, Australia

3. Sydney West Translational Cancer Research Centre, Sydney, NSW, Australia

4. Sydney Medical School, University of Sydney, Sydney, NSW, Australia

Tumour heterogeneity poses significant challenges in the development and implementation of personalised cancer treatments, and the optimisation of sampling strategies for diagnostic assays. Expanding our understanding of cancer proteomic heterogeneity is essential to improving cancer treatment (Dagogo-Jack et al, 2018).Ovarian cancer is currently the most lethal gynaecological malignancy, of which high-grade serous ovarian cancer (HGSOC) is the most common subtype. Studies of HGSOC have reported spatial heterogeneity at a genomic level (Schwarz et al, 2015), but this had yet to be explored at a proteomic level. In this study, we explored within-patient and between-patient heterogeneity in eleven HGSOC patients using sequential window acquisition of all theoretical fragment ion spectra mass spectrometry (SWATH-MS).We analysed matched samples taken from primary ovarian tumours and omental metastasis. The MS data acquisition was performed in duplicate, with technical replicates being run on independent 6600 triple TOF instruments. The dataset consisted of 1,354 SWATH-MS runs, the majority representing 454 sections of cryopreserved cancerous tissue. A spectral library was created from 145 LC MS/MS runs in DDA mode. The raw wiff files were searched with ProteinPilot against the Uniprot database, to generate nine different libraries. The libraries were filtered at 1% peptide FDR. A merged library, used for the SWATH-MS runs, contained 76,209 peptides. The SWATH data was filtered at 1% PSM FDR, and 67,775 peptides were found in more than ten SWATH-MS runs.

Our analysis shows that the variation between patients is greater than the local variation within a tumour. The local variation within a tumour can, to some extent, be explained by histological content variation. We show that patient specific proteomic signatures can distinguish between samples from different patients. Our large dataset also highlights the effects of confounding variables, such as differences between mass spectrometers, and sample preparation methods. These confounding variables often cause differences in the proteins identified as having non-zero intensities, rather than causing relative differences in intensities. This has important consequences for the interpretation of proteomic analysis of large cohorts.

1. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol. 2018;15(2):81–94.
2. Schwarz RF, Ng CKY, Cooke SL, Newman S, Temple J, Piskorz AM, et al. Spatial and Temporal Heterogeneity in High-Grade Serous Ovarian Cancer: A Phylogenetic Analysis. PLOS Med. 2015;12(2):e1001789.

67

High-throughput sample preparation for proteomic analysis using 3d-printing

Dylan J Harney[1], Amy T Hutchinson[2], Luke Hatchwell[1], Sean J Humphrey[1], David E James[1], Samantha Hocking[3], Leonie K Helibronn[2], 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

Human plasma derived from clinical trials is one of the most difficult sample sets to analyze using mass spectrometry-based proteomics, due to the extensive sample preparation required and the need to process many samples to achieve statistical significance. Here, we describe an optimized and accessible device (Spin96) to accommodate up to 96 StageTips, a widely used sample preparation medium enabling efficient and consistent processing of samples prior to LC-MS/MS. Spin96 is a 3D-printed device, designed to be compatible with any centrifuge with a deep-well adaptor. The design of the Spin96 minimizes sample loss and eliminates cross-contamination while enabling use of all commonly used chromatography resins such as C18, mixed-mode SDB-RPS and titanium dioxide. To demonstrate the efficacy of Spin96, we analysed human plasma from a clinical trial of Intermittent Fasting (IF). IF increases lifespan and decreases metabolic disease phenotypes and cancer risk in model organisms, but the health benefits of IF in humans is less clear. In this longitudinal study employing 8-weeks IF, we identified significant abundance differences induced by the IF intervention, including increased apolipoprotein A4 (APOA4) and decreased apolipoprotein C2 (APOC2) and C3 (APOC3). These changes correlated with a significant decrease in plasma triglycerides after the IF intervention. Given that these proteins have a role in regulating apolipoprotein particle metabolism, we propose that IF had a positive effect on lipid metabolism through modulation of HDL particle size and function. In addition, we applied a novel human protein variant database to detect common protein variants across the participants. We show that consistent detection of clinically-relevant peptides derived from both alleles of many proteins is possible, including some that are associated with human metabolic phenotypes. Together, these findings illustrate the power of our Spin96 device for high throughput proteomics to yield biological insight of clinical significance.

68

Deep neuroproteomics reveal defective synaptic signalling in excitotoxicity

Syeda Sadia Ameen[1], Ching-Seng Ang[1], Joe Ciccotosto[1], Ashfaqul Hoque[2], Bruno Catimel[3], Antoine Dufour[4], Nicholas Williamson[1], Heung-Chin Cheng[1]

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

2. St. Vincent's Institute of Medical Research, , , Fitzroy, Vic, Australia

3. Walter and Eliza Hall Institute of Medical Research, Parkvile, Melbourne, Victoria, Australia

4. University of Calgary, Calgary, Canada

One of the main underlying mechanism of stroke, trauma and other neurodegenerative diseases is excitotoxicity which remain nebulous to date. Excitotoxicity is caused by over-stimulation of ionotropic glutamate receptors mainly NMDA receptors and lead to uncurbed accumulation of intracellular calcium ions. As a result there is substantial imbalances in the intracellular environment (regulation and signalling) that activate various programs of neuronal death. While many studies have done to find the whole proteome of mouse brain none have scrutinized the functional impairment of regulatory signalling proteins in excitotoxicity. We implemented for the first time a combination of phosphoproteomics and N-terminomics approach to understand two important post translation modifications (phosphorylation and truncation) of proteins during neuronal excitotoxicity in a time dependent manner. We quantified a total of 4755 phosphopeptides and 3843 N-termini. To unravel global phosphoproteome changes we deliberately choose a higher confident set of phosphopeptides (identified in two search engines) 34.1% of the total and ultimately found 156 and 397 phosphopeptides that significantly (±2 fold) fluctuate at two different time points in neuronal excitotoxicity. Besides changes in phosphoproteomes in excitotoxicity we found 98 and 509 cleavage sites in N-terminomics experiment. All the significantly dismayed proteins direct synaptic signalling defects (axonal guidance and signalling pathway ($p <$ 0.01) during excitotoxicity. From these findings we can pinpoint the group of proteins that have dysregulated phosphorylation and truncation in excitotoxicity. Therefore, those proteins perturbed in excitotoxicity may suggest new directions in the investigation of the underlying molecular mechanisms accountable for the pathophysiological consequences of excitotoxicity.

69

The sweet connections in neurotransmission – investigating sialylated glycosylation in the active zone of nerve terminals

Inga Boll[1], Pia Jensen[1], Martin R Larsen[1]

1. University of Southern Denmark, Odense, Denmark

Even though the importance of sialic acids and sialylated glycoproteins for brain development was observed about three decades ago, most global studies regarding the cell biology of the neuron focused on genomics or proteomics, but barely touched upon PTMomics and especially not on glycoproteomics. Examining the signalling of nerve terminals during synaptic transmission, most studies investigated the alteration of phosphorylation, because it is known to be fast and reversible. Nevertheless, glycan structures can be even more versatile than other PTMs due to their different isomeric monosaccharides, being combined by different linkages and branching. Our study applying sialylated glycopeptide enrichment and MS-based quantitative proteomics clearly shows that the role of sialylated glycoproteins in neuronal signalling upon depolarization has been underestimated. By using isolated nerve terminals (synaptosomes) from rat brains and subsequent depolarization with KCl, the modulation of sialylated glycosylation was determined by using state-of-the-art methods for enrichment of sialylated glycopeptides combined with quantitative proteomics. Mapping the alterations after 5 seconds of depolarization led to the identification of 2419 unique

sialylated formerly N-linked glycopeptides whereof 158 were upregulated and 328 downregulated with respect to sialylation. Furthermore, the enrichment of an active zone-like fraction revealed changes at the site of neurotransmitter release in detail. In this fraction, important synaptic proteins such as calcium, potassium and sodium channels, glutamate receptors, ATPases, sodium and potassium transporters as well as extracellular matrix, cell migration and adhesion proteins were regulated, supporting the importance of sialic acids in the signalling taking place in nerve terminals. Especially for ion channels and transporters, the addition or removal of sialic acids seemed to be dynamic suggesting that the negative charge modulates structure and function of these proteins. This would provide a hitherto unknown level of regulation of neuronal signalling.

70

Rapid qualitative and absolute quantification of plasma based proteins using a novel scanning quadrupole DIA acquisition method

Lee A Gethings[1], Chris J Hughes[1], Sarah Lennon[1], Sebastian Muller[2], Robert Plumb[3]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Biognosys, Schlieren, Switzerland

3. Waters Corporation, Milford, MA, USA

Quantitative proteomics often incorporates the use of stable isotope labels (SILs) to provide absolute quantification. Recent advancements have seen the introduction of peptide panels allowing the quantification of over 500 proteins in plasma sample sets. However, this is technically challenging when attempting to acquire the data using more traditional MS acquisition modes such as multiple reaction monitoring (MRM), since the duty cycle of the instrument is compromised and therefore results in under sampling. An alternative approach is to apply a data independent analysis (DIA) methodology, allowing for high throughput whilst ensuring high rates of data acquisition and specificity. We describe the use of a SIL kit capable of quantifying >500 plasma-based proteins in conjunction with SONAR, to quantify proteins of interest for patient cohorts diagnosed with respiratory disorders.

Undepleted human plasma originating from controls and patients diagnosed with chronic obstructive pulmonary disorder (COPD) and asthma were reduced, alkylated and tryptically digested overnight. Prior to LC-MS analysis, samples were spiked with PQ500 SIL peptides (Biognosys). Samples were separated using a 1mm scale column over various LC gradients (15, 30 and 60 minutes). MS data were acquired using SONAR, whereby the quadrupole (MS1) was continuously scanned between *m/z* 400-900 using a quadrupole transmission width of approximately 20 Da, whilst the TOF scanned between m/z 50-2000. In all cases, precursor and product ion information were collected. Data were processed with Spectronaut Pulsar X (Biognosys AG and Progenesis QI for Proteomics (Non-linear Dynamics). Acquiring data over a 1mm scale demonstrates greater robustness and highlights flexibility of the workflow for multi-OMICS based applications. Multivariate statistical analysis showed distinct differences between all three cohorts and proteins corresponding to >86% of the spectral library were quantified. CV's for each group were all found to be <10% in all cases and was maintained at high throughput (i.e. 15 minutes).

71

Glycopeptide fragmentation optimization and quantitation by multi collision energy ramp scanning quadrupole data independent acquisition

Lee A Gethings[1], Christopher J Hughes[1], YiJu Chen[2], YuJu Chen[2], Robert S Plumb[3], Johannes PC Vissers[1]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Academia Sinica, Taipei, Taiwan

3. Waters Corporation, Milford, MA, USA

Data-independent acquisition (DIA) approaches have gained increasing popularity due to their high reproducibility. Both targeted and discovery workflows are now routinely used as they can provide both qualitative and quantitative information in a single experiment. Here, we describe and evaluate a novel DIA acquisition method, termed scanning quadrupole DIA, whereby the resolving quadrupole of a high-resolution benchtop Q-TOF instrument is repetitively scanned using overlapping precursor m/z windows to enable both high sensitivity and reproducible acquisition of all fragment ions. LC-MS data were collected in a modified data independent mode of acquisition using continuous quadrupole scanning between *m/z* 800 to 1600. oa-TOF mass spectra were recorded for each quadrupole position and stored into 200 discrete bins. Three alternating data functions (modes) were acquired: in the low energy MS1 mode, data were collected at 6eV collision energy without quadrupole scanning. In two separate elevated energy MS2 modes, the collision energy was ramped from 12 to 23 eV and from 27 to 50 eV. The spectral acquisition times were between 0.1 s and 0.5 s for each mode. Assessment of the raw data shows that implementation of the hybrid mode of acquisition results in improved fragmentation when compared with data collected using a single collision energy ramp. For each raw data file, in case of untargeted, library-independent searches, two processed peak list files were generated from the individual high-energy CID MS2 data streams. . In-solution digest from tryptic haptoglobulin (HP), a defense response regulator, was used as a model system for method development. Immunoaffinity and HILIC glycopeptide HP enriched plasma sample from three human lung cancer cell lines were contrasted in terms of their relative quantitative glycopeptide profiles. Over 30 potential different glycan structures were characterized, with a significant number showing over expression for cell lines resulting from patients with non-small cell carcinomas.

72

Multi-omic characterisation of bladder and lung carcinomas using a novel scanning quadrupole DIA acquisition method

Lee A Gethings[1], Adam King[1], Heather Patsiouras[2], Robert Plumb[3]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Sydney, NSW, Australia

3. Waters Corporation, Milford, MA, USA

Cancer is one of the most complex, life threatening diseases, existing in many forms which have unknown pathogenesis. A combination of genetic and lifestyle factors are known to contribute towards increasing the probability of encountering cancer. Plasma samples were prepared for proteomic, metabolomic and lipidomic analyses. Label-free LC-MS data were acquired using a oa-QTof platform utilizing a broadband acquisition technique (SONAR) data independent acquisition workflow. Small molecule analysis consisted of using Progenesis QI for data processing to provide normalised values prior to statistical analysis. Unsupervised MVA of the data showed clear distinction between cohorts. OPLS-DA was used to filter for features of significant correlation and covariance prior to identification using HMDB (metabolites) and LipidMaps (lipids). Identifications matching the following criteria, mass accuracy <5 ppm, ANOVA p <0.05, %CV <30 and fold change >2 were considered for further interrogation. SONAR-based analysis indicates that the scanning quadrupole DIA enables over an order of magnitude more specificity than a static quadrupole operated with the same resolution and it was found that a quadrupole transmission window of approximately 10 Da provided optimum identifications. Proteomic data were processed using Progenesis QI for Proteomics and searched against a Uniprot *Homo sapien* database, containing reviewed entries and limited to 1% FDR. Additionally, we also searched the data against a spectral library using Spectronaut and comparatively analysed the results from both workflows. A number of significant proteins with differential regulation were exhibited for a number of protein groups involved in antigen and lipid binding. Proteins occurring in a minimum of two out of three replicates and with ANOVA p <0.05 were considered. Biological significance of the results was established by merging data from all three omic experiments and performing pathway analysis. A number of significant pathways including complement activation, B cell mediated immunity and receptor signalling were identified as key pathways.

73

Proteomic profiling analysis of Australian seagrass, *Zostera muelleri* under copper stress

Nasim Shah Mohammadi[1], Manoj Kumar[1], Matt Padula[1], Mathieu Pernice[1], Peter Ralph[1]

1. University of Technology Sydney, Broadway, NSW, Australia

Seagrasses are subjected to a wide range of environmental stresses, including pollution, as a consequence of population growth. Industrial waste as well as agricultural and domestic run-off, often rich in trace elements, have been shown to be major sources of environmental pollutants with deleterious effects on seagrasses mainly at physiological level. The molecular response to trace metal toxicity in seagrasses remains poorly understood compared to terrestrial plants. The emergence of omics techniques such as transcriptomics and proteomics allows investigation of the molecular mechanisms and pathways driving the stress response. Therefore, valuable information can be provided for identification of biomarkers for early detection of toxicity to elevated levels of trace elements, including copper. In this project, the leaf-specific transcriptome of an Australian seagrass, *Zostera muelleri* was initially investigated after 7 days exposure to 0.25 and 0.5 mg/L of copper concentrations. The results showed a dose-dependent response in genes involved in photosynthesis, energy production and defence mechanisms (enzymatic and chemical). Quantitative proteomics was subsequently performed using iTRAQ (isobaric tags for relative and absolute quantification). Around 170 proteins were identified as differentially expressed under 0.5 mg/L of copper stress. The results of this study provide new insights on the molecular mechanisms driving heavy metal toxicity in seagrasses at both translational and post-translational level, and assist in the identification of potential biomarkers for early detection of toxic exposure in seagrass. This new knowledge can provide valuable information for implementing more effective conservation strategies for seagrass meadows.

74

Larval development of the barber's pole worm is under tight post-transcriptional control

Guangxu Ma[1], Tao Wang[1], Pasi K. Korhonen[1], Ching-Seng Ang[2], Nicholas A. Williamson[2], Neil D. Young[1], Andreas J. Stroehlein[1], Ross S. Hall[1], Anson V. Koehler[1], Andreas Hofmann[3], Robin B. Gasser[1]

1. Department of Veterinary Biosciences, Melbourne Veterinary School, The University of Melbourne, Parkville, Victoria, Australia

2. The Bio21 Institute of Molecular Science and Biotechnology, The University of Melbourne, Melbourne, Victoria, Australia

3. Griffith Institute for Drug Discovery, Griffith University, Nathan, Queensland, Australia

We explored molecular alterations in the developmental switch from the L3 to the exsheathed L3 (xL3) and to the L4 stage of the barber's pole worm (*Haemonchus contortus*) using an integrated proteomic, transcriptomic and bioinformatic approach. About 16% of 9842 transcripts in the transcriptome were expressed as proteins (n = 1596). There was a positive correlation (r = 0.39-0.44) between transcription and protein expression for distinct developmental stages of the nematode. There was a negative correlation (r = -0.6 to -0.5) in the differential protein expression between developmental stages upon pairwise comparison. Changes in protein expression from the free-living to the parasitic phase of *H. contortus*related to enrichments in biological pathways associated with metabolism (e.g., amino acid metabolism and carbohydrate and lipid degradation), environmental information processing (e.g., signalling and interactions) and/or genetic information processing (e.g., translation). Fatty acid degradation and steroid hormone biosynthesis were suppressed, whereas translation and protein processing in the endoplasmic reticulum were upregulated in the transition from the free-living L3 to the parasitic larval stages of the nematode. Post-transcriptionalregulation was inferred to elicit these alterations, with particular miRNAs likely having roles in environmental

adaptations and/or stress responses during developmental transitions. The results of this study provide a comprehensive insight into the developmental biology of this economically important worm at the molecular level, and the methodologies used are readily applicable to other parasites.

75

Chemical and random additive noise elimination (crane)

Akila Seneviratne[1], Brett Tully[1]

1. Children's Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

Introduction - Data generated by Data Independent Acquisition mass spectrometry (DIA-MS) is used to both identify and quantify peptides within a biological sample. The output of DIA-MS is influenced by multiple sources of noise that may be due to impurities introduced during sample preparation, reagents used in the liquid chromatography column, or from particles in the atmosphere. Most of these impurities cause chemical noise; its signal is distributed over a wide range in the retention time, has constant in mass over charge (m/z) and usually has a single charge. Random noise is generated by the machine and depends on its operating conditions such as gain, temperature and age. Random noise is of high frequency and could occur at any point in m/z vs retention time space. In micro-channel plate detectors, the collision of an ion with the detector generates a signal. When a host of ions collide with the detector there could be an accumulation of residue at the detector which results in a shift in the baseline. Many attempts have been made to de-noise MS data. Most attempts try to de-noise spectra [1-8], yet important spatial information is lost when the MS data is summed over retention time. Other techniques [9, 10] perform de-noising of single ion chromatograms, and unlike [1-8] some of the multi-dimensional information of the MS data is explored by these methods. However, these are not strictly two-dimensional de-noising techniques. Method - Crane is a technique for de-noising raw MS data via two-dimensional undecimated wavelet transform inspired by image de-noising techniques. Preliminary results show that it is a promising technique for removing chemical and random noise. Ultimately, this technique should increase the accuracy and reduce false detection rates of peptide / fragment detection pipelines such as OpenSWATH.

1. 1. Halima Bensmail, Jennifer Golek, Michelle M. Moody, John O. Semmes and Abdelali Haoudi, "A novel approach for clustering proteomics data using Bayesian fast Fourier transform", Bioinformatics, Volume 21, Issue 10, May 2005, Pages 2210–2224
2. 2. Kevin R. Coombes, Spiridon Tsavachidis, Jeffrey S. Morris, Keith A. Baggerly, Mien-Chie Hung and Henry M. Kuerer, "Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform" Proteomics, Volume 5, Issue 16, 2005, Pages 4107-4117
3. 3. Chao Yang and Weichuan Yu, "A brief review of signal processing issues in mass spectrometry-based proteomics studies," 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, 2011, Pages 1036-1040
4. 4. Deukwoo Kwon, Marina Vannucci, Joon Jin Song, Jaesik Jeong and Ruth M. Pfeiffer, "A novel wavelet-based thresholding method for the pre-processing of mass spectrometry data that accounts for heterogeneous noise", Proteomics, Volume 8, Issue 15, Aug 2008, Pages 3019-3029
5. 5. Peicheng Du, Gustavo Stolovitzky, Peter Horvatovich, Rainer Bischoff, Jihyeon Lim and Frank Suits, "A noise model for mass spectrometry based proteomics", Bioinformatics, Volume 24, Issue 8, Apr 2008, Pages 1070-1077
6. 6. Ole Schulz-Trieglaff, Egidijus Machtejevas, Knut Reinert, Hartmut Schlüter, Joachim Thiemann and Klaus Unger, "Statistical quality assessment and outlier detection for liquid chromatography-mass spectrometry experiments", BioData Mining, 2009;2:4. doi:10.1186/1756-0381-2-4
7. 7. Jan Schlenkea, Lars Hildebranda, Javier Morosb and J. Javier Laserna, "Adaptive approach for variable noise suppression on laser-induced breakdown spectroscopy responses using stationary wavelet transform", Analytica Chimica Acta, Volume 754, 2012, Pages 8–19
8. 8. Xiaoli Li, Jin Li and Xin Yao, "A wavelet-based data pre-processing analysis approach in mass spectrometry", Computers in Biology and Medicine, Volume 37, 2007, Pages 509–516
9. 9. Salvatore Cappadona, Fredrik Levander, Maria Jansson, Peter James, Sergio Cerutti, and Linda Pattini, "Wavelet-Based Method for Noise Characterization and Rejection in High-Performance Liquid Chromatography Coupled to Mass Spectrometry", Analytical Chemistry, Volume 80, Issue 13, 2008, Pages 4960-4968
10. 10. Xiaoran Ning, Ivan W. Selesnick and Laurent Duval, "Chromatogram baseline estimation and denoising using sparsity (BEADS)", Chemometrics and Intelligent Laboratory Systems, Volume 139, 2014, Pages 156–167

76

Multi-omic approaches for elucidating the assembly and function of mitochondrial respiratory chain complexes

Boris Reljic[1], Joanna Sacharz[1], Daniella Hock[1], David P. De Souza[2], Dedreia Tull[2], Michael T. Ryan[3], David A. Stroud[1]

1. The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, Victoria, Australia

2. Metabolomics Australia, The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, Victoria, Australia

3. Department of Biochemistry and Molecular Biology, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Mitochondrial oxidative phosphorylation (OXPHOS) disorders represent the most common genetic form of inherited metabolic diseases, affecting ~1/5000 births with a diagnosis rate of only ~60%. In addition, only very limited proven treatments are available and those fortunate enough to receive a molecular diagnosis often wait years. Our lab is developing the use of computational proteomics and metabolomics tools to complement existing diagnostic strategies, discover new disease genes, and understand the impact of mitochondrial dysfunction on a systems-wide level. The presented work will provide an update on our project to

generate CRISPR/Cas9 gene disrupted cell lines of every known OXPHOS subunit and assembly factor – in total this represents >120 nuclear encoded genes. In particular I will be discussing the molecular roles of individual complex II and III subunits within wider mitochondrial function, as well as new data describing the roles of existing and novel assembly factors in respiratory chain complex assembly. Using quantitative mass-spectrometry analyses of mitochondria from individual knockout cell lines we quantified the levels of ~80% of the known mitochondrial proteome, allowing us to understand the impact of the different classes of complex II and III dysfunction on complex assembly, as well as the response made by various other mitochondrial biological processes. Knockouts could be grouped by hierarchical clustering, likely into structural 'modules' giving insights into assembly of both complexes. Complementary to this, I will present our latest metabolic profiling data demonstrating the unusual roles of complex II subunits and assembly factors on the tricarboxylic acid cycle (TCA cycle). The presented data not only confirms but also expands on the pivotal role of complex II/III within the electron-transport chain, respiratory supercomplex and mitochondrial biogenesis. Deeper interrogation of our data into how the mitochondrial proteome and metabolome responds to respiratory chain dysfunction will potentially allow the identification of novel assembly factors which represent potential mitochondrial disease markers.

77

Sexual dimorphism and epileptogenesis in developing and adult rat brains

Jonas B Albinas[1], Merridee A Wouters[1], Madeleine J Otway[1], Peter G Hains[1], Martin R Larsen[2], Pablo Espinosa[3], Terence O'Brien[3], Nigel Jones[3], Qing Zhong[1], Phil J Robinson[1]

1. Children's Medical Research Institute, Faculty of Medicine and Health, University of Sydney, Westmead, NSW, Australia

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 levels in epileptogenesis during seizures,[1] as well as changes in expression of actin, profilin 1 and a-synuclein.[2] We performed a dual study of gender differences and epileptogenesis in developing rat brains by comparing samples from 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 (Genetic Absence Epilepsy Rats) have a deficit in Cacng2, a T-type voltage-dependent calcium channel. The WAG/Rij (Wistar Albino Glaxo from Rijswijk) rats, on the other hand, have hyperpolarized HCN channels that need higher intracellular concentrations of cAMP for activation. 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 mutant rat models are known to display AE, the age of onset differs. GAERs have regular seizures from the onset of puberty at around 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, on the other hand, 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 specific to the two different epileptic rat models. The data show both that deep proteomic brain analysis needs to consider gender and that proteins changes in the brain can occur well before seizure onset.

1. 1. Thind KK, Yamawaki R, Phanwar I, Zhang G, Wen X, Buckmaster PS (2009) Initial loss but later excess of gabaergic synapses with dentate granule cells in a rat model of temporal lobe epilepsy. Journal of Comparative Neurology 518: 647-667. DOI: 10.1002/cne.22235
2. 2. Li A, Choi Y-S, Dziema H, Cao R, Cho H-Y, Jung YJ, Obrietan K (2010) Proteomic profiling of the epileptic dentate gyrus. Brain Pathology 20: 1077-1089. DOI: 10.1111/j.1750-3639.2010.00414.x.

78

Fully automated all-in-one sample preparation for high-throughput spectral library generation and SWATH-based quantitation

Pawel Sadowski[1], Selvam Paramasivan[1, 2], Nana Satake[2], Rajesh Gupta[1], Paul Mills[2]

1. Queensland University of Technology, Brisbane, QLD, Australia

2. The School of Veterinary Science, The University of Queensland, Brisban, QLD, Australia

Next generation clinical mass spectrometry that relies on SWATH data acquisition is transforming translational research. 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 quantitation. Here, we have developed 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 96-well 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 membrane without compromising reproducibility. We leveraged the improved throughput to create a compendium of cattle proteins from 24 clinically relevant tissues and body fluids and demonstrated uniform and effective digestion performance across all tissue types that allowed us to reproducibly quantify nearly 6,000 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.

79

Tissue-specific immunopeptidome of HLA-B*57:01 in drug hypersensitivity reactions

Johanna E E Tuomisto[1], Asolina Braun[1], Patricia T Illing[1], Nicole A Mifsud[1], Anthony W Purcell[1]

1. Monash University, Clayton, VIC, Australia

Publish consent withheld

80

Protein O-fucosyltransferase 2 -mediated *O*-glycosylation of MIC2 is dispensable for *Toxoplasma gondii* tachyzoite infection

Sachin Khurana[1], Michael J Coffey[1], Alan John[1], Alessandro D Uboldi[1], My-Hang Huynh[2], Rebecca J Stewart1[1], Vern B Carruthers[2], Christopher J Tonkin[1], Ethan Goddard-Borger1[1], Nichollas E Scott[3]

1. The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia

2. Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI

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

Toxoplasma gondii is a ubiquitous obligate intracellular eukaryotic parasite that causes congenital birth defects, disease of the immunocompromised and blindness. Protein glycosylation plays an important role in the infectivity and evasion of immune response of many eukaryotic parasites and is also of great relevance to vaccine design. Here, we demonstrate that MIC2, a motility-associated adhesin of *T. gondii*, has highly glycosylated thrombospondin repeat domains (TSR). At least seven *C*-linked and three *O*-linked glycosylation sites exist within MIC2, with >95% occupancy at *O*-glycosylation sites. We demonstrate that the addition of *O*-glycans to MIC2 is mediated by a Protein *O*-fucosyltransferase 2 homologue (TgPOFUT2) encoded by TGGT1_273550. While POFUT2 homologues are important for stabilizing motility associated adhesins and host infection in other apicomplexan parasites, in *T. gondii* loss of TgPOFUT2 has only a modest impact on MIC2 levels and the wider proteome. Consistent with this, both plaque formation and tachyzoite invasion are broadly similar in the presence or absence of TgPOFUT2. These findings demonstrate that TgPOFUT2 *O*-glycosylates MIC2 and that this glycan is dispensable in *T. gondii* tachyzoites.

81

A slimy situation: using integrated label-free quantitative proteomics to uncover potential neuropeptides within the defensive slime secretion of the striped pyjama squid (*Sepioloidea lineolata*)

Nikeisha J Caruana[1], Jan M Strugnell[2], Pierre Faou[3], Julian Finn[4], Kim Plummer[5], Ira R Cooke[6]

1. Department of Ecology, Environment and Evolution, La Trobe University, Melbourne, VIC, Australia

2. Centre for Sustainable Tropical Fisheries and Aquaculture, James Cook University, Townsville, QLD, Australia

3. Department of Biochemistry and Genetics, La Trobe Institute for Molecular Sciences, Melbourne, VIC, Australia

4. Sciences, Museums Victoria, Melbourne, VIC, Australia

5. Department of Animal, Plant and Soil Sciences, AgriBio, La Trobe University, Melbourne, VIC, Australia

6. Department of Molecular and Cell Biology, James Cook University, Townsville, QLD, Australia

Cephalopods exhibit an extensive range of secretions, including ink, mucus and venom. Sepiadariidae, one family of benthic bobtail squids, possess specialised systems of secretion, secreting a viscous slime from their underside. It is believed that these secretions are involved in defending the species, and exhibit unique biochemical and biophysical properties including dramatic volume expansion, adhesion, and antimicrobial defence. In order to further understand the mechanisms behind the secretion of the slime and the proteins involved, six tissues from four individuals of a species of Sepiadariidae, the striped pyjama squid *(Sepioloidea lineolata)*, underwent label-free quantitation. The proteomic results were combined with a *de novo* transcriptome and analysed using computational bioinformatics methods, including a proteogenomic pipeline developed in Galaxy. Twenty-eight highly differentially expressed proteins were identified within the slime of *S. lineolata,* with peptidases and proteases being the most prominent. This was further evidenced by the highly differentially expressed enzymes involved in the post-translational modifications of neuropeptides. While none of the proteins commonly expressed through these pathways were identified (catechloamines and insulin among others), it is possible the novel, short and secreted proteins that were highly differentially expressed, are unclassified neuropeptides. These identified proteases and potential neuropeptides may have further implications for biomimetics within the design of new pharmaceuticals, antimicrobials and antifungals. While traditionally quantitative proteomics is not used to analyse complete proteomes on non-model organisms, our work illustrates that the combination of *de novo* assembled transcriptomes and comprehensive bioinformatic analysis can provide rapid and informative investigation into species, which have high potential for biological applications without the need for a full genome. Our work incorporates the first label-free quantification analysis of a cephalopod secretion and the first proteomic analysis of the tissues from Sepiadariidae.

82

Optimizing red blood cell protein extraction for biomarker discovery

Stephan Klatt[1], Anne Roberts[1, 2], Blaine Roberts[1, 2]

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

2. Cooperative Research Centre for Mental Health, Parkville, Victoria 3052, Australia

Red blood cells (RBCs) are the most common cell type found in the blood. However, not many biomarker assays have been conducted on RBCs. This can be attributed to the large dynamic range of proteins, high abundance of lipids and metabolites, and difficulties accompanying hemoglobin interference particularly with colorimetric and ELISA assays. In this study, we developed a semi-quantitative assay by comparing the efficiency of urea, sodium deoxycholate, acetonitrile and HemoVoid™ to extract the RBC proteome. Protein lysates were analyzed on an Agilent 6495 QQQ instrument, targeting 144 proteins. The results confirm that RBC protein extraction with HemoVoid™ leads to a tremendous reduction of hemoglobin and subsequently to a detection increase of most remaining proteins. In addition, peptide normalization to the RBC housekeeping proteins resulted in very stable coefficient of variation (CV) values. Hemoglobin interference after deoxycholate and urea extraction is still strong, but signal intensities even for low abundant proteins are in a satisfying range, and most CV values are strong. Extraction with Acetonitrile leads to an overall decreased protein abundance probably due to protein precipitation. High-throughput feasibility is most promising after urea extraction. We therefore recommend RBC protein extraction with either HemoVoid™ or urea.

83

How bad is my child's burn. Differences in blister fluid biochemistry to assess paediatric burn wounds

Tuo Zang[2, 1], Leila Cuttle[3, 2, 1], Daniel A Broszczak[2, 4, 1], James A Broadbent[2, 1], Catherine Tanzer[5, 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. 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

4. School of Science, Faculty of Health Sciences, Australian Catholic University, Brisbane, Queensland, Australia

5. Griffith University, Gold Coast, Qld, Australia

Burn blister fluid analysis provides an opportunity to non-invasively investigate the biology of the initial response to burn injury; uncover novel diagnostics or prognostics to assist in clinical decision making and enable the identification of new therapeutic approaches to enhance healing. We performed a proteomic analysis of 87 paediatric burn blister fluid samples using liquid chromatography - tandem mass spectrometry with SWATH (data independent) acquisition, which allows for the large scale relative quantification of multiple proteins between samples. The blister fluid proteomes of all samples were compared to the key clinical features of burn depth classification and time-to-reepithelialisation. Both of these clinical parameters are critical for enabling accurate clinical decisions regarding early burn treatment options. Subsequent analyses 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. Importantly, this highlights the utility of diagnostic markers for burn depth and / or prognostic markers of time-to-reepithelialisation. While full-thickness burns are often grafted and superficial-partial thickness burns are often not, it is more difficult to determine if deep-partial thickness burns should be grafted or not. Thus, 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.

84

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, NSW, Australia

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 inflammatory 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 (1), and is a major target for existing and novel antibiotics (2). *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 (3, 4). Proteome analysis revealed several glycoproteins in the *C. jejuni* 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. Abundance changes and structural features of peptidoglycans in each strain were determined. Our results showed that changes in protein glycosylation impacts the structural composition of the PG layer commensurate with changes to PG-associated proteins observed at the proteome level. We also observed altered cell morphology and reduced *C. jejuni* pathogenesis. Unravelling PG structures and the factors involved in their biosynthesis may help in defining why *N*-glycosylation is critically required for *C. jejuni* pathogenesis.

1. Frirdich E et al (2014) *J Biol Chem*. 289(12):8007-18
2. Katayama N et al (2003) *Plant Cell Physiol*. 44(7):776-81
3. Szymanski CM et al (2002) *Infect Immun.* 70:2242-4
4. Karlyshev AV et al (2004) *Microbiol.* 150:1957-64

85

Insulin-related deaths: analysis of insulin and synthetic analogues in coronial exhibits by QE Plus Orbitrap

Nicola Beckett[1], Rebecca Tidy[1], Bianca Douglas[1], Nathan Sumner[1]

1. ChemCentre, Perth, WA, Australia

An overdose of insulin will produce a hypoglycaemic effect in the body, which may prove to be fatal.[1] Insulin, used clinically in the treatment of diabetes, includes synthetic analogues such as Apidra®, Humalog®, Lantus®, Levemir®, Novorapid®. The ability to effectively identify and differentiate the various types of insulin can be very important to an investigation e.g. the presence of synthetic or exogenous insulin in a non-diabetic may be indicative of a poisoning attempt, while abnormally high levels of insulin (exogenous/synthetic or endogenous) in a victim could point to homicide or suicide.[2] Commonly, clinical laboratories analyse insulin in blood by radioimmunoassay. In a forensic context, limited information is obtained via this technique thus highlighting the need for methodology that can provide greater sensitivity, specificity and robustness, and which is also capable of the unequivocal identification of the various insulin analogues at pmol levels — paramount for coronial investigations. The QE Plus Orbitrap can deliver these requirements and therefore provide invaluable information to forensic pathologists and Coroners to assist in their inquiries into insulin-related death investigations.

Forensic toxicology laboratories do not routinely conduct protein analyses for the likes of insulin. Therefore, this project is aimed at bridging those gaps with the development of mass-spectrometry-based proteomics for the extension of ChemCentre's analytical capability into the analysis of 'large' molecules of forensic interest. An extraction method, QE Orbitrap screening and LCMS-QQQ confirmation method was developed to analyse blood and urine of insulin-related WA coronial cases. In circumstances where routine exhibits (blood/urine) are not available an alternative method for the analysis of insulin in vitreous humour was developed and validated. This study highlights an advancement in forensic technology and the promising ability of the QE Plus Orbitrap as an excellent screening tool for insulin-related forensic toxicology and proteomic investigations.

86

Identification of isoaspartylated peptides by electron-transfer/higher-energy collision dissociation

Clifford Young[1, 2], Frank Kjeldsen[3], Michael L Nielsen[1]

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

2. Future Industries Institute, University of South Australia, Adelaide, South Australia, Australia

3. 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.

87

Mass spectrometry-based large-cohort proteomics for precision medicine – An international Cancer Moonshot multiple site study

Andreas Huhmer[1], Yue Xuan[1]

1. Thermo Fisher Scientific, SAN JOSE, CA, United States

To successfully elevate discovery proteomics to translational research in the pipeline of precision medicine, large-cohort studies are essential in discovery and verification of protein biomarkers. However, to reproducibly and reliably quantify large numbers of proteins across different laboratories remain challenging. To address this, we present a high-throughput and streamlined analytical workflow using high resolution MS1-based quantitative data-independent acquisition (HRMS1–DIA) mass spectrometry. HRMS1-DIA workflow is standardized with well-defined experimental steps and systematically applied to set of test samples. The study was benchmarked across multiple Cancer Moonshot sites utilizing identical instrument, methods and software, demonstrating stability in a 24/7 operation mode for 7 consecutive days.

Successful proteome profiling workflows must address highly complex samples. Our approach is to increase chromatographic and mass spectral resolution, utilize HRAM MS1 data for quantitation and interspersed DIA for qualitative analysis, spiking quality control peptides, and creating in-depth spectral libraries. Besides setting FDR at 1%, roll-up statistic strategy was applied to improve quantitation precision. Robust and straightforward SOPs are created for HRMS1-DIA workflow, to define instrumental aspects such as retention time stability, spray stability, and product ion distribution overlap with spectral library.

The workflow was transferred to multiple sites to perform similar measurements for assessment. Each laboratory was equipped with similar instruments, columns, methods, spectral libraries and software. To evaluate robustness of the workflow, the study was carried out in a 24/7 operation mode for 7 consecutive days. The resulting data were processed individually and combined to evaluate proteome coverage and quantitative capabilities. At 1% FDR, > 5,000 proteins from > 40,000 peptides of the QC sample, as well as > 7,000 proteins from > 50,000 peptides of the mixed proteome sample, are consistently detected and reliably quantified across sites. The ratios of the mixed three proteomes accurately reflect the expected values.

88

SWATH-MS proteomics enabled studying the semen quality in Brahman bulls

Rajesh Gupta[1], Nana Satake[2], Boe-Hansen G. B[2], Rêgo J. P. A[3], Venus B[4], Nouwens A. S[5], Li Y[6], Pawel Sadowski[1]

1. Cental Analytical Research Facility, IFE, QUT, Brisbane, QLD, Australia

2. School of Veterinary Science, The University of Queensland, Brisbane, QLD, Australia

3. Federal Institute of Education, Science and Technology of Ceará, Boa Viagem Campus, Brazil

4. Agri-Science Department of Agriculture, Fisheries and Forestry, Brisbane, QLD, Australia

5. School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia

6. CSIRO Livestock Industries Animal, Food and Health Sciences, Brisbane, QLD, Australia

Over the last two decades, proteomics based mass spectrometry (MS) techniques have emerged as powerful tools and have advanced our understanding of the complex biological systems of living organisms. However, the adoption of proteomic approaches in the clinical veterinary field lags behind than in the human medical field. Traditionally, veterinary proteomics investigations are conducted using 2D gel techniques that is laborious and require extensive optimization to achieve reproducible results. Recently, SWATH-MS approach has emerged as reliable alternative but application into veterinary filed is compromised by the limited protein databases available for many animal species. We have enabled SWATH-MS based quantitation in a recent study assessing the change in sperm characteristics occurring after an acute temperature increase obtained through scrotal insulation (SI), in regularly electroejaculated Bos indicus Brahman bulls. A total of six Brahmans were included and SI was applied to three of the bulls for 48 h, and semen was collected by electroejaculation at three days intervals from before (-10 d), until 74 d after initiation of SI. In particular, 418 proteins were identified from a pool of seminal plasma samples using DDA acquisition in order to obtain the spectral library. Subsequent SWATH-MS analysis reproducibly quantified 158 proteins. Significant difference magnitude of change (adjusted $P<0.05$) in protein concentration between the two treatment groups were found in 29 out of 158 proteins. SWATH-MS result was compared to previously acquired 2D gel data on the same set of samples revealing vast agreement between these two techniques.

89

Delineation of the monocyte chemoattractant protein-dependent CCR2 signalling network using phosphoproteomics

Simon Foster[1], C Huang[1], A Shah[1], O Kleifeld[1], R_B Schittenhelm[1], M.J Stone[1]

1. Biochemistry and Molecular Biology, Monash University, Clayton, VIC, Australia

A universal feature of inflammatory diseases is the excessive migration of leukocytes to the affected tissues. Leukocyte recruitment is regulated by chemokines, which are secreted at the site of injury or infection and then activate chemokine receptors, G protein-coupled receptors expressed on the target leukocytes. CCR2 is the major chemokine receptor on monocytes and macrophages, for which the primary cognate ligands are the monocyte chemoattractant proteins (CCL2/MCP-1, CCL7/MCP-3 and CCL8/MCP-2). Genetic deletion of CCR2 or silencing of CCL2 protects mice from developing atherosclerosis, yet numerous clinical trials of CCR2 antagonists have failed, in many cases due to lack of efficacy in phase II or III. One contributing factor for this failure is the complexity and potential compensatory mechanisms inherent in chemokine signalling networks. Thus, to identify appropriate targets for more effective therapeutic approaches, we need a deeper understanding of the chemokine-mediated signalling networks. Here, we describe a large-scale proteomic study using data-independent acquisition (DIA) mass spectrometry to quantify the differences in protein and phosphopeptide levels between untreated and CCL2-treated human endothelial kidney cells stably expressing CCR2. We also conduct a time course study of CCL2 activation to investigate and characterise the kinetics of signalling events. To the best of our knowledge, this is the first time that DIA has been applied to globally quantify the phosphoproteome between multiple conditions. We successfully confirmed the involvement of known canonical pathways (MAPK, JAK/STAT and Akt/mTOR) as downstream effectors of CCL2/CCR2 signalling and we also identified, mapped and manually curated additional signalling networks that have not been associated with the CCL2/CCR2 signalling cascade. These include Rho guanine nucleotide exchange factors (ARHGEFs), nuclear pore complex (NPC) proteins and many actin cytoskeletal proteins.

90

Profiling of ascites proteome in ovarian cancer patients

Dilys TH Leung[1], George Kannourakis[1], Nuzhat Ahmed[1]

1. Fiona Elsey Cancer Research Institute, Ballarat Central, VIC, Australia

Background: Ovarian cancer (OC) is the leading cause of death from gynaecological cancers among women. Vague clinical manifestation attributes to late-stage diagnosis in approximately 75% of OC patients. Surgical debulking and chemotherapy offers marginal benefit to improve clinical outcomes. As OC progresses, ascites (peritoneal malignant tumour fluid) containing detached tumour cells, cancer stem cells (CSC) and pro-oncogenic components accumulate and facilitate tumour spread and survival. We

previously reported CSCs-like phenotypes including enrichment of metabolic pathways and distinct immune surveillance process in the ascites-derived tumour cells of OC patients after (chemoresistant, CR) but not before (chemo-naïve, CN) chemotherapy. We hypothesise that extracellular solutes in ascites contribute to chemoresistance and promote survival of residual CSCs leading to recurrence. The aim of this study is thus to identify the proteome signature of the ascites unique to chemotherapy-resistant OC patients.

Methods: Ascites were collected from advanced-stage CN (n=5) and CR (n=5) serous OC patients. All samples were processed with filter-aided sample preparation (FASP) protocol. Peptides were fragmented and analysed using Thermo liquid chromatography (LC)-Orbitrap Elite™ electron transfer dissociation (ETD) mass spectrometer. Raw data were processed using MaxQuant (v1.5.3.30) and Perseus (v1.5.3.1) workflow and searched against the human UniProt Swiss-Prot database (December 2018, 20,412 protein sequences). Label-free quantification intensities were calculated and fold changes were determined between CN and CR samples. Gene enrichment analysis was performed using FunRich (v3.1).

Results: We have identified a total of 324 proteins, 207 of which are commonly found in both CN and CR ascites. Alpha-1-antitrypsin (SERPINA1) was upregulated by 2-fold in CN compared to CR samples (p<0.001). Gene ontology (GO) analysis revealed an overrepresentation of proteins associated with the regulation of complement activation (GO:0030449), cellular protein metabolic process (GO:0044267) and innate immune response (GO:0045087).

Conclusion: Proteins identified in ascites highlight the existence of metabolic and immune regulatory signatures in the tumour microenvironement that nurtures OC. Further functional analyses need to be performed to determine the importance of the reduced SERPINA1 expression in the development of chemoresistance in OC patients.

91

A multidisciplinary approach to the analysis of complex MSI data

Matthew O'Rourke[1], Ben Roediger[2], Matt Padula[3]

1. The University of Sydney, St Leonards, NSW, Australia

2. Centenary institute, Sydney

3. University of Technology Sydney, Sydney

Mass Spectrometry Imaging (MSI) is a hybrid technique combining the visualisation features of traditional histology with accompanying mass spectrometric data. MSI uses a Matrix Assisted Laser Desorption/ Ionisation (MALDI) mass spectrometer to create spatial maps of targeted biomolecule classes. The spatial maps take the form of 2D images that show the location and relative intensity of a given molecule and the dynamic range of the instrument allows for analyses of virtually any class of ionisable biomolecule without the need for adding detection reporter probes. Though this may seem straightforward the issue becomes the translation of mass spectrometric data; a complex list of numbers, to literal pictures, and how these pictures need to be interpreted. When MSI data is viewed from the perspective of microscopy, MSI data is low resolution, pixelated and does not allow the visualisation or distinction of the whole tissue. By all these counts, the data is not of a publishable quality. Conversely from a mass spectrometry perspective, the mass resolution is very low and the number of detected/ detectible molecules and the confidence that molecular identities are statistically confident is tenuous.

We therefore describe a new approach to the analysis of MSI data that converts it to 2 dimensional images, and then processes them using an Image J based processing pipeline. The pipeline is then able to report R^2 values for the correlation of the detected peptides, thereby confirming their co-localisation on tissue. By using a pixel to pixel correlation of both intensity and presence and absence of a series of peptides for a single protein, peptides can be matched together based on location, to form peptide mass fingerprints (PMF's) which can then be searched through traditional PMF databases to achieve reliable protein identifications. Therefore, this approach allows MSI data to be analysed in isolation without requiring complementation with other instrumentation such as LCMS.

92

Improved lipid profiling workflow using novel iterative MS/MS mode and *in silico* annotation

Mark Sartain[1], Sarah Stow[1], Xiangdong Li[1], Adithya Murali[1], Christine Miller[1]

1. Agilent Technologies, Santa Clara, CA, United States

While shotgun lipidomics has advanced the field of lipid analysis, it has limitations including ionization suppression and failure to distinguish isobaric species of possible biological importance. This has led to chromatographic-based lipid profiling approaches using HPLC coupled to high resolution mass spectrometry. For best quantitative results and reproducibility, lipid profiling is done in MS1 mode, however confident lipid annotation requires MS/MS data acquisition to enable product-ion spectral matching against *in silico* generated databases. While chromatography helps elucidate isomeric lipid species and reduce complexity, it is still not possible to acquire all the MS/MS of interest in a single analysis for complex samples. In this study, a new, fully-automated Q-TOF iterative acquisition mode was used where precursors previously selected for MS/MS fragmentation are excluded on a rolling basis over multiple injections.

Another challenge in lipid profiling is the annotation of lipid MS/MS spectra. Due to the lack of authentic standards and the large diversity of lipid species, MS/MS annotation is done using *in silico* matching. In this study, a novel software tool was employed which uses a Bayesian-probability-theory algorithm and a theoretical lipid library (LipidBlast) to annotate the iterative mode MS/MS spectra. The tool takes special care not to over-annotate lipid entities by only providing the level of structural information confidently informed by the MS/MS spectra. The tool quickly and automatically generates an accurate mass-retention time database, and the resulting database annotates the MS-only lipid profiling data. This workflow provides more comprehensive lipid annotation than can be achieved by traditional approaches.

93

Developing an automated plasma sample preparation method for LC/MS analysis of metabolites

Weston Spivia[1], Koen Raedschelders[1], Manuel Gomez[1], Mark Sartain[2], Steve Fischer[2], Christine Miller[2], Jennifer Van Eyck[1]

1. Cedars-Sinai Medical Center, Los Angeles, California, USA

2. Agilent Technologies, Santa Clara, CA, United States

Translational research often involves the analysis of large numbers of plasma samples which are inherently complex and require significant sample preparation for metabolomics analysis. Manual sample preparation for metabolomics analysis involves quenching (protein precipitation) and lipid removal (liquid-liquid extraction) which is both time-consuming and inherently error-prone. Biological variation in plasma samples impacts the levels and composition of both proteins and lipids which can further complicate sample preparation and analysis. We have developed a semi-automated sample preparation workflow for metabolites in plasma to facilitate translational studies. As part of the development, different approaches to protein precipitation were tested to assess both quenching effectiveness and sample stability. For lipid removal, classic liquid-liquid extraction is known to be ineffective in completely removing some classes of lipid which can cause ionization suppression in subsequent LC/MS analysis. In this study, a sorbent-based approach was found to be more effective in removing the majority of the lipids thus improving the analysis results. The sorbent combines size exclusion and hydrophobic interactions between the sorbent and the long aliphatic chain of lipids to efficiently remove lipids without unwanted loss of metabolites. After optimization of quenching and lipid removal, the protocol was automated on the Bravo workstation and reproducibility was assessed using a targeted metabolomics approach on a triple quadrupole LC/MS system.

94

Chemotyping of medicinal cannabis phytocannabinoids

Aaron Elkins[1], Simone Rochfort[1], Delphine Vincent[1], Vilnis Ezernieks[1], German Spangenberg[1]

1. Agriculture Research Victoria, AgriBio, Centre for AgriBiosciene, Bundoora, Vic, Australia

Cannabis is an herbaceous flowering plant of the Cannabis genus (Rosale) that has been used for thousands of years to treat a wide range of medical conditions. Despite the illicit status of cannabis, the social push for the therapeutic use of cannabis extracts has steadily increased over recent years resulting in legislation legalising the cultivation of cannabis for medicinal use. This has opened the field of medicinal cannabis research. The biochemistry of cannabis is complex including phytocannabinoids, terpenes and phenolics. Each of these classes have biologically active compounds that contribute to the medical efficacy of cannabis. The most abundant of these classes are the phytocannabinoids derived from the glandular trichomes of the flowering heads of the female plant. We have undertaken the targeted and untargeted chemotypic analysis of 70 diverse cannabis strains using liquid chromatography mass spectrometry (LCMS) to characterise the major and minor cannabinoids. Multivariate hierarchical clustering has been employed to assess similarities and differences between the strains. This provides an opportunity to target strains with specific chemotypic profiles to be used in the treatment of specific medical conditions.

95

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[3, 4], Pamela Soh[1], Zeynep Sumer-Bayraktar[1, 2], Paula Niewold[5, 2], Ashleigh Dale[1, 2], Stuart J. Cordwell[1, 5, 2, 6]

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

4. Centre for Blood Research, University of British Columbia, Vancouver, Canada

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

6. 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 assaccharolytic 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 adherence and 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.

96

Early detection of honey bee malnutrition and colony collapse

Clara Castaños Sánchez de la Barquera[1], Julia Grassl[1], Mary Boyce[2], Harvey Millar[1]

1. School of Molecular Science , The University of Western Australia, Perth, WA, Australia

2. Centre for Integrative Metabolomics and Computational Biology, Edith Cowan University, Perth, WA, Australia

The nutritional status of managed honey bees depends on the attention that beekeepers provide them and their hive management skills. One major issue beekeepers have to deal with, is the early detection of what is termed the "skinny bee syndrome", where bees suffer nutritionally and become so unwell that they cannot recover easily. At this stage, the colony is unable to continue foraging and support their brood, commonly by the time this is detected it is too late for beekeeper intervention. It is essential for beekeepers to maintain their bee colonies in a healthy nutritional condition for the optimal production of honey and pollination services. The objective of this project is to identify the time point when honey bees are on a nutritional decline and determine the best nutritional management strategies to minimize the risk of productivity loss. Gas Chromatography and Liquid Chromatography-Mass Spectrometry (GC-MS or LC-MS) analyses have become effective tools in biochemistry and biology to study the metabolism of insect systems and their regulation in disease, genetic and environmental variations. Therefore, GC-MS and LC-MS, metabolomic profiles of the bees at different time points of nutritional starvation and supplemental feeding will be compared to evaluate changes in the bee's physiology. We are measuring variations in amino acids, lipids, carbohydrates and proteins expression, that will be correlated with hive condition measurements; such as brood quantity (counting brood cells through photos of the frames), bee noise, hive weight and humidity.

97

Improved metabolome coverage and increased confidence in unknown identification through novel automated acquisition strategy combining sequential injections and MS^n

Amanda Souza[1], Ioanna Ntai[1], Iman Mohtashemi[1], Jenny Berryhill[1], Ralf Tautenhahn[1], Graeme McAlister[1], Derek Bailey[1], Linda Lin[1], Ryo Komatsuzaki[1], Caroline Ding[1], Seema Sharma[1], Tim Stratton[1], Vlad Zabrouskov[1], Andreas Huhmer[1]

1. Thermo Fisher Scientific, San Jose, California, United States

Introduction

Compound identification is a bottleneck in untargeted metabolomics, hindering biological interpretation of results. Here, we describe a data-informed workflow that maximizes the number of metabolites interrogated by MS/MS and MS^n, while minimizing the acquisition of uninformative spectra. Human plasma was analyzed using this workflow resulting in high confidence identifications, deeper metabolome coverage, and enhanced biological knowledge generation.

Methods

Human plasma was extracted with methanol and injected on a Thermo Scientific™ Hypersil GOLD™ column. Instrumentation included a Thermo Scientific™ Vanquish™ UHPLC system and a Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer. Data were analyzed using Thermo Scientific™ Mass Frontier software and Thermo Scientific™ Compound Discoverer™ software.

Results

During data-dependent MS/MS, ions are selected based on abundance, without any knowledge of biological relevance or ion type. In a typical DDA experiment, we determined, that >40% of MS/MS spectra could be attributed to background ions. By enabling the automatic generation and implementation of an exclusion list based on real-time feature detection in LC-MS data, background ion MS^2 spectra were practically eliminated (<0.1%), allowing for the analysis of true sample components.

Small molecules form different adducts and cluster ions during electrospray ionization. Highly abundant compounds may prevent the fragmentation of metabolites of lower abundance. By populating the inclusion list with the preferred ion for each metabolite, more compounds can be sampled by MS/MS and MS^n in a single run. Additionally, by automatically updating inter-run inclusion and exclusion lists, we can ensure that compounds not selected for MS/MS and MS^n will be prioritized during a subsequent injection.

Conclusions

The combination of MS^n and automatically generated inter-run inclusion and exclusion lists resulted in fragmentation of more unique metabolites and a greater number of metabolites confidently annotated. This innovative workflow addresses the identification bottleneck in untargeted metabolomics and enables confident biological interpretation of the results.

98

Flavonoid annotation using a product ion-dependent MS^n data acquisition method on a Tribrid Orbitrap mass spectrometer

Amanda Souza[1], Reiko Kiyonami[1], Iwao Sakane[2], Seema Sharma[1], Graeme McAlister[1], Caroline Ding[1], Andreas Huhmer[1]

1. Thermo Fisher Scientific, San Jose, California, United States

2. ITO EN, LTD, Tokyo, Japan

Introduction

Flavonoid annotation from natural products remains challenging because of their structural diversity and lack of authentic standards. Simple MS2 based analyses are often not sufficient for complete structural annotation of flavonoids. We present a new flavonoid annotation workflow that uses comprehensive fragment ion information from HCD MS-MS and higher order $FTMS^n$ for rapid flavonoid annotation.

Methods

Flavonoid extracts from different types of natural products were analyzed using LC-MS. A Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer was used for collecting HRAM MS and MS^n data. The data were processed using Compound Discoverer™ 3.0 software. A novel structure ranking algorithm, mzLogic, was applied to the MS and MS^n data for confident structure elucidation of the unknown flavonoids.

Results

The flavonoid extracts from multiple tea and fruit/vegetable juice samples were analyzed using the developed flavonoid annotation workflow. The MS^n (up to 5) scans were only triggered when one of the sugar neutral loss product ions (162.0523, 180.0628, 146.0574, 164.0685, 176.0315, and 194.0421) was detected in the MS/MS scan. The collected data were processed using Mass Frontier 8.0 and Compound Discoverer 3.0 software. A novel structure ranking algorithm included in the Compound Discoverer 3.0 software was applied to the MS-MS and MS^n data for confident structure characterization of the unknown flavonoids based on ChemSpider database and custom flavonoids database. The MS^n data were critical, especially for the annotation of flavonoid glycoconjugates and flavonoid isomers.

Conclusion

The new LC-MS^n workflow allows intelligent MS/MS fragment ion dependent MSn data collection. In combination with dedicated software for data processing, the new workflow enables improved throughput, identification coverage and confidence for flavonoids annotation from natural products by using comprehensive fragment ion information from HCD MS-MS and higher order $FTMS^n$.

99

Bespoke deuterated lipid molecules for structure function applications from the National Deuteration Facility

Karyn L Wilde[1], Nageshwar R Yepuri[1], James K Howard[1], Anwen M Krause-Heuer[1], Peter J Holden[1], Tamim A Darwish[1]

1. National Deuteration Facility - ANSTO, Lucas Heights, NSW, Australia

Deuterium (^{2}H or D) is a naturally occurring stable isotope of hydrogen (^{1}H or H) which contains a neutron in addition to the proton and electron found in the naturally abundant protium ^{1}H. Molecular deuteration of molecules significantly increases the options in structure function investigations using MS, NMR, neutron scattering and other techniques but there have been limited global initiatives in the production and thus availability of complex deuterated small molecules such as lipids and fatty acids for such experiments.

Utilising catalysed ^{1}H/^{2}H exchange followed by custom chemical synthesis, the Chemical Deuteration laboratories of the National Deuteration Facility (NDF) of the Australian Nuclear Science and Technology Organisation are able to produce and supply relevant and bespoke deuterated compounds unavailable from commercial sources. We will describe the synthesis and application of new deuterated molecules with the NDF now able to produce deuterated cholesterol, head and/or tail deuterated lipids or selectively deuterated triglycerides,[1] including phospholipids based on deuterated oleyl chains with a range of head groups. These include perdeuterated 1,2-oleoyl-*sn*-glycero-3-phosphocholine (POPC)[2], selectively deuterated POPC, branched chain (phytanic) phospholipids 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhyPC)[3], perdeuterated phytantriol[4] and perdeuterated mono and di-oleoyl glycerol and phytanoyl monoethanolamide.[5] These lipids have been extensively used in constructing biologically more relevant lipidic matrices for investigations using neutron, NMR, IR and MS techniques.

The National Deuteration Facility is partly supported by the National Collaborative Research Infrastructure Strategy – an initiative of the Australian Government.

1. Salentinig, S.; Yepuri, N. R.; Hawley, A.; Boyd, B. J.; Gilbert, E.; Darwish, T. A. Chemistry and physics of lipids 2015, 190, 43.
2. Yepuri, N. R.; Darwish, T. A.; Krause-Heuer, A. M.; Leung, A. E.; Delhom, R.; Wacklin, H. P.; Holden, P. J. ChemPlusChem 2016, 81, 315.
3. Yepuri, N. R.; Holt, S. A.; Moraes, G.; Holden, P. J.; Hossain, K. R.; Valenzuela, S. M.; James, M.; Darwish, T. A. Chem. Phys. Lipids 2014, 183, 22.
4. Yepuri, N. R.; Clulow, A. J.; Prentice, R. N.; Gilbert, E. P.; Hawley, A.; Rizwan, S. B.; Boyd, B. J.; Darwish, T. A. Journal of colloid and interface science 2018.
5. van 't Hag, L.; de Campo, L.; Garvey, C. J.; Feast, G. C.; Leung, A. E.; Yepuri, N. R.; Knott, R.; Greaves, T. L.; Tran, N.; Gras, S. L.; Drummond, C. J.; Conn, C. E. J. Phys. Chem. Lett. 2016, 7, 2862.

100

A high throughput single platform for *quantitative multiOmics*

Billy Molloy[1], Sarah Lennon[1], Lee Gethings[1], Robert Plumb[2]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Milford, MA, USA

A high throughput targeted UPLC-MS/MS single platform, employing a reversed-phase gradient separation, has been developed for the quantification/monitoring of small molecule metabolites, lipids and peptides. The platform employs a single LC column and mobile phase combination which allows the analysis of multiple analyte classes with either positive or negative ion MRM detection. The use of metabolic profiling (metabonomics/metabolomics) to discover biomarkers of organismal response to environmental and physiological change is now widespread. In biomedical applications metabolic profiling is being deployed as a method for finding novel, mechanistic, biomarkers of disease with obvious potential for improving diagnosis, and patient stratification. Hypothesis driven metabolomics delivers detailed qualitative and quantitative analysis on specific pathways or classes of metabolites, allowing researchers to analyse the effects of disease or treatments in greater detail. These targeted assays usually employ “bespoke” methods which are optimized for each pathway or metabolic class making multiplexing assays difficult. We have developed a single analytical LC-MS/MS platform which is rapid, simple and reliable. The methodology employs a single LC column / mobile-phase combination which facilitate the bile acids, biogenic amine, free fatty acids, acyl carnitines, lipids and 100 protein panel. This single platform approach has been employed for the analysis of plasma from a lung and bladder cancer study, showing excellent throughput and sensitivity.

101

High throughput lipidomics using ion mobility enabled rapid LC-MS profiling shows promise for the analysis of human plasma samples obtained from breast cancer patients

Adam King[1], Robert Trengove[2], Lauren Mullin[3], Paul Rainville[3], Giorgis Isaac[3], Robert Plumb[3], Lee Gethings[1], Ian Wilson[4]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Murdoch University, Perth, WA, Australia

3. Waters Corporation, Milford, MA, USA

4. Imperial College, London, UK

The study of lipids can be traced back to the 19th century but, with recent advancements in analytical chemistry and biomolecular research the importance of lipids in biochemical processes and function continues to expand. This is especially true in the area of human disease progression where lipids are shown to provide important information on disease inception and progression. MS based lipidomics can be subdivided into qualitative and quantitative analysis. The quantitative analysis of lipids typically employs more rigorously developed methodology and the use of authentic standards to provide concentration information on specific lipid species. majority of research in this area has been focused on developing improved high resolution, information rich and comprehensive methods, which typically results in analysis times in the 15-30 minute time scale. However, these timescales mean that the application of LC-MS-based lipid phenotyping studies on large epidemiological cohorts and human sample biobanks can be limited by resource constraints. A sub 4-minute microbore LC-ion mobility-accurate mass MS (LC-IMS-MS) method has been developed for the rapid, profiling of lipids in biological fluids and tissues. The method was scaled directly from a conventional, 12 min, LC-MS analysis maintaining the chromatographic performance and lipid class separation observed in the longer methodology. A 75% saving in mobile phase consumption and analysis time was achieved by employing this microbore high throughput methodology, with the capability of detecting in-excess of 5,500 lipid features from a human plasma sample. The incorporation of ion mobility into the LC-MS analysis resulted in the generation of superior quality MS and MS/MS spectral data as well as enhanced resolution in IMS dimension based on lipid class. The rapid methodology was applied to the analysis of a pilot set of commercially sourced breast cancer plasma samples. The assay was capable of differentiating healthy control samples from diseased using multivariate statistical analysis based on their lipid phenotypes. The data showed phosphatidylcholines, triglycerides, diglycerides exhibited lower expression and phosphatidylserine showed increased expression in the breast cancer samples compared to healthy subjects.

102

Peptides and Penicillins: Identifying the immunogenic β-lactam ligand(s)

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

Penicillin hypersensitivity is recognised as a major public health concern with up to 11.5% of recipients reporting adverse reactions. Traditionally, penicillin allergies have been considered to be primarily IgE-mediated, however, there is emerging evidence for a key role for T cell-mediated reactions. Possessing the ability to form covalent bonds (known as haptenation) with serum proteins, penicillins create neo-antigens that have the potential to be immunogenic. Upon processing and presentation by antigen-presenting cells, these haptenated neo-antigens may activate drug-responsive T cells, leading to a range of cutaneous adverse drug reactions from mild rashes to more severe pathologies involving organ damage such as drug-induced liver injury. A group of 10 patients were recruited from The Alfred Hospital Allergy Clinic based on their history of adverse drug reaction to penicillins. Of these patients, only 40% tested positive for serum IgE against the causative drug, suggesting that other

factors/cellular populations (i.e. T cells) contribute to the adverse reaction. Here we focus on one hypersensitive patient, who exhibited drug-responsive T cells activated via the human leukocyte antigen (HLA) class-I molecule HLA-A*02:01. We hypothesised that this common Caucasian HLA allotype presented a haptenated neo-antigen to T cells responsible for the ensuing drug reaction. Using an immunopurification LC-MS/MS workflow, we explored the HLA-A*02:01 immunopeptidome isolated from patient-derived B-cells treated with or without penicillin. Utilising haptenated human serum albumin, we defined a set of robust criteria for identification of penicillin-haptenated peptides, which when applied to this data set showed evidence for three HLA-A*02:01 bound penicillin-haptenated peptides. Together, these results set the basis for further characterisation of haptenated peptides in other HLA-A*02:01^{+} penicillin hypersensitive patients and will provide novel molecular insights into this common drug reaction.

103

Using genomics and proteomics to understand the antibiotic resistance capabilities of an isolate

Stephanie E L Town[1], Matt P Padula[1], Steven P Djordjevic[2]

1. School of Life Sciences, University of Technology Sydney, Broadway, NSW, Australia

2. Infection, Immunity and Innovation (i3) Institute, University of Technology Sydney, Broadway, NSW, Australia

Background: Recently, next-generation genomic sequencing has shown potential as an accurate predictor of phenotype and for understanding genes conferring antibiotic resistance (AMR) in bacterial communities. In the clinic, antibiotic disc assays can provide a visual indication of the type and strength of AMR in an isolate, but do not observe the phenotype on a molecular level. Aim & Objectives: Our research aims to understand the AMR capabilities of an isolate on a molecular level through examination of its genome and proteome, with and without antibiotic challenge. This research seeks to connect AMR genes to gene end-products on the proteoform-level (peptides, proteins, post-translational modifications), which has yet to be experimentally shown. Methods: Long- and short-read genomic sequencing and assembly was conducted in-house on multi-drug resistant *E. coli* isolates to ascertain the presence of AMR-related genes and antibiotic disc assays were performed to visualise phenotypic effectiveness for related-AMR. To correlate phenotype with genotype on a molecular level, a shotgun proteomics pipeline using LC-MS/MS was used to generate data on proteome changes with and without antibiotic challenge. PEAKS Studio was used to analyse this data, and UniProt and STRING databases were used to assess the significance of reported protein and protein interactions implicated in AMR. Results: Several interesting proteins related to AMR were upregulated, despite no antibiotic challenge, including proteins previously annotated as hypothetical and multi-drug resistance proteins. Conclusion: This research establishes that proteomic techniques validate the potential for genomic sequencing to replace current clinical tests and provide more specificity in selecting the right antibiotic in the clinic. Our research findings are significant for being the first to experimentally link an AMR-related gene to the AMR-related protein using a systems biology approach.

104

Extractomes profiling reveals potential role of key proteins in *Staphylococcus aureus* biofilm using TMT-based quantitative mass spectrometry

Md Arifur Rahman[1], Ardeshir Amirkhani[2], Durdana Chowdhury[1], Mark Molloy[2], Dana Pascovici[2], Maria Mempin[1], Xiaomin Song[2], Mark Baker[1], Honghua Hu[1], Karen Vickery[1]

1. Department of Biomedical Sciences, Macquarie University, Sydney, NSW, Australia

2. Australian Proteome Analysis Facility, Sydney, NSW, Australia

Staphylococcus aureus and coagulase-negative staphylococci comprises approximately 65% of infections associated with medical devices and are well known for their biofilm formatting ability. Currently, there is no efficient method for early biofilm detection. Therefore, we aimed to construct a comprehensive reference map followed by identifying marker proteins in the different growth phase of hydrated biofilm, and then perform pathway analysis, subcellular localisation and protein-protein interaction network mapping. S. *aureus* reference strain (ATCC 25923) was grown in tryptic soy broth to produce a 24-hour planktonic, 3-day wet biofilm (3dwb), and 12-day wet biofilm (12dwb). The Centres for Disease Control biofilm reactor was used to grow biofilms. Tandem Mass Tag (TMT)-based mass spectrometry was performed, and protein identification and relative quantitation of protein levels were performed using Proteome Discoverer (version 1.3). Statistical analysis was done using the TMTPrePro R package. We identified 1636 total biofilm extractomes. Among the significantly (>2-fold) up-regulated proteins in 3dwb, hyaluronidase encoded by hysA an extracellular enzyme involved in dispersing established biofilms by degradation of hyaluronic acid. Proteins significantly up-regulated in 12dwb are mostly involved in energy metabolism and cell wall formation. Among the significantly down-regulated proteins, chitinase encoded by SA0914 an extracellular enzyme involved in quorum sensing and responsible for preventing initial attachment of biofilm formation. Interestingly, chitinase can hydrolyse N-acetyl-D-glucosamine which is the structural component of hyaluronic acid. Therefore, hysA in combination with chitinase may play a potential role in the eradication and/or prevention of biofilm formation. In addition, protein-protein interaction network showed significantly more interactions for 3dwb and 12dwb. Current study showed a significant range of quantitative proteomic shifts and changes in metabolic process in biofilm. HysA and chitinase may be potential targets in the biofilm research *in vitro* and *in vivo*. The proteins identified might be helpful in designing advanced targeted candidates for vaccines, anti-biofilm agents, diagnostic biomarkers for *S. aureus* biofilm-related infections associated with implantable medical devices.

105

Accessing the mannose-6-phosphate glycoproteome using titanium dioxide (TiO_2) enrichment

Sayantani Chatterjee[1], Rebeca K. Sakuma[1, 2], Morten Thaysen-Andersen[1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Department of Parasitology, Institute of Biomedical Sciences, University of Sau Paulo, Sau Paulo, Brazil

Lysosomal hydrolases receive mannose-6-phosphate (M6P) modifications while trafficking the ER-Golgi pathway, an *N*-glycosylation tagging process that directs enzymes to their lysosomal destination. Previously unknown extracellular roles of M6P-glycoproteins were recently suggested including the involvement in cell-to-cell communication and cell growth in inflammation and cancer. However, solid structural data is unavailable to determine the cellular expression of the "M6P-glycoproteome" and therefore the extracellular role of M6P-glycoproteins. Comprehensive characterisation of the M6P-glycoproteome is challenging owing to their low abundance and anionic character, which, in concert, limit their MS-detection using conventional analytical strategies. We have developed a novel method that enables isolation and thereby enhanced detection of both free reduced *N*-glycans and *N*-glycopeptides utilising titanium dioxide (TiO_2)-SPE based enrichment followed by LC-MS/MS detection. Optimisation of the mobile phases (loading and elution conditions) was performed using isolated M6P-glycoproteins (myeloperoxidase and SMPDL3A). Firstly, it was demonstrated that M6P-glycans and glycopeptides quantitatively survive the short exposure to the acidic and alkaline conditions of the mobile phases. Efficient depletion of neutral and, surprisingly, sialylated glycans and glycopeptides were shown demonstrating preferential affinity of the M6P-glycoconjugates to TiO_2 under the optimised loading condition. Phosphopeptides were efficiently depleted from the peptide samples using HILIC-SPE prior to the TiO_2 enrichment to avoid analyte interference. Finally, the enrichment method was tested on a complex biological protein mixture derived from prostate cancer tissues and different neutrophil granules. Identification of a spectrum of M6P-glycans ($Phos_{1-2}Man_{3-8}GlcNAc_2$) and M6P-glycopeptides (containing $Phos_{1-2}Man_{3-8}GlcNAc_2$) originating from lysosome-like glycoproteins e.g. myeloperoxidase, cathepsin G, neutrophil elastase from neutrophil granules and extracellular glycoproteins including fibronectin, collagen and prostatic acid phosphatases from prostate cancer tissues demonstrated the potential of this enrichment method when applied to complex biological samples. In conclusion, this novel enrichment method offers parallel characterisation of M6P-glycans and M6P-glycopeptides enabling deep characterisation of the M6P-glycoproteome in a biologically relevant context.

106

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 (four 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 eight sites are novel. By purifying these proteins from single gene knockout yeast strains, we now seek to determine the upstream kinases responsible for the phosphorylation, and potential regulation of Set5, Dot1, Gis1 and Jhd2. This will facilitate the integration of these enzymes into intracellular signalling pathways, and thereby aid the assembly of the first regulatory network of histone methylation. Given the evolutionary conservation of this histone modification, the foundational insights gained through this work will be relevant to other eukaryotes and may prompt similar in-depth analyses in these organisms.

107

Characterisation of human adipose stem cells

Naomi Koh Belic[1], Fiona Bicknell[1], Jerran Santos[1], Bruce Milthorpe[1], Matt 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 lack of scientific evidence in regard to their safety let alone efficacy. These are unproven stem cell treatments being offered by predatory clinics, and there are more than 60 of these clinics 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. 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.

108

EasyPep - A new simplified and optimized workflow for MS sample preparation

Michael MM Mariani[1], Amarjeet Flora[2], Sergei Snovida[2], Ryan Bomgarden[2]

1. ThermoFisher Scientific, Melbourne, VIC, Australia

2. Thermo Fisher Scientific, Rockford, IL, USA

Advances in mass spectrometry has 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, digestion efficiency and reproducibility. Here, we describe a simplified sample prep kit containing pre-formulated reagents and standardized protocol that can be used to efficiently process 10µg to 100µg protein samples in less than 2 hours. In this study, we evaluated the scalability, compatibility, and reproducibility of this sample preparation kit compared to published methods. Protein extraction from cells and tissues was evaluated using a cell lysis buffer and universal nuclease for nucleic acid disruption compared to standard sonication methods. A rapid (<10min.) reduction/alkylation solution was developed as well as a combined trypsin/LysC protease mixture for protein digestion. A mixed mode peptide clean-up procedure using a novel spin column format was used for detergent removal. Peptides were quantified and normalized using the Pierce™ Quantitative Colorimetric Peptide Assay prior to analysis using a Thermo Scientific™ Q Exactive™ HF MS.
Our new standardized workflow yielded 10-20% higher number of peptides and proteins with lower missed cleavages (<90%) compared to other commercial kits or homebrew methods. We demonstrate that our optimized protocol reduces hands on time to less than 30 minutes with total sample processing time from intact cells to cleaned-up peptides under 1.5 hours. Finally, we show this procedure is compatible with isobaric labeling reagents such as TMT and LFQ methods to reproducibly quantify protein abundances. The protocol has been successfully tested with several cell lines (HeLa, CHO, A549, 293), purified proteins and various mouse tissues (heart, liver, lung).
Overall, our kit provides superior method in terms of time saved, peptide/protein identification rates, and reproducibility compared to previous proteomic methods and greatly simplifies proteomic sample preparation for protein identification and quantitation.

109

Dietary selenium significantly effects rat liver protein expression

Larissa Lago[1], Eugene Kapp[1], Roger Sude[2], Blaine Roberts[1]

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

2. Nutritional Sciences , University of Wisconsin-Madison, Madison, Wisconsin, USA

Selenium is an essential trace element incorporated as selenocysteine (Sec) in selenoproteins. Selenium has enjoyed increased attention in studies due to it playing an essential role in a variety of cellular processes. It is involved in protection against excess H_2O_2, heavy metal detoxification, and regulation of both the immune and reproductive systems. Selenium has been found to have been protective in a range of cancers, but also in many neurodegenerative disorders, including both Alzheimer's and Parkinson's disease, particularly when populations are deficient. On the other side having too much Se can be toxic and there are documented cases of acute exposure through poor manufacturing and quality control of supplements and through environmental exposure due to selenium-rich geological deposits. In this study, we investigated the effect of dietary selenium deficiency and supernutritial doses and how effects incorporation of SeCys into different selenoproteins in rat liver. We additionally used shotgun proteomics and identified nonselenoproteins regulated by dietary changes in Se. Our results confirm that selenium treatment results in increased selenium incorporation. Selenium-treated animals also showed a significant alteration of the expression of 30 separate proteins. We found that an increase in dietary selenium caused an upregulation of essential proteins such as glutathione peroxidase (Gpx1, an essential protein and the major selenoprotein in most animal tissue). There also changes in non-selenium containing proteins linked to energy production, lipid metabolism or protein synthesis pathways. This indicates the broad effect on cellular metabolism that can be perpetuated through changes in selenium exposure or deficiencies and begins to explain how this essential element can be involved in a wide variety of diseases.

110

Development of an immune response assay for honey bees

Shannon Holt[2, 1, 3], Harvey Millar[3], Julia Grassl[2, 1, 3]

1. Honey Bee Health Research Group, School of Molecular Sciences, The University of Western Australia, Perth

2. Cooperative Research Centre, Honey Bee Products, Yanchep, WA, Australia

3. ARC Centre of Excellence in Plant Energy Biology, School of Molecular Sciences, The University of Western Australia, Perth, WA, Australia

Honey bee (*Apis mellifera*) health is under increasing threat, with population declines observed in both managed and wild bees. As demands for pollination services are increasing, understanding and protecting bee health is essential. Honey bees are susceptible to a large range of parasites and pathogens, and infections can significantly affect colony health and productivity. As a result, weak colonies can become increasingly vulnerable to additional infections. Current practice around the world often involves treating colonies with pesticides in order to kill pathogens. This is harmful to the bees and is also an ineffective method for long term management. A more sustainable approach is to selectively breed for disease resistant traits. This project aims to develop a panel of protein markers which could be used to identify immune competent bees. Recent research identified an array of immune response proteins in male honey bees. These were found to change in abundance when male bees were infected with the pathogen, *Nosema apis*. Using targeted protein quantitation through MRM mass spectrometry analysis, we will develop

an assay to quantitate proteins from the four immune pathways: Toll, immune deficiency (IMD), c-Jun N-terminal kinases signalling (JNK), and Janus kinase/signal transducers and activators of transcription (JAK/STAT). We will then test the markers in other diseases; both present and absent from Australia. Once confirmed, the panel of markers could be used to investigate immune response, providing a platform for marker-directed breeding of more immune competent bees.

111

Potential protein biomarkers for the diagnosis of early stage 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, 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 as 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. Rahman et al. World J Gastroenterol, 20 (2014) 4483–4490
2. Digklia and Wagner. World J Gastroenterol, 22 (2016) 2403-2414
3. Humphries et al. Biochimica et Biophysica Acta, 1844 (2014) 1051-1058

112

A mass spectrometry quality control website: monitoring SCIEX instruments to maximise productivity

Katherine E Scull[1], Jerico Revote[2], Sri H Ramarathinam[3], Nathan Croft[3], Anthony W Purcell[3]

1. Biochemistry & Molecular Biology, Monash University, Clayton, Victoria, Australia

2. Monash eResearch Centre, Monash University, Clayton, Victoria, Australia

3. Biochemistry & Molecular Biology, Monash University, Clayton, Victoria, Australia

To maximise productivity and the amount of information proteomics laboratories can derive from precious samples, it is essential to keep mass spectrometers working at peak efficiency. Ideally, laboratories need a user-friendly method of monitoring the instruments' performance over time to allow early detection of any problems arising, thereby minimising instrument down-time and preventing samples being wasted in sub-optimal runs.

We have designed a website which accesses stored data from routine quality control (QC) experiments on our SCIEX 5600 and 6600 TripleTOF mass spectrometers. This is hosted on R@CMon, the Monash node of the Nectar research cloud, and is accessible from Monash University client computers. The data comes from standard SCIEX calibration runs, which run automatically several times a day, as well as complex biological standard experiments representative of our lab's typical immunopeptidomic samples, which are run manually several times a week and searched using ProteinPilot. Several times daily, the QC server accesses the resulting data files via a network shared drive and extracts the relevant data, along with the date and time, for storage in MySQL databases using a program written in C. The QC server utilises Apache tomcat to run Java servlets, which request information regarding specified dates from the MySQL databases, and present useful graphs with Plotly.js. In this way, it updates itself on-the-fly, allowing the user to see at a glance whether recent runs succeeded or failed, and notice subtle deterioration or fluctuations in signal, ideally in time to pinpoint and address issues before they develop into major problems.

While this QC server is tailored to our lab's SCIEX instruments and specific needs, the individual components may be adapted to other mass spectrometers and other sample-specific readouts. Thus, it may be of interest to many groups seeking tools for monitoring their mass spectrometers for proactive quality control.

113

Desorption electrospray ionization and data independent analysis profiling of the lipid complement of lysosomal storage disorders

Emmanuelle Claude[1], Lee Gethings[1], Johannes Vissers[1], Mina Mirzaian[2], Johannes Aerts[2], David Heywood[1]

1. Waters Corporation, Wilmslow, United Kingdom

2. School of Biosciences , University of Birmingham, Birmingham, United Kingdom

Gaucher disease (GD) is a lysosomal storage disorder that is caused by a deficiency in the lysosomal beta-glucosidase. A hallmark of GD is the massive accumulation of glucosylceramide (GlcCer) in lysosomes of tissue macrophages transforming into viable Gaucher cells. However, tissue macrophages do not manage to prevent GlcCer accumulation by conversion to GlcSph. Elucidating the chemical composition and distribution of GlcCer in primary storage cells, especially the fatty acyl composition and its glycosylation supplement, is therefore of importance to test the hypothesis that modified GlcCer is not a suitable substrate and subsequently accumulates.

Gaucher and control spleen tissue samples were either directly analyzed using DESI or total and Bligh and Dyer extracts by LC-MS on the same oa-ToF MS platform (Xevo G2-XS QTof). Frozen tissue slices were directly mounted on a modified Prosolia 2D stage and analyzed by DESI (without any additional sample pre-treatment). The LC-MS experiments were conducted using reversed phase based chromatographic methods. A typical workflow comprised of the relative quantification of the DIA LC-MS data using Progenesis QI and development informatics, providing candidate *m/z* target lists for the tissue lipid distribution analysis using High Definition Imaging and DriftScope software.

The multi-dimensional processed LC-MS data were submitted for untargeted searches and the detected features that demonstrated statistical relevance based on OPLS-DA were selected for compound database searching using various lipid discovery search tools. Fatty acids and glucosylceramides were found to be significantly over expressed, typically >10 fold, in GD patient samples. Despite vast morphological differences, localization/distribution and abundance differences of fatty acyls and glycosylated lipid species between the healthy control and Gaucher spleen tissue samples were readily noted. The analysis of the spleen tissue of GD patients and healthy control spleen tissue by DESI MSI and DIA LC-MS provided complementary information. The initial results suggest that the applied methodologies provide novel insights in lysosomal storage disorders, GD in particular, by revealing metabolite abnormalities, which are currently undergoing validation.

114

Optimised Desorption Electrospray Ionisation Mass Spectrometry Imaging (DESI-MSI) method for the analysis of proteins/peptides directly from tissue sections

Emmanuelle Claude[1], Mark Towers[1], James Hughes[2], Patricia Lalor[2], Helen Cooper[2], David Heywood[1]

1. Waters Corporation, Wilmslow, United Kingdom

2. School of Biosciences , University of Birmingham, Birmingham, United Kingdom

DESI-MSI is typically known for the mapping of small molecules such as lipids, directly from tissue sections. Whilst the detection of spotted protein standards from target plates has previously been shown, the detection of large biomolecules directly from tissue sections has presented difficulty. Here we describe a newly developed method combining a series of optimized parameters and conditions which allow the extraction of large biomolecules from the tissue in droplet form. Detection of these multiply charged ions is enhanced using ion mobility which enables separation from the intense endogenous ion species and chemical background aiding in visualisation.

All experiments were carried out on a SYNAPT HDMS G2-S*i* Q-ToF (Waters) with a Prosolia2D DESI stage. A modified spray head assembly and an inlet capillary with a heated sheath connected to an adjustable power supply were used; voltages between 0 and 12V were applied to generate a variable heating effect.

To remove lipids and salts, the tissues were subjected to a series of ethanol based washes followed by chloroform. The heated inlet capillary was particularly beneficial giving an enhanced ionisation of proteins and peptides from the tissue with further optimisation in terms of sprayer inlet geometry, solvent composition, nebulising gas pressure and solvent flow rate. The orthogonal separation of ion mobility was found to be essential to observe the most abundant charge series in the highly complex data set. The most abundant charge series observed were those belonging to Haemoglobin with charge states from 9+ to 21+ observed. Extraction of the trendlines via the Driftscope software and subsequent use of MaxEnt 1 algorithm allowed a putative assignment of four Haemoglobin subunits from rat liver tissue. Additional charge series putatively assigned to FABP , HSP10 and COX8A for the human liver tissue section and MBP for the brain tissue section were observed. Overall further trend lines were found relating to background solvent peaks, residual lipids and more importantly lower abundant small proteins / endogenous peptides with charge states from +2 to +8.

115

Top-down, middle-down and bottom-up proteomics analysis of medicinal cannabis

Delphine Vincent[1], Simone Rochfort[1], German Spangenberg[1]

1. Agriculture Victoria Research, AgriBio, Centre for AgriBioscience, Bundoora, VIC, Australia

Medicinal cannabis is used to relieve the symptoms of certain medical conditions, such as epilepsy. Cannabis is a controlled substance and until recently was illegal in many jurisdictions. Consequently, the study of this plant has been restricted. Proteomics studies on *Cannabis sativa* reported so far have been based primarily based on plant organs and tissues other than buds, such as roots, hypocotyl, hempseeds, and flour. Therefore, we set out to develop mass spectrometry-based proteomics methods to analysis the reproductive organs of medicinal cannabis.

In this study, we first optimised protein extraction from medicinal cannabis apical buds and trichomes exploring six methods for top-down analyses by UPLC-MS. We then optimised protein digestion on the highest yielding protein extract by testing four

endoproteases independently or combined for middle-down and bottom-up analysis by nLC-MS/MS. The most efficient methods will be applied to the comparison of different medicinal cannabis strains.

116

Identifying PTM changes from DIA data – a computational workflow

Jemma Wu[1], Conor McCafferty[2], Thiri Zaw[1], Vera Ignjatovic[2, 3], Xiaomin Song[1], Mark Molloy[4], Dana Pascovici[1]

1. Australian Proteome Analysis Facility, Macquarie University, Sydney, NSW, Australia

2. Hematology Research Laboratory, Murdoch Childrens Research Institute, Melbourne, VIC, Australia

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

4. Kolling Institute of Medical Research, The University of Sydney, St. Leonards, NSW , Australia

Post-translational modifications (PTMs) help regulate protein folding, activity, signalling and their interactions. PTMs are crucial to our understanding of biological functions and are particularly important for clinical research. With the fast improvement in high throughput MS technology such as DIA, large scale PTM studies have emerged, in which thousands of modifications can be found from a single DIA/SWATH experiment. Discovering reliable and biologically interesting PTMs has become a challenging problem, hence detailed and robust PTM workflows are now crucial to further our understanding of PTMs. Here we describe a computational workflow which aims to facilitate the discovery of reliable and relevant PTMs from DIA experiments.

The workflow starts from extracting the SWATH experiment data by using a peptide library containing the identified PTM peptides. The extracted Swath results can be optionally filtered by using FDR pass filtering criteria; then multivariate and differentially expression analysis will be performed on the normalised PTM data. Often hundreds of PTM peptides are found to be differentially expressed, but many of these changes are simply due to the underlying protein's expression changes. In order to exclude these baseline cases, the relative stoichiometry is calculated and applied as an additional filtering criterion for each PTM. The reliability of the results can be further improved by checking the modification residues against the known residues in public repositories such as SwissProt. We evaluated this workflow with a previously published human plasma study searched with ID focus allowing Biological Modifications using ProteinPilot V5.0. Over 75% of the PTM peptides and 50% of the PTM proteins were filtered out by using both the differentially expressed and relative stoichiometry over threshold criteria. This workflow can be easily extended to PTM studies from other types of MS techniques.

117

From raw data to biological insights: A computational pipeline for SWATH-MS-based proteomics

Srikanth S. Manda[1], Max Wittmann[1], Michael Dausmann[1], Michael Hecker[1], Brett Tully[1], Peter Hains[1], Phil Robinson[1], Roger Reddel[1], Qing Zhong[1]

1. Children's Medical Research Institute, University of Sydney, Westmead, New South Wales, Australia

SWATH-MS is a type of data-independent acquisition (DIA) method that combines the advantage of both shotgun and targeted proteomics, attaining deep proteome coverage, reproducibility, and accurate protein quantification. The analysis of SWATH-MS data involves generation of a spectral library that records prior knowledge about the mass spectrometric and chromatographic behaviour of peptides, which are used to deconvolute highly complex SWATH-MS data in a targeted manner. Tools such as PeakView®, Spectronaut X™, OpenSWATH or DIAUmpire can be used to analyze the SWATH-MS data against the spectral library of interest. Upon controlling false discovery rates (FDR), protein inference and quantification could be achieved by tools like aLFQ , Diffacto and mapDIA. To our knowledge, currently, there is no pipeline consolidating these various steps.

In this study, we propose a computational pipeline that integrates various open source tools to generate a final protein matrix starting with the raw input. The pipeline starts with the conversion of wiff files to MGF for data-dependent acquisition and mzML for SWATH-MS runs respectively, followed by the application of three complementary search engines (MS-GF+, Mascot and X!Tandem) against the UniProt database with decoys. The search results are combined using the PeptideShaker algorithm . Results are subsequently filtered at 1% peptide and protein levels to generate the final spectral library. OpenSWATH pipeline is then used along with the PyProphet and TRIC algorithms, with FDR being controlled for both peptides and proteins at either run-specific, experiment-wide or global contexts. Finally, the Diffacto method is used for protein inference and quantification. Exploratory data visualisation using various plots is further provided on the resulting protein matrix.

We have applied this pipeline to a benchmark cancer dataset and compared with a commercial pipeline, showing comparable results both at the peptide and protein levels. We further demonstrated that the pipeline is robust and easily scalable to any large dataset generated by SWATH-MS.

118

Transpeptidation is a major mechanism of haemoglobin proteolysis by the gingipain cysteine proteinases

Paul D Veith[1], Lianyi Zhang[1], Eric C Reynolds[1]

1. Melbourne Dental School, Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia

The gingipains are the major cell-surface proteinases and virulence factors of *Porphyromonas gingivalis*, considered a keystone pathogen of chronic periodontitis (severe gum disease). We recently showed that these cysteine proteinases prefer to catalyse transpeptidation over hydrolysis and utilise a wide range of amino acids and peptides as acyl acceptors (1). Furthermore, haemoglobin (Hb) was shown to be digested faster in the presence of added peptides implying that transpeptidation was involved in the rate-limiting step which is the initial digestion of the protein. The transpeptidation products include discontiguous Hb peptides suggesting that the pathway of Hb digestion may involve transpeptidation within the Hb structure using the free N-

terminus and other N-termini newly created by transpeptidation or hydrolysis events. To investigate this, both circular and linear transpeptidation products were identified in Hb and myoglobin and analysed with respect to their known crystal structures. Partial protein digests were analysed by orbitrap LC-MS/MS. Transpeptidation products were identified by Mascot and custom made transpeptidation databases or alternatively using pLink cross-linking software.

1. Zhang L*, Veith PD*, Huq NL, Chen YY, Seers CA, Cross KJ, Gorasia DG and Reynolds EC (2018). Porphyromonas gingivalis Gingipains Display Transpeptidation Activity J Proteome Res. 17(8): 2803–2818

119

Using quantitative proteomics to resolve genomic diagnosis of mitochondrial disease

Daniella H Hock[1], Alison Compton[2, 3], Joanna Sacharz[1], Boris Reljic[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 5000 live births. These disorders can manifest at any stage of life and compromise energy production via oxidative phosphorylation (OXPHOS) via myriad mechanisms. The current diagnosis for mitochondrial disease relies on target exome sequencing of genes encoding known mitochondrial proteins combined with clinical notes and basic biochemical measurements such as respiratory chain enzyme activity. Due to its complex genetics and phenotypic heterogeneity, over 40% of patients with suspected mitochondrial diseases remain undiagnosed. Quantitative proteomics is an untargeted approach that allows not only discovery of the disease gene but also can provide insights into the molecular function of the mutation within the patient cellular proteome.In this study, primary fibroblasts from two patients with suspected mitochondrial disease were subjected to quantitative proteomics using Tandem Mass Tag (TMT) labels. Patient 1 (P1) had no likely pathogenic variants detected after Whole Exome Sequencing (WES) or mitochondrial DNA sequencing, and no clear enzymatic defect detected in fibroblasts. Quantitative proteomics results suggest decreased abundance of proteins belonging to the large subunit of the mitoribosome compared to multiple control fibroblasts. Re-analysis of sequencing data identified two novel variants in mitoribosome genes that were targeted for follow-up studies. Patient 2 (P2) has two missense variants identified as possible causative mutations: *MT-ATP6* p.Leu49Pro (81% heteroplasmy in blood) and a heterozygous *ATAD3A* p.Thr228Met. Proteomics results demonstrate unchanged levels of ATAD3A protein across samples while MT-ATP6 is among the least abundant mitochondrial proteins in P2, together with other Complex V subunits. This suggests co-dependency between the proteins for a fully functional complex. Further investigations using Blue Native PAGE (BN-PAGE) shows strong defect in mitochondrial Complex V assembly, confirming the *MT-ATP6* p.Leu49Pro missense variant as P2 diagnosis. Hence, our study demonstrates the utility of quantitative proteomics to complement the diagnosis of mitochondrial diseases.

120

An integrated -omics approach to overtraining biomarker discovery

Jessica van Haeften[1], Tony Parker[1], Daniel Broszczak[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 and utilising inadequate recovery periods may become overtrained. This debilitating condition is characterised by persistent fatigue and an inability to maintain high-performance athletic outputs. Metabolites, lipid mediators and proteins harvested from body fluids are physiological indicators of the state of tissues, organs and general health of the individual analysed. These biomarkers could be used to monitor the state of stress and recovery from exercise when high intensity training has caused exercise-induced fatigue, stress or injury.

This discovery project will utilize quantitative protein 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]. These same samples will also be analysed for metabolites and lipid mediators to develop an in-depth view of athlete health status after being excessively trained without a recovery period. 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 a altered physiological state based upon effected biological pathways post-overtraining and better elucidate the mechanisms that underpin the processes of post-exercise muscle fatigue.

In addition, the identification of biomarkers indicating the onset of overtraining syndrome will allow for improved monitoring of athletic training, and determine key points of intervention to prevent long-term and career threatening performance decrements.

[1] Le Meur, Y., et al, "A multidisciplinary approach to overreaching detection in endurance trained athletes," Journal of Applied Physiology (2013). 114(3): 411-420.

121

Defining the contribution of linear and spliced peptides to the anti-influenza T cell response

Patricia Illing[1], Pouya Faridi[1], Nicole Mifsud[1], Grace Khuu[1], Ziyi Huang[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

Despite existing vaccine strategies, influenza viruses inflict a high level of mortality and morbidity worldwide and are a significant economic burden. Current vaccine formulations induce predominantly antibody-based immune responses against the surface hemagglutinin (HA) which can differ greatly between strains resulting in limited cross-strain protection. CD8 T cells recognise peptidic antigens generated through the breakdown of viral proteins within infected cells and displayed by the HLA class I molecules (HLA-I) at the cell surface. Thus, they have the capacity to initiate responses against more conserved internal viral proteins and therefore broad recognition across strains. Rational vaccine formulation requires knowledge of which viral components instigate immune responses. Therefore, we have utilised a mass spectrometry-based epitope discovery program, interrogating the immunopeptidome (array of peptides presented by the HLA) of diverse HLA molecules to define the virus-derived peptides available for immune surveillance. Empowered by an in-house developed workflow that combines data-driven *de novo* sequencing (PEAKS 8.5) with proteome database searching, we have mapped the contribution of both genome templated (linear) and a newly recognised subset of non-genome templated (spliced) peptides. Current work moves to assess the contribution of both linear and spliced peptide subsets to the anti-influenza immune response in healthy donors. Not only are these data sets helping define the rules for peptide splicing, but they will inform the design of vaccine strategies that maximise T cell responses and the progress towards a universal influenza vaccine.

122

Development of an optimized end to end phosphoproteomic workflow on the timsTOF PRO

Rune H Larsen[1, 2], Jarrod J Sandow[1, 2], Giuseppe Infusini[1, 2], Laura F Dagley[1, 2], Andrew I Webb[1, 2]

1. Walter and Eliza Hall Institute, Parkville, VIC, Australia

2. Department of Medical Biology, University of Melbourne, Melbourne, Victoria, Australia

Phosphorylation is an important post-translational modification (PTM) that occurs on proteins to modulate how intracellular signals propagate within cells in response to different stimuli or pathological conditions. Phosphoproteomics has established itself as the tool of choice to investigate the complexities of intracellular signalling cascades in an unbiased manner.

However, the utility of phosphoproteomics has been hampered by the large amount of input material normally required to enable high levels of phosphopeptide identifications and accurate quantitation (>1 mg). This limitation has restricted the majority of phosphoproteomics studies to *in vitro*experiments whereas the true power of phosphoproteomics is likely to be revealed from studies of primary cells. Traditional phosphopeptide enrichment strategies have often relied on large amounts of material due to sample losses during desalting steps, lower phosphoenrichment specificity and the need to perform comprehensive fractionation which often results in reduced yields and increased technical variability. We have developed an optimised end-to-end phosphoproteomic workflow by improving cell lysis, protein digestion, phosphopeptide enrichment and identification by mass spectrometry to vastly decrease the required sample input amount. The increase in speed and sensitivity of the PASEF (Parallel Accumulation–Serial Fragmentation) acquisition strategy on the timsTOF Pro (Bruker) has greatly aided the identification of phosphorylated peptides without the need for extensive fractionation.

We have optimised collision energies, gradient times and trapped ion mobility spectrometry (TIMS) settings in order to refine the PASEF strategy for ultrasensitive phosphoproteomics. Using our end-to-end phosphopeptide enrichment workflow, we demonstrate the capabilities of the timsTOF by enabling robust phosphopeptide quantitation with minimal starting material.

123

Development of proteomic tools to identify telomere maintenance mechanisms in human cancer

Yangxiu Wu[1], Sadia Mahboob[1], Asim Anees[1], Andrew Robinson[1], Jane Noble[1], Steve Williams[1], Peter Hains[1], Qing Zhong[1], Roger Reddel[1]

1. Children's Medical Research Institute, Westmead, NEW SOUTH WALES, Australia

Publish consent withheld

124

Recognising the role of post-translational modification in immunopeptidomic peptides by using mass spectrometry and bioinformatics

Shutao Mei[1], Nathan P Croft[1], Jiangning Song[1], Anthony W Purcell[1]

1. Department of Biochemistry and Molecular Biology and Infection and Immunity Program, Biomedicine Discovery Institute, Monash University, Clayton, VIC, Australia

The peptides modified by post-translational modification (PTM) then presented by the human leukocyte antigen (HLA) encoded molecules can be recognised by T cells as these modified peptides are not encoded by the human genome. The presence of such modified peptides in autoimmune diseases have been demonstrated can be associated with the development of some autoimune diseases, such as type 1 diabetes, celiac disease. Therefore, predicting PTM-peptides presented by HLA molecules is crucial for understanding the role of PTM-peptdies in the development of the autoimmune diseases. However, compared to the development of many PTM predictors that predict the PTM sites of proteins. So far, little is known about the characteristics of PTM-peptides in the immunopeptidome repertoire - i.e. the repertoire and sequences of PTM-peptides after proteolysis and presented by HLA class I (HLA-I) molecules.

In our work, we have found the deamidated peptides presented by HLA molecules show different characteristics compared to the deamidated peptides identified from proteolysis by using mass spectrometry. This finding may be helpful to better understand the role of deamidation in the development of some autoimmune diseases, such as celiac disease and type 1 diabetes. Moreover, predictor trained by deamidated peptides derived from immuno-peptidomics will be designed and built to predict deamidated peptides for further immunogenic screening.

This work aims to understand the characteristics of PTM-peptides presented by HLA molecules. Using mass spectrometry following bioinformatic works for building the next generation of PTM prediction models to screen immunogenic PTM peptides.

125

Proteomics unravel novel molecular mechanisms of Alzheimer's disease in retina

Mehdi Mirzaei[1, 3, 2], Liting Deng[3], Kanishka Pushipitha[2], Ardeshir Amirkhani[1, 3], Yunqi Wu[1, 3], Karthik Kamath[1, 3], Matthew McKay[1, 3], Vivek Gupta[1, 3]

1. Australian Proteome Analysis Facility, Macquarie University, Sydney, NSW, Australia

2. Clinical Medicine , Macquarie University, Sydney, NSW, Australia

3. Molecular Science , Macquarie University, Sydney, NSW, Australia

Publish consent withheld

126

Analysis of cysteine redox post-translational modifications in a rat model of type II diabetes mellitus

Shivanjali J Lingam[1], Desmond K Li[1], Alexander W Rookyard[2], Stuart J Cordwell[3], Melanie Y White[3]

1. Discipline of Pathology, School of Medical Sciences, Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

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

3. 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) and associated complications remains a major challenge world-wide. Several animal/human and tissue/cell-based studies have found reactive oxygen and nitrogen species (ROS/RNS) as key mediators of signal transduction during T2DM and T2DM-mediated diseases. It has been suggested that cellular antioxidant defenses are incapable of compensating for T2DM-induced ROS/RNS production. The associated oxidative stress modulates protein function by accelerating redox protein post-translational modifications (PTM) on cysteine residues (e.g. S-glutathionylation, S-nitrosylation and S-sulfenylation). Our group has defined changes in several protein PTMs in T2DM and T2DM-mediated diseases and aim to understand the interplay between these various PTMs. Here, we examined the redox-modified renal proteome in T2DM. Rats were fed 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 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. These analyses further the scope of PTMs associated with DM and may further understanding of molecular mechanisms underlying metabolic disturbances of this disease.

127

Detection of collagen proteins of bovine and porcine origin in food products using targeted LC-MS

Matthew J Fitzhenry[1], Ardeshir Amirkhani[1], Thiri Zaw[1], Xiaomin Song[1], Matthew J McKay[1]

1. Australian Proteome Analysis Facility, Department of Molecular Sciences, Macquarie University, North Ryde, NSW, Australia

The presence of animal derived collagen proteins in various foodstuffs poses concern for some consumers, whether for ethical (vegetarian or vegan) or religious reasons. Food labelling laws require all ingredients to be listed, however there is no requirement to indicate the species of origin for products containing protein based additives such as gelatin. Gelatin is composed primarily of type I, II and III collagen, and is derived from acid or alkaline hydrolysis of skin, bones and hide, from porcine and bovine sources. Each collagen proteoform represents a unique opportunity to assess the origin of gelatin used in food manufacturing, and confers proteomics a direct advantage in detecting foodstuffs adulterated with gelatin.

We have developed sample preparation procedures to optimise the recovery of collagen from various yoghurt and confectionery products, and have also developed both untargeted and targeted LC-MS procedures for the identification of proteotypic peptides which uniquely differentiate collagen isoforms at the species level. Using targeted peptide assays we report on the detection of species specific collagen proteoforms using multiple reaction monitoring – high resolution (MRM-HR) and parallel reaction monitoring (PRM) mass spectrometry, and evaluate the detection of porcine and bovine derived gelatin used in the preparation of twenty five yoghurt and confectionery products commonly available in Australian supermarkets.

Our findings indicate that targeted LC-MS can be used to differentiate different collagen proteoforms from bovine and porcine sources, and provides species specific information to distinguish the source of gelatin used in the preparation of various foodstuffs.

1. Detection of gelatin adulteration using bio-informatics, proteomics and high-resolution mass spectrometry, C. T. Yang, D. Ghosh, and F. Beaudry, Food Additives & Contaminants: Part A, 2018.

128

Understanding the roles of subunits and assembly factors implicated in the biogenesis of human Complex II

Joanna Sacharz[1], Boris Reljic[1], Daniella H Hock[1], David de Souza[2], Michael T Ryan[3], Dedreia Tull[2], David A Stroud[1]

1. Department of Biochemistry and Molecular Biology, Bio21, The University of Melbourne, Melbourne, VIC, Australia

2. Metabolomics Australia, Bio21, The University of Melbourne, Melbourne, VIC, Australia

3. Department of Biochemistry and Molecular Biology, Monash University, Melbourne, VIC, Australia

We sought to explore the roles of assembly factors (SDHAF1-4) in the biogenesis and function of Complex II (CII), namely SDH (Succinate Dehydrogenase) in human cell lines. Most of our knowledge about the function of CII assembly factors comes from studies performed in the yeast *S. cerevisiae*. We employed gene editing using CRISPR/Cas9 to generate knockout (KO) cell lines of the four core subunits of CII (SDHA-D) and its 4 known assembly factors. We carried out a comprehensive biochemical, metabolic and proteomic analysis of each KO cell line. In order to confirm the direct interactions, we performed affinity enrichment mass-spectrometry (AE-MS) of tagged subunits.

Our data suggests that the CII core subunits SDHA, SDHC, SDHD, but not SDHB, are essential for cell viability. We show altered metabolism in a mutant of SDHA expressing <10% of the native protein, as well as strong turnover of other CII subunits. Loss of SDHB results in turnover of SDHC and SDHD, but not SDHA. Knockout of SDHAF1-4 assembly factors had no impact on the levels of CII core subunits and only minimal impact on CII assembly. Conversely loss of SDHAF1 and SDHAF3, as well as SDHA resulted in turnover of ALDH1L1 and ALDH1L2 (cytosolic and mitochondrial isoforms of 10-formyltetrahydrofolate dehydrogenase) implicated in one-carbon metabolism, and PCK2 (phosphoenolpyruvate carboxykinase) involved in the conversion of oxaloacetate (OAA) to phosphoenolpyruvate in the TCA cycle. This data suggests that CII subunits and assembly factors have yet to be elucidated roles in mammalian mitochondria.

129

Improved lipid annotation utilizing positive and negative ion MS^2 / MS^3 HCD and CID spectra

Amanda Souza[1], David A Peake[1], Reiko Kiyonami[1], Daniel Gachotte[2], Gavin Reid[3], Yasuto Yokoi[4], Andreas Huhmer[1]

1. Thermo Fisher Scientific, San Jose, California, United States

2. Corteva Agriscience, Indianapolis, Indiana, United States

3. The University of Melbourne, Parkville, Victoria, Australia

4. Mitsui Knowledge Industry, Atago, Tokyo, Japan

Introduction

We present improvements to Thermo Scientific™ LipidSearch software. New algorithms were introduced to reduce false positives, improve quantitation and automate searching of $LC\text{-}MS^n$ data obtained by higher collisional energy (HCD) and linear ion trap collisional induced dissociation (CID). The use of $LC\text{-}MS^n$ is applied for more complete characterization of triacylglycerols in total lipid extracts.

Methods

Bovine liver and insect larvae extracts were analyzed by data dependent LC/MS^2 and MS^3 using a Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer. The MS^n data were searched against the m/z of selected lipid precursor ions and their predicted product ions and neutral losses. Each lipid adduct ion is ranked by mass tolerance, match to the predicted fragmentation and fraction of total MS-MS intensity.

Results

The number of lipid species annotated in each experiment was assessed at the sum composition (MS) and isomer (MS^n) levels. LC-dd-MS^2/MS^3 spectra for potential lipid species were annotated separately from positive and negative ion adducts and then merged into a single lipid result. This approach provides lipid annotation that reflects the appropriate level of MS^2/MS^3 product ions and neutral losses from giving higher confidence in lipid annotations. Results were filtered by the minimum number of data points, signal-to-noise ratio, adduct ion, match score, ID quality, and coefficient of variation from replicate injections. Compared to the results generated only from dd-MS^2 HCD, a combination of HCD and CID LC-MS^n gave significantly higher quality lipid identifications. From bovine liver, the number of PC and TG results with complete acyl chain information improved by 12% to 23%, respectively.

Conclusion

Lipid annotation is often over-reported because the current software/database approaches match mass spectral evidence to exact lipid structures [1]. The lipidomics community is working towards reporting the correct level of annotation based upon the available mass spectral information. Utilizing a combination of HCD MS^2 and intelligent product ion or neutral loss triggered CID MS^2/MS^3 data improves lipid annotation available within a single LC/MS^n experiment.

1. Rustam Y. H., Reid, G. E. Anal. Chem., 90(1), 374-397 (2018).

130

A multiplexed enrichment and targeted PRM platform for absolute quantitation of AKT/mTOR, Ras, and p53 signaling pathways targets

Aaron S Gajadhar[1], Steve Binos[2], Bhavin Patel[3], Penny Jensen[3], Leigh Foster[3], Sebastien Gallien[4], Jonathan R Krieger[5], Rosa Viner[1], Andreas Huhmer[1], Kay Opperman[3], John C Rogers[3]

1. Thermo Scientific, San Jose, CA, USA

2. Thermo Fisher Scientific, Parkville, VIC, Australia

3. Thermo Fisher Scientific, Rockford, IL, USA

4. Thermo Fisher Scientific PMSC, Cambridge, MA, USA

5. SPARC Biocentre, The Hospital for Sick Children, Toronto, ON, Canada

The AKT/mTOR and RAS/ERK pathways represent mechanisms for cells to regulate survival, proliferation, and motility. These signaling pathways play a central role in tumor progression and drug resistance. Highly accurate monitoring of these pathway proteins has not been achieved, due to poor reproducibility, unreliable quantitation, and lack of standardized methods and reagents. To overcome these challenges, the novel SureQuant™ pathway panels have been applied, which utilize an optimized multiplex immunoprecipitation to a targeted mass spectrometry (mIP-tMS) workflow. mIP-tMS assays can quantitate multiple proteins, PTMs and interacting partners, which creates new possibilities for a broad range of applications, including cancer diagnosis and prognosis, drug development, and precision medicine.

The SureQuantTM AKT pathway (total or phospho), RAS, or TP53 workflow includes a multiplex IP module (antibodies and lysate), MS sample prep, absolute or relative quantitation modules (AQUA Ultimate peptides standards), and standardized data analysis pipeline. Serum-starved, inhibitor-treated (LY294002/NVP-BEZ235/Rapamycin) HCT116, A549, and MCF7 cells were stimulated with hIGF-1. IP-enriched, digested samples were spiked with heavy peptides and analyzed using optimized targeted MS (nanoLC-PRM/MS) and Skyline software. The panels were benchmarked against Western blotting (WB) using three unstimulated, hIGF-1 stimulated or inhibited cell lysates, as well as several tissue lysates.

Previously, we showed the feasibility of optimized mIP-tMS assays to quantitate AKT and RAS pathway proteins across 2 cancer cell lines ± LY294002. The SureQuant™ multiplex pathway panels allowed absolute quantitation of multiple total and phosphorylated targets in low to sub-nanogram concentrations across three unstimulated, hIGF-1 stimulated and inhibited cell lysates as well as tissue/xenograft lysates. Analysis by mass spectrometry allowed for more accurate and informative data with the determination of fmol levels of protein expression and capability to discriminate between isoforms of many proteins that are unable to procure with western blot analysis.

SureQuant™ pathway panels allowed simultaneous absolute quantitation of AKT pathway, RAS proteins and PTMs in a streamlined, standardized workflow.

131

Degradomic analysis of a putative C-terminal processing peptidase from *Campylobacter jejuni*

Joel A Cain[2, 1], Paula Niewold[2, 3], Sherif Abouelhadid[4], Brendan W Wren[4], Stuart J 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. The London School of Hygiene & Tropical Medicine, London

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

Campylobacter jejuni is one of the leading causes of acute gastroenteritis in the developed world. The serine peptidase Cj0511, homologous to C-terminal processing peptidases, has been shown to be required for optimum virulence in its influences towards biofilm formation, stress tolerance and pancreatic amylase triggered α-dextran secretion. There is also the implication that Cj0511 may be involved in host protein degradation during infection, given it is a component of secreted outer membrane vesicles. Extensive characterisation of this protein has to date been lacking, with our previous indirect attempts via quantitative proteomic and degradomic analysis of wild-type *C. jejuni* and a Cj0511 knock-out (Δ*cj0511)* failing to find identify putative endogenous protein targets against a background of elevated proteolytic activity in its absence.

Here we chose to take a direct approach by characterising the cleavage specificity of recombinant Cj0511, using *N*-terminal amine isotopic labelling of substrates (*N*-TAILS). On a defined substrate of B-casein we show that Cj0511 possesses distinct proteolytic activity. Putative native substrates of Cj0511 were also identified through analysis of differential proteolytic activity between treated and untreated *C. jejuni* whole cell lysates derived from Δ*cj0511*.

132

Detect more proteins with decreased false positives using filtered SWATH peptide libraries to improve plasma biomarker studies

Xiaomin Song[1], Jemma Wu[1], Thiri Zaw[1], Dana Pascovici[1], Mark P Molloy[2]

1. Australian Proteome Analysis Facility, Macquarie University, NSW, Australia

2. Bowel Cancer and Biomarker Laboratory, The University of Sydney, St Leonards, NSW, Australia

Data-independent acquisition of peptide mass spectrometry data has the potential to enable improved quantitative reproducibility for plasma biomarker studies. One approach for peptide identification is to take advantage of existing peptide spectral libraries, as these can be merged with a local seed library to make an extended reference library using software such as *SWATHXtend,* as we have previously reported. Important to recognise when merging libraries is that the concomitant larger extended library yields increased probability of false-positive extraction. In this study, we explored optimising plasma SWATH library generation aiming to maximum protein coverage, while minimising false-positive detections.

We used a locally acquired plasma library as a seed to make two extended libraries by merging spectral data downloaded from the plasma dataset published by Liu et al [1] (1885 proteins) and from the human SWATH library in SWATHAtlas after selecting for plasma proteins reported in the HPP-2017 update [2] (3286 proteins). Data was acquired on a TripleTOF 6600 using 60min LC SWATH runs from five human plasma samples. We used PeakView for peptide extraction with protein FDR set at 99% confidence. Combining only the proteins detected from the two extended libraries with our local seed library, we obtained a new SWATH plasma library containing 1161 proteins. This library is 38% and 65% smaller than the original extended two libraries respectively. It is more specific to the plasma samples therefore detected more proteins with fewer false positives than using any of the individual local or extended reference libraries.

References:

1. Liu, Yansheng, et al. "Quantitative variability of 342 plasma proteins in a human twin population." *Molecular systems biology* 11.2 (2015): 786.
2. Schwenk, JM., et al. "The human plasma proteome draft of 2017: building on the human plasma PeptideAtlas from mass spectrometry and complementary assays." *Journal of proteome research* 16.12 (2017): 4299-4310.

133

Exploring the bacteriophage φX174 infection response

Bradley W Wright[1], Dominic Y Logel[1], Paul R Jaschke[1], Mark P Molloy[1, 2]

1. Molecular Sciences, Macquarie University, Sydney, NSW, Australia

2. Northern Clinical School, Sydney University, Sydney, NSW, Australia

We have investigated the transcriptomic and proteomic response of *Escherichia coli* C122 to φX174 infection and found that there is a large and rapid cellular response present during the later stages of the 60-minute infection cycle. Using TMT-labelling and mass spectrometry we quantified approximately 46% of the *E. coli* proteome and observed large fold-change inductions of bacterial proteins involved in membrane biogenesis, in particular the Sec protein translocation and lipoprotein processing and trafficking pathways. The cell's heat shock response and protein folding machinery were largely upregulated. "Holdases" IbpA and IbpB had fold-changes equivalent to φX174 proteins indicating their significance during infection. Separate Crispr-Cas9 mediated knockouts of these two genes led to no observable changes in infection dynamics, however, double knockouts resulted in lower efficiency of plating and significantly larger plaques. Interestingly, Δ*IbpA/B* also resulted in a slower growth phenotype at a wide range of temperatures, relative to the wild-type and single knockout strains, and plasmid complementation with the native genes only partially resolved the defect. These preliminary results are in contrast to gene disruptions in the K12 strain, and will require further exploration.

A further experiment aims to understand the significance of φX174's highly compressed genome. We have a genome decompressed variant which we have observed to have the same lysis timing as the wild-type, but displays fitness defects in burst size and plaque formation. We endeavour to investigate the effects of genome decompression on protein expression using the targeted mass spectrometry method of parallel reaction monitoring on the wild-type phage and decompressed variant and will present preliminary results.

134

Proteomic profiling of the cell surface of senescent cells

Birgit Schilling[1], Herbert Kasler [1], Eric Verdin[1], Judith Campisi[1], Nathan Basisty[1]

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

Introduction: Cellular senescence is a stress response that causes an irreversible arrest of cell proliferation and development of a senescence-associated secretory phenotype. Pre-clinical mouse studies have demonstrated that targeted removal of senescent cells mitigates a variety of aging phenotypes. By current methods, senescence must be induced in culture to examine the characteristics of senescent human cells. However, a comprehensive understanding of the phenotype(s) of these cells in humans is lacking because they cannot, thus far, be isolated from tissues. Therefore, senescent cell-surface markers are needed to identify, isolate, characterize and ultimately target senescent cells for therapeutic interventions.

Methods: Senescent and non-senescent primary human fibroblasts (IMR90) were generated in cell culture by X-irradiation or mock irradiation. Cell-surface proteins were isolated using 'Cell-Surface Capture', an in-vivo biotinylation/affinity enrichment approach. Data-independent acquisitions (DIA) were performed on an orthogonal quadrupole time-of-flight TripleTOF 6600 mass spectrometer. Protein identification was performed with ProteinPilot and Spectronaut Pulsar. Secondary validation of surface proteins was performed using flow cytometry (BD LSR II).

Results: We identified 59 proteins significantly increased and 44 proteins significantly decreased on the surface of senescent compared to non-senescent human fibroblasts. Senescent cell-surface proteins were enriched for proteins in exosomal machinery, DNA damage response and cell death pathways. Three surface proteins – CALR, VCAM1, and CD44 – were independently validated by flow cytometry.

Conclusions: Our comprehensive proteomics approach has revealed novel, differentially expressed proteins for the cell-surface, exosomes, and the secretome of senescent cells. We will further focus on validating surface protein candidates and assessing their utility as biomarkers in order to isolate senescent cells from intact human tissues.

135

The role and regulation of Atg1 in autophagy and DIF-1 induced autophagic cell death of *Dictyostelium discoideum*

Melodi G Sen[1], Sarah J Annesley[1], Paul R Fisher[1]

1. La trobe University, Bundoora, VIC, Australia

Autophagy is a central process for the maintenance of cell homeostasis and survival of mitochondrially diseased cells. It is a basally active, catabolic process, targeting cytosolic macromolecules and organelles that are unnecessary or dysfunctional. These are degraded and recycled through the activity of lysosomal enzymes. In neurodegenerative and mitochondrial diseases, cells upregulate autophagy in a bid to remove protein aggregates and damaged organelles. After a point of no return, this process results in autophagic cell death. The presence of apoptotic machinery in mammalian cells obscures the study of autophagy in relation to complex diseases. Therefore, the social amoeba, *Dictyostelium discoideum,* provides a simple model which lacks these apoptotic genes. Furthermore, *D. discoideum* has homologues for autophagy proteins, which makes it an ideal candidate for this research. Autophagic pathways are crucial for the multicellular development of *D. discoideum,* as cells targeted for differentiation into stalk cells undergo autophagic cell death. An important regulator of this process is AMP-activated protein kinase (AMPK), which plays a central role in cellular processes such as growth, cell cycle progression and photosensory signal transduction. Downstream of AMPK lies TOR complex 1 (TORC1), which inhibits autophagy, and stimulates cell growth and proliferation. The regulatory activity of TORC1 on autophagy has been found to occur through the ULK1 complex in mammalian cells. In *D. discoideum*, the role of the ULK1 counterpart, Atg1, has yet to be elucidated. Therefore, this project aims to characterise the role of Atg1 and determine the cellular processes effected by altered expression of Atg1. Furthermore, a thorough understanding of the interactions between Atg1, TORC1 and AMPK in the regulation of autophagic cell death, could provide future targets for therapeutics. Preliminary data suggests which proteins are upregulated or downregulated with a change in Atg1 expression. Furthermore, Atg1 has proven to be essential for normal aggregation, stalk cell differentiation and fruiting body formation in *D. discoideum.* This suggests that Atg1 is crucial for long term viability of the organism.

136

Using integrative multi-omics to explore *Pseudomonas aeruginosa* cellular physiology after adaptation to the Cystic Fibrosis lung

William Klare[1], Paula Niewold[2], Joel Cain[1], Emily King[1], Hannah Nicholas[1], Jamie Triccas[3], Stuart Cordwell[1, 2, 4]

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

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

3. Department of Infectious Diseases and Immunology, 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 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.

137

Development of a quantitative proteomic standard for Tandem Mass Tags (TMT)

Samah Issa[1], Jae Choi[2], Aaron Robitaille[3], Tabiwang Arrey[4], Rosa Viner[3], Andreas Huhmer[3], John Rogers[2]

1. ThermoFisher Scientific, Melbourne, VIC, Australia

2. Thermo Fisher Scientific, Rockford, IL, USA

3. Thermo Fisher Scientific, San Jose, CA, USA

4. Thermo Fisher Scientific, Bremen, Germany

Quantitative proteomics strategies using Tandem Mass Tags (TMT) enablable sample multiplexing and precise measurement of protein abundance. However, co-isolated ion interference can suppress accurate ratio quantification. Employing MS3 methods with synchronous precursor selection (SPS) on Orbitrap Tribrid mass spectrometers can minimize ion interference. Therefore, we developed a standardized commercially available TMT11plex yeast digest standard to detect co-isolation interference and enable MS method optimization.
Here, we utilized TMT11plex to label peptides from four strains of Saccharomyces cerevisiae; a parental line and three lines respectively lacking the non-essential protein Met6, His4, or Ura2. Tryptic peptides from the strains lacking gene MET6, HIS4, OR URA2 were labeled in triplicate, while the parental line was labeled in duplicate.
We demonstrate that a TMT11plex yeast digest standard can be used as a proteomic reference standard to measure protein/peptides identification and optimize acquisition and data analysis methods to limit co-isolation interference, as well as diagnosis MS instrument status by monitoring mass accuracy, ion injection time, and reporter ion signal to noise. We then used the standard to establish a standardized workflow including two LC methods (50min or 120min gradients) for a variety of nano-spray liquid chromatography setups including Easy-nLC 1200 and 1000, and Dionex U3000, optimized MS acquisition settings for Hybrid or Tribrid Orbitrap mass spectrometers, and data analysis in Proteome Discoverer 2.2. The TMT11plex yeast digest standard provides mass spectrometry users a tool to measure the accuracy, precision, and dynamic range assessments for different mass spectrometry approaches, and is an excellent quality control assay to the assessment of the LC and MS instrument status when combined with a standardized workflow.

138

An "OMICS" approach to food profiling, authenticity, and classification

Katherine Hyland[1]

1. SCIEX, Redwood City, CA, United States

Food analysis has been expanded to incorporate all manner of post-genomic profiling, classification, nontargeted screening, and authentication of any of a variety of food, environmental, or agricultural products. We will demonstrate workflows utilizing SCIEX high resolution mass spectrometric technology, novel data acquisition techniques, spectral libraries, and software tools to utilize nontargeted and statistical approaches to classify such samples as juices, herbs, and cannabis strains, and identify unknown constituents in complex matrices like citrus fruits extracts and even environmental water samples.

139

AI-based solutions for MS-based immunopeptidomics with high sensitivity and accuracy

Paul Shan[1]

1. Bioinformatics Solutions, Waterloo, ONTARIO, Canada

Identification of tumor-specific antigens (neoantigens) is needed for development of effective cancer immunotherapy and a good source for such antigens are the pools of HLA-bound peptides presented exclusively by the tumor cells. Mass spectrometry (MS) has evolved as the method of choice for the exploration of the human immunopeptidome (HLA class-I and class-II peptides). Workflows for immunopeptidomics are different from ones for more established shotgun proteomics, yet inherent differences between these two fields create significant drawbacks of current data analysis algorithms for the former. We provide AI-based solutions to address the barriers for data analysis, e.g., diverse C-termini of HLA-peptides, lack of sequence library for spliced peptides, no peptide de novo sequencing algorithms for data independent acquisition (DIA) method, etc. With deep learning technology, we proposed a new approach integrating motif-constrained database search and de novo sequencing for HLA-peptide identification that would increase peptide sequencing coverage, depth, and confidence, collectively enhancing the capabilities of the field of immunopeptidomics. Our approach was tested with several public data sets for the identification of HLA-peptides or neopeptides, including DDA (PXD007596, PXD006939, PXD004964, and PXD002431) and DIA (PXD001094). On average, 60% more HLA-peptides were identified by validation with the public immunopeptidome databases, IEDB and NeuroPep. Preliminary results showed that the AI-based data analysis would provide a novel solution for immunopeptidomics with high sensitivity and accuracy.

Reference

1. Secher, A., Kelstrup, C.D., Conde-Frieboes, K.W., Pyke, C. *et al*., *Nat Commun*, 2016, 7:11436
2. Abelin, J.G., Keskin, D.B., Sarkizova, S., Hartigan, C.R., *et al*., *Immunity*, 2017, 46:315-326
3. Faridi, P., Purcell, A.W., Croft, N.P., Proteomics, 2018, 18:1700464

140

MS evidence acquisition of missing proteins in chromosome 9 using halo tag purification system and identification of cellular roles.

HuiSu Kim[1], Dong Wook Kim[1], Hyoung-Min Kim[1], Yong-In Kim[2], J. Eugene Lee[2], Je-Yoel Cho[1]

1. BK21 PLUS Program for Creative Veterinary Science Research and Research Institute for Veterinary Science College of Veterinary Medicine, Seoul National University, Seoul, South Korea

2. Bioanalysis center, Korea Research Institute of Standards and Science, Daejeon, South Korea

The NeXtProt release 2018-09-17 reports 2,890 of proteins are remaining in missing proteins level, also known as PE 2,3,4 level (14.2%). Our goal is finding MS evidence of missing proteins and identification of biological roles in cells. First, we tried to find missing proteins with no MS evidences in Chromosome 9 using NeXtProt database. High obstacles acquiring MS evidences of missing proteins due to the very low expression level or specific expression pattern in certain cell types or tissues drive us to take advantage of overexpression system. In brief, we constructed 5 plasmids DNA (FOXD4, ARID3C, OR1J1, ANKRD18A, ZNF510) harboring cDNA of missing proteins fused to Halo Tag and artificially transfected into HEK293T cells. Overexpressed missing proteins were purified with Halo tag purification system and subjected to LC-MS/MS analysis. This strategy gave us great advance in detection of MS evidences for missing proteins. Using this method, we definitely get 2 MS evidences of missing protein (FOXD4, and ARID3C) under the rule of Human Proteome Project Data Interpretation Guidelines. In short, we used followed search parameter; peptide FDR < 0.01, protein FDR < 0.01, and 2 or more unique peptides with 7 amino acids. Additional cellular analysis identified that FOXD4 and ARID3C were localized in cytosol and nucleus respectively. In conclusion, overexpression and purification system we used in the study may be alternative method to identify MS evidences of missing proteins.