1

A proximity map of a human cell

Anne-Claude Gingras[1, 2]

1. Mount Sinai Hospital, Toronto, ON, Canada

2. Molecular Genetics, University of Toronto, Toronto, ON, Canada

Compartmentalization is essential for all complex forms of life. In eukaryotic cells, membrane-bound organelles, as well as a multitude of protein- and nucleic acid-rich subcellular structures, maintain boundaries and serve as enrichment zones to promote and regulate protein function, including signalling events. Consistent with the critical importance of these boundaries, alterations in the machinery that mediates protein transport between these compartments have been implicated in a number of diverse diseases. Understanding the composition of each cellular “compartment” (be it a classical organelle or a large protein complex) remains a challenging task. For soluble protein complexes, approaches such as affinity purification other biochemical fractionation coupled to mass spectrometry provide important insight, but this is not the case for detergent-insoluble components. Classically, both microscopy and organellar purifications have been employed for identifying the composition of these structures, but these approaches have limitations, notably in resolution for standard high-throughput fluorescence microscopy and in the difficulty in purifying some of the structures (e.g. p-bodies) for approaches based on biochemical isolations. Prompted by the implementation *in vivo* biotinylation approaches such as BioID, we report here the systematic mapping of the composition of various subcellular structures, using as baits proteins (or protein fragments) which are well-characterized markers for a specified location. We defined how relationships between “prey” proteins detected through this approach can help to understand the protein organization inside a cell, which is further facilitated by newly developed computational tools. We will discuss our low-resolution map of a human cell containing major organelles and non-membrane bound structures, but also describe how this map can be harnessed to uncover new signalling components implicated in cancer and vascular diseases.

2

Strategies to improve reproducibility of large-scale and longitudinal proteomics

Rebecca C Poulos[1], **Peter G Hains**[1], **Rohan Shah**[1], **Natasha Lucas**[1], **Dylan Xavier**[1], **Srikanth Manda**[1], **Asim Anees**[1], **Jennifer MS Koh**[1], **Sadia Mahboob**[1], **Max Wittman**[1], **Steven G Williams**[1], **Erin K Sykes**[1], **Michael Hecker**[1], **Michael Dausmann**[1], **Merridee A Wouters**[1], **Keith Ashman**[2], **Jean Yang**[3], **Peter Wild**[4, 5], **Anna deFazio**[6, 7, 8], **Rosemary Balleine**[1], **Brett Tully**[1], **Ruedi Aebersold**[9, 10], **Terence P Speed**[11, 12], **Yansheng Liu**[13, 14], **Roger R Reddel**[1], **Phillip J Robinson**[1], **Qing Zhong**[1]

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

2. Sciex, 2 Gilda Court, Mulgrave, VIC, Australia

3. School of Mathematics and Statistics, The University of Sydney, Sydney, NSW, Australia

4. Dr. Senckenberg Institute of Pathology, University Hospital Frankfurt, Frankfurt am Main, Germany

5. Department of Pathology and Molecular Pathology, University Hospital Zurich, Zurich, Switzerland

6. Centre for Cancer Research, Westmead Institute for Medical Research, Westmead, NSW, Australia

7. Faculty of Medicine and Health, The University of Sydney, Westmead, NSW, Australia

8. Department of Gynaecological Oncology, Westmead Hospital, Westmead, NSW, Australia

9. Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, Zürich, Switzerland

10. Faculty of Science, University of Zürich, Zürich, Switzerland

11. Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia

12. Department of Mathematics and Statistics, University of Melbourne, Melbourne, VIC, Australia

13. Department of Pharmacology, Yale University School of Medicine, New Haven, CT, USA

14. Yale Cancer Biology Institute, Yale University, West Haven, CT, USA

Reproducibility of research results is the bedrock of experimental science. Reproducibility is particularly difficult to achieve in large-scale studies on inherently variable clinical samples. SWATH mass spectrometry (MS) has emerged as a robust proteomic method with the capacity to acquire data at high sample throughput. The reproducibility of SWATH-MS proteomic measurements of human tissue samples, acquired across multiple instruments over time in a single laboratory, has not been established. We aimed to assess the reproducibility of industrial- and clinical-scale SWATH-MS and develop computational methods for improving quantitative accuracy. To this end, we performed 1,560 SWATH-MS runs of eight samples comprising a dilution series of prostate cancer tissue in fixed proportion (50%), with a variable fraction of ovarian cancer tissue (3-50%) offset by yeast cells, and a control cell line. Replicates were run on six mass spectrometers operating continuously with varying maintenance schedules over four months, interspersed with more than 5,000 runs from unrelated studies. We first applied a normalisation strategy that utilises negative controls and replicates to remove unwanted variation and elevate biological signal with greater success than existing approaches frequently applied in proteomics. We next developed a strategy for replacing missing values in normalised data by leveraging measurements acquired from replicates spanning multiple instruments. We integrated these new computational modules into a pipeline called *ProNorM* (Proteomics Normalisation and Missingness Removal). With *ProNorM*, we could mitigate technical variation between instruments across extended periods and rescue approximately 20% of values missing for non-biological reasons. *ProNorM* enabled the detection of peptide intensity changes at low concentrations in a dilution series comprising complex human cancer tissues, and allowed the prediction of tissue content in mixed samples with high accuracy by machine learning. Taken together, we demonstrate large-scale SWATH-MS data to be comparable over extended periods, providing a pathway toward reproducible clinical proteomics.

3

The use of natural products in the leather industry: Depilation without damage.

Yi-Hsuan Tu[1], Meekyung Ahn[1, 2], Mark L Patchett[1], Rafea Naffa[2], Dragana Gagic[1], Jasna Rakonjac[1], Gillian E Norris[1]

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

2. New Zealand Leather and Shoe Research Association, Palmerston North, Manawatu, New Zealand

The removal of hair from skin (depilate) is traditionally done by mechanical means, (shaving) or using chemicals. Although chemicals are very effective, they can damage or burn the skin and the waste products are an environmental problem. More recently enzymes have been trialled, but these are difficult to control and also often damage the skin.

We have found a simple solution; a dairy by-product that results in easy depilation without apparent skin damage and have used a number of different analyses to compare the molecular components of our depilated skins with those depilated using chemicals. Firstly SEM showed there was no obvious damage to skin depilated using the natural product and TEM showed the arrangement of collagen fibres was similar to those in skins depilated with chemicals. To assess any less visible damage to the skins, quantitative biochemical analyses were carried out to measure any changes in amino acid composition, collagen crosslink composition and concentration, and glycosaminoglycan concentration and compared to the levels in chemically depilated skins. Label free quantitative proteomic analyses were then carried out on skin samples using gelLC MSMS to identify any differences in the skin proteomes of each sample, including any differences in glycosylation and hydroxylation of the collagen molecules. In order to determine any enzyme activity that could be responsible for depilation, and any protein based compounds that might be responsible for the preservation of the skins during the process, the solutions used for depilation were first analysed by SDS PAGE, and the resulting bands also subjected to proteomic analysis. This presentation will describe the progress that has been made towards understanding the scientific rational of this depilation process and the differences in the protein components of skin before and after depilation using this new environmentally friendly method and compared to the differences seen before and after chemical depilation.

4

Functional modification of matrix metalloproteinase 9 by substrate glycosylation

Elizabeta Madzharova[1], Fabio Sabino[1], Ulrich auf dem Keller[1]

1. Department of Biotechnology and Biomedicine , Technical Univesity of Denamark, Kgs. Lyngby, Denmark

Matrix metalloproteinase (MMP)9 is a secreted protease, primarily functioning in the extracellular matrix. It is the most extensively glycosylated MMP, having two putative N-glycosylation sites in the prodomain and the catalytic domain and a heavily O-glycosylated linker domain. Importantly, also around 50% of known extracellular and membrane substrates of MMP9 are found to be either N- or O-glycosylated. However, the interplay of glycosylation and proteolytic processing of MMP9 substrates and its consequences for the modulation of biological activities have not been fully understood. Assessing the modulation of MMP9 proteolytic processing by substrate glycosylation through a combination of a multiplexed mass spectrometry-based substrate discovery approach, TMT-TAILS, and cellular engineering will allow systematically exploring the regulatory crosstalk between limited proteolysis and glycosylation in the extracellular space. Novel glycosylation-dependent processing events will be identified by a comparative assessment of MMP9 substrate cleavages in glycosylated and non-glycosylated cell secretomes, and the selected candidate substrates will be further validated and functionally characterized. This study will not only assess the fine-tuning of limited proteolysis in the extracellular space by glycosylation but also provide a better understanding of MMP9 regulation with strong implications for basic biology and translational approaches that aim at targeting aberrant protease activities in human diseases.

5

The mitochondrial unfolded protein response: a global translational program that repairs local mitochondrial damage

Louise Uoselis[1], Marvin Skulsuppaisarn[1], David A Stroud[2], Ralf B Schittenhelm[3], Michael Lazarou[1]

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

2. Department of Biochemistry and Molecular Biology, Bio21 Institute, University of Melbourne, Melbourne, Victoria, Australia

3. Monash Proteomics & Metabolomics Facility, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Mitochondria are essential organelles that provide the energy to power most biochemical reactions in our cells. Dysfunctional mitochondria have been implicated in the pathology of many diseases from cancer to diabetes, while defects in quality control pathways that maintain mitochondrial health can cause Parkinson's disease. One major mitochondrial quality control pathway is the mitochondrial unfolded protein response (mtUPR). The mtUPR is a transcriptional program that responds to mild mitochondrial damage by inducing the expression of proteins that boost both mitochondrial health and cellular function. Using a small molecule inhibitor of mitochondrial HSP90 to induce protein aggregation within mitochondria, we have shown that the mammalian mtUPR consists of three separate signalling arms mediated by the transcription factors CHOP, ATF4 and ATF5. CRISPR/Cas9 knockouts of these three transcription factors revealed that each arm of the mammalian mtUPR functions in a non-redundant manner. To understand how each transcription factor contributes to the folding state of the mitochondrial proteome during stress, we have characterised the mitochondrial stress 'aggresome'. Comparative analysis of the mitochondrial aggresome in CHOP, ATF4 and ATF5 knockout lines has revealed at a functional level how each signalling arm of the mammalian mtUPR promotes mitochondrial repair.

6

Quantitative proteomics of irreversible cysteine post translational modifications in myocardial ischemia / reperfusion (I/R) injury

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

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

2. School of Life and Environmental Sciences, 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 reactive nitrogen species (ROS/RNS). This increase in ROS/RNS overwhelms cellular antioxidant defence and can alter protein structure / function via various protein post-translational modifications (PTMs). One target of ROS/RNS is the most redox active amino acid cysteine (Cys). Cys redox PTMs can be broken down into two classes, those that are biologically reversible (e.g. *S*-glutathionylation) or those considered 'irreversible' (sulfinic and sulfonic acid; Cys-SO_2H/SO_3H). Irreversible Cys redox PTM occur with prolonged exposure to high levels of ROS/RNS and are associated with protein dysfunction and/or degradation. A mass spectrometry technique based on parallel reaction monitoring was employed to quantify changes in irreversible Cys modification in a Langendorff model of myocardial ischemia/reperfusion injury (I/R). Due to the low abundance of Cys, and low abundance of Cys PTMs, an enrichment strategy utilising strong cation exchange and hydrophilic interaction chromatography was used to better profile the changes in irreversible Cys PTM. I/R significantly increased the abundance of Cys-SO_2H/SO_3H-modified peptides from proteins involved in the tricarboxylic acid (TCA) cycle. By using a targeted metabolomic workflow we also observed concurrent perturbations in the abundance of metabolites involved in the TCA cycle 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.

7

Glycopeptide variable window SWATH for improved Data Independent Acquisition glycoproteomics

Chun Zhou[1], Ben Schulz[1, 2]

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

2. Centre for Biopharmaceutical Innovation, Australian Institute of Bioengineering and Nanotechnology, The University of Queensland, St Lucia, QLD, Australia

N-glycosylation plays an essential role in regulating protein folding and function in eukaryotic cells. *N*-glycan structures and occupancy can be impacted by the physiological state of cells and in disease. Sequential window acquisition of all theoretical fragment ion spectra mass spectrometry (SWATH-MS) is a powerful data independent acquisition (DIA) MS method for qualitative and quantitative analysis of glycoproteins and their glycan modifications. By separating the entire m/z range into consecutive isolation windows, DIA-MS allows comprehensive MS data acquisition and high-sensitivity detection of molecules of interest. The use of variable width DIA windows allows optimal analyte measurement, as peptide ions are not evenly distributed across the full m/z range. However, the m/z distribution of glycopeptides is different to that of unmodified peptides because of their modification with large complex glycan structures. Here, we improved the performance of DIA glycoproteomics by using variable width windows optimized for glycopeptides. This method allocates narrow windows at m/z ranges rich in glycopeptides, improving analytical specificity and performance for DIA glycoproteomics. We demonstrate the utility of the new variable window DIA method by comparing the glycoproteomes of wild type and *N*-glycan biosynthesis pathway deficient yeast. Our results highlight the importance of appropriately optimized DIA methods for measurement of post-translationally modified peptides.

8

Plasma proteome profiling to assess human health and disease

Philipp E Geyer[1, 2, 3, 4], Johannes Bruno Mueller[1], Peter V Treit[1], Sophia Doll[1, 2, 4], Lili Niu[1, 3], Eugenia Voytik[1], Lesca M Holdt[4], Daniel Teupser[4], Matthias Mann[1, 3]

1. Proteomics and Signal Transduction, Max Planck Insitute of Biochemistry, Martinsried, Bavaria, Germany

2. OmicEra Diagnostics , Planegg, Bavaria, Germany

3. Clinical Proteomics, NNF Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

4. Institute of Laboratory Medicine, LMU Munich, Munich, Bavaria, Germany

In the face of several challenges, mass spectrometry (MS)-based proteomics is now starting to live up to its initial promise as a generic technology for the discovery and quantification of proteins that reflect an individual´s health or disease state. We dramatically streamlined the labor-intensive proteomic workflow by novel sample preparation strategies and robotic automation, resulting in a rapid, robust and highly reproducible pipeline called 'Plasma Proteome Profiling'. The pipeline reduces the analysis time to less than 3 hours and standard operating procedures combined with robotics strongly improve the robustness. Such an analysis covers at least 50 known biomarkers that were approved by the U.S. Food and Drug Administration (FDA), while simultaneously quantifying hundreds of less described proteins with the potential to become biomarkers. Moreover, the

implementation of novel chromatography concepts such as pre-formed gradients in the Evosep liquid chromatography system, will further move plasma proteomics to the fast lane and will ultimately enable clinical applications.

The development of Plasma Proteome Profiling and its optimization for high-throughput screening allowed us to develop new concepts within the field of biomarker discovery. After we repeatedly observed proteins that tended to emerge as groups of statistically significant outliers independent of a particular study, we developed contamination marker panels for the quality assessment of individual samples and entire studies. We successfully applied these contamination marker panels to several studies, provided sample preparation guidelines and an online resource (www.plasmaproteomeprofiling.org) to assess sample-related bias. Moreover, we calculated in a longitudinal dataset that the majority of all protein levels (69%) are individual-specific. A consequence for biomarker discovery studies is that longitudinal studies are preferable. However, if not available, large-scale cohorts should be considered, in a `rectangular strategy` of biomarker discovery. So far, we have applied our pipeline to follow life-style intervention in the case of a weight loss study, the effect of surgery and extreme weight loss on the human plasma proteome, discovered new biomarkers for liver diseases and also adapted it to define Alzheimer's biomarkers in CSF.

9

Precision cancer immunotherapy - can proteomics rise to the challenge?

Anthony Purcell[1, 2], Pouya Faridi[1], Sri Ramarathinam[1], Nathan Croft[1]

1. Monash University , Clayton, VIC, Australia

2. Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

Immunotherapy, and in particular immune checkpoint blockade, has revolutionised clinical practice and improved outcomes in many cancer types. Immune checkpoint blockade essentially takes the brakes off the immune system allowing the full complement of immune effector cells to enter and attack the tumour. This begs the question – what is targeted by the infiltrating immune cells? And if there is a specific target can this response be augmented or even targeted by therapeutic or prophylactic vaccination strategies.

Key to the destruction of tumours is the activity of the killer T cells which recognise peptide antigens bound to the cell surface HLA class I molecules of the tumour cell. Here I present our efforts to develop a clinically relevant workflow for the identification of different classes of peptide antigens available for precision cancer immunotherapy and the validation of these targets from tumour biopsy material using a novel approach that is amenable even to just a few hundred cancer cells. These approaches will enable to next generation of T cell receptor-based and T cell receptor like therapies.

10

Caveolae formation alters intracellular organelle composition and function

Harley Robinson[1, 2], Jeffrey Molendijk[1], Mriga Dutt[1], Alok Shah[1], Ahmed Mohamed[1], Leonard Foster[3], Michelle Hill[1, 4]

1. QIMRBerghofer, Herston, QLD, Australia

2. Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia

3. Michael Smith Laboratories, University of British Colombia, Vancouver, BC, Canada

4. UQ Diamantina Institute , Brisbane, QLD, Australia

Abnormal expression of caveolae proteins, Caveolin-1 (CAV1) and Cavin-1 (CAVIN1), are linked with tumour progression, lipodystrophy, neurodegeneration and cardiac diseases. While these proteins are co-expressed and co-localized to the plasma membrane in healthy conditions, their dysregulation has been found to alter individual organelle content in some of the aforementioned conditions. Here, we assess the roles of CAV1 and CAVIN1 in regulating organelle content using the PC3 cell line which recapitulates the CAV1+ CAVIN1- expression in advanced prostate cancer. We previously reported ectopic expression of CAVIN1 in PC3 cells reduces oncogenicity, alters exosomal content and cholesterol distributions. As CAVIN1 remains at the plasma membrane in PC3-CAVIN1 cells, we hypothesized organelle remodelling as a mechanism for CAVIN1 in regulating PC3 cell aggressiveness. Using our integrated Protein and Lipid Organelle Profiling (iPLOP) method, we assessed the organelle membrane composition of the aggressive PC3 (non-caveolar) and tumour suppressed PC3-CAVIN1 (caveolae forming) cells. From the proteomics data, sucrose gradient profiles of 5 organelles (mitochondria, endoplasmic reticulum, lysosome, multivesicular bodies and endosomes) were defined based on 5-10 marker proteins. The lipidome of organelles were then mapped by matching lipidomics profiles. This analysis revealed striking changes in the composition of the mitochondria and endoplasmic reticulum. Specifically, we identified enrichment in long unsaturated lipid species in the mitochondrial membranes in the aggressive PC3 cells, corresponding to localization of various proteins relating to energy production, apoptosis and autophagy. Furthermore, iPLOP analysis reveal potential organelle remodelling mechanisms through lipid and protein translocation between organelles. Future work will validate these discovery data using orthogonal methods.

11

Rapid, unbiased identification of protein inclusion components from patient post-mortem brain tissue using Biotinylation by Antibody Recognition (BAR)

Stephanie L Rayner[1], Rowan Radford[1], Albert Lee[1], Roger Chung[1]

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

Many neurodegenerative diseases are characterised by the formation of insoluble protein inclusions in the brains of affected patients [1]. Although these inclusions form at late stages of disease, the identification of these components is integral to understanding the molecular underpinnings of disease.

The insolubility of aggregate components has limited the use of standard antibody-based approaches frequently used to study protein-protein interactions. In addition, typical biochemical fractionation methods (used to isolate the insoluble proteome)

requires homogenisation of post-mortem brain tissue prior to isolation and identification of insoluble proteins. A recently developed proximity-ligation method however, has enabled targeted biotinylation and subsequent isolation of aggregating components directly from fixed, post-mortem tissue of patients [2]. This method specifically labels aggregate components *in situ* prior to tissue homogenisation, thereby providing greater chance of identifying these proteins and significantly reducing the chance of non-specific contaminants.

Here, we apply Biotinylation by Antibody Recognition (BAR) followed by mass spectrometry to specifically identify the composition of cytoplasmic phosho-Tau aggregates found in Progressive supranuclear palsy (PSP) patients. BAR is a recently developed method, by which a primary antibody recognises the target of interest in fixed samples. A secondary antibody conjugated to horseradish peroxidase recognises the primary antibody, and, with the addition of biotin phenol and hydrogen peroxide, facilitates the rapid deposition of biotin onto proteins within the vicinity of the antibody complex [2]. Biotinylated proteins are subsequently identified following reverse cross-linking, homogenisation and a streptavidin-conjugated bead pull-down. To identify the isolated proteins, an on-bead trypsin digest is conducted before tryptic peptides are analysed by mass spectrometry.

Using BAR in fixed, post-mortem brain tissue from PSP patients, we identified several known aggregate components found in Tauopathies including Synapsin 1, Syntaxin-1B, beta-synuclein and gamma-enolase. Our data also identified confidently assigned phosphorylation sites on Tau that are associated with Tauopathies. Together these data validate our approach for rapidly revealing the aggregate components of Tauopathies whilst also identifying many novel components that may provide valuable insight into disease mechanisms upon careful validation.

1. Ross, C. A., Poirier, M. A., Protein aggregation and neurodegenerative disease. Nature medicine 2004, 10 Suppl, S10-17.
2. Bar, D. Z., Atkatsh, K., Tavarez, U., Erdos, M. R., et al., Biotinylation by antibody recognition-a method for proximity labeling. Nature methods 2018, 15, 127-133.

12

Splicing up your sex life

Mark A Baker[1], Jacob K Netherton[2, 1], Rachel A Ogle[1]

1. Environmental and Life Science, The University of Newcastle, Callaghan, NSW, Australia

2. Genea, Newcastle, NSW, Australia

Male infertility is a very common condition, with reports suggesting that one in 15-20 men of reproductive age are affected. Understanding why or how men produce defective sperm is a question that has remained elusive. We have used proteomic screens to identify mechanisms responsible for building defective sperm in men. Significantly, we have found regulators of alternate splicing appear to be a major key; being more abundant within infertile spermatozoa. To understand this, we overexpressed specific alternate-splicing regulators within *Drosophilia.* Amazingly, our data show that sperm overexpression RNA-splicing regulators produced typical patterns of “male-factor” infertility, including (i) decreased amounts of sperm production, (ii) head morphology defects and (iii) poor sperm motility. Furthermore, fertility data demonstrate changes to alternate splicing have dramatic consequence. Fly strains ranged from completely infertile to extremely subfertile. This data strongly suggest that aberrant alternate splicing is likely to play a major role when it comes to the production of poor quality spermatozoa, and male factor infertility.

13

The role of ubiquitylation on the phagosome: a tale of inflammatory responses and vesicle trafficking

Tiaan Heunis[1], Orsolya Bilkei-Gorzo[1], Anetta Härtlova[1], José Marin Rubio[1], Julien Peltier[1], Matthias Trost[1]

1. Newcastle University, Newcastle-upon-Tyne, NOT US OR CANADA, United Kingdom

Phagocytosis is an evolutionarily conserved key process required for innate immunity and homeostasis. During phagocytosis, particles are internalised into a *de novo*, membranous organelle, the phagosome, which fuses with early and late endosomes and, finally, lysosomes in a process called phagosome maturation. Recently, we have shown that proinflammatory activation of macrophages by Interferon-gamma (IFN-ɣ) greatly changed phagosome functions. As phagosomes are enriched in polyubiquitylation, which is further enhanced by IFN-ɣ, we wanted to explore the role of ubiquitylation on phagosome functions. We applied a targeted mass spectrometry approach by which we quantified ubiquitin chain linkage peptides from phagosome samples. All chain linkages apart from linear chains are present on phagosomes and IFN-ɣ activation enhanced K11, K48 and K63 chains significantly. In order to identify the molecular function of this polyubiquitylation, we characterized the ubiquitinome of phagosomes of IFN-γ activated macrophages and can demonstrate that ubiquitylation is preferentially attached to proteins involved in vesicle trafficking, thereby delaying fusion with late endosomes and lysosomes. We further show that phagosomal recruitment of the E3 ligase RNF115 is enhanced upon IFN-γ stimulation, which is responsible for most of the increase of K63 phagosomal polyubiquitylation. Loss of RNF115 also affected proinflammatory cytokine production and tissue damage during *in vivo* infection with *S. aureus*. Detailed understanding of phagosomal ubiquitylation could not only increase our understanding of vesicle trafficking but could also serve as possible targets for antibacterial host-directed therapies.

14

Proteogenomic characterization of einkorn wheat cultivars with low celiac disease response

Angela Juhasz[1], Mitchell Nye-Wood[1], Zsofia Birinyi[2], Gyongyver Gell[2], Michelle Colgrave[1]

1. Edith Cowan University, Joondalup, WESTERN AUSTRALIA, Australia

2. Applied Genomics, Centre for Agricultural Research, Martonvasar, Hungary

Wild wheat species and wheat genome donors represent an important source in breeding wheat with altered gluten content and composition. Previously, nearly 200 wild and cultivated einkorn accessions have been characterized. Gluten protein content has been measured using gluten specific G12 and R5 monoclonal antibodies in commercial sandwich ELISA methods. Genotypes with significantly lower gluten (below 500 ppm) content compared to bread wheat have also been analysed for gluten-specific IgA and IgG antibody response using human blood sera collected from celiac diseases (CD) patients. Altogether four genotypes have been identified as potentially suitable for low gluten breeding programs of which two cultivars showed relatively stable expression pattern across environments. These lines were grown, harvested and milled in a confirmed gluten-free environment.

In this study we aim to explore the differences at the genetic and protein levels of these genotypes in order to understand the variation in gluten protein composition and to explain the potentially low CD response. Grain proteomes of seven *Triticum monococcum* cultivars and one bread wheat were compared using information dependent data acquisition and SWATH-MS analysis. Transcriptome data collected from developing seeds of the two low CD responsive genotypes at 20 days post-anthesis was used to relate the cultivar specific allelic information to the expressed protein characteristics. Although a significant variation was observed in the detected number of gluten proteins when *T. monococcum* accessions were compared to the bread wheat cultivar, there was no significant variation in the number of gluten proteins within the *T. monococcum* cultivars. Mapping of the RNAseq read data confirmed the presence of the major gluten protein genes, however a significant variation was observed in their expressed amounts. In this presentation, we provide an insight in the unique genetic characteristics of these *T. monococcum* accessions by exploring the expression differences observed in the gluten proteins and their regulating genetic factors.

15

Multi-omic analysis and functional validation of *Campylobacter jejuni* grown with intestinal bile salts

Lok Man[2, 1], Joel A. Cain[2, 1], William Klare[2, 1], Nestor Solis[2], Pamela Soh[2], Zeynep Sumer-Bayraktar[2, 1], Paula Niewold[3, 1], Ashleigh Dale[2, 1], Stuart Cordwell[2, 3, 1, 4]

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

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

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

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

Campylobacter jejuni is the leading cause of bacterial gastroenteritis in the developed world, with infection is predominantly caused by the consumption of undercooked poultry. *C. jejuni* exists mainly as a commensal within the intestines of chickens, but is pathogenic in humans. The presence of bile salts in the human intestinal environment play an important role as antimicrobials, and gut pathogens such as *C. jejuni* often respond strongly to compounds such as deoxycholate, adapting their protein synthesis in an attempt to counteract the toxic effects of these bile salts. As a microaerophilic human pathogen, *C. jejuni*, particularly the iron-sulfur clusters found in many of its key enzymes, is susceptible to oxidative stresses posed by its environment, and as such has evolved many proteins dedicated to protecting it against these stresses. In this study, we characterised changes in intracellular protein and metabolite levels of *C. jejuni* when grown under 0.1% sodium deoxycholate, then validated select protein and metabolite changes using functional assays to describe the altered phenotype induced by exposure to deoxycholate. The results showed a marked decrease in abundance of proteins and metabolites that have known functions in protection against oxidative stress, and a marked increased in abundance of various membrane- and transport-related proteins. Functional validation assays confirmed that growth under deoxycholate greatly sensitised cells to oxidative damage, with cells exposed to the bile salt more than 20 times more susceptible to hydrogen peroxide treatment. This was a specific effect, as deoxycholate did not result in an overall weakened phenotype, with growth in the bile salt increasing adherence and invasiveness of mammalian intestinal cells, motility, and resistance to several antibiotics, consistent with proteomics showing increased abundance of drug transport proteins.

16

Proteomics tools for medicinal cannabis

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

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

Cannabis sativa has been dubbed “the plant of the thousand and one molecules” owing to its propensity to produce a plethora of phytochemicals with myriad biological activities as well as fibrous components. Out of the 500 compounds that have been described thus far, more than 90 are phytocannabinoids including cannabidiolic acid (CBDA) and delta 9-tetrahydrocannabinolicacid (THCA) accumulating within the trichomes of mature buds. One of the key phytocannabinoid-related enzymes is olivetolic acid (OLA) cyclase (OAC) acting in succession to convert hexanoyl-CoA into OLA.

Cannabis is a controlled substance and until recently was illegal in many jurisdictions. The recent revised legislation on medicinal cannabis has triggered a surge of medical and clinical research studies evaluating the effect of major cannabis components on human health. Medicinal cannabis is used to relieve the symptoms of certain medical conditions, such as epilepsy.

The state of Victoria in Australia was the first jurisdiction to legalise access to medicinal cannabis under the Medicinal Cannabis Act in 2016. In this context, Agriculture Victoria Research (AVR) have controlled access to medicinal cannabis material grown to full maturity in the state-of-the-art Victorian government medicinal cannabis cultivation facility.

A comprehensive systems biology approach including genome sequencing, transcriptomics, proteomics and metabolomics was undertaken by AVR to fully characterise the various medicinal cannabis cultivars developed in house.

In this study, we present three complementary proteomics strategies, namely bottom-up proteomics (BUP) through the use of the most commonly used protease trypsin[1], middle-down proteomics (MDP) by exploiting alternative orthogonal proteases[2], and top-down proteomics (TDP) which required innovative data mining methods and revealed numerous proteoforms[3]. We use OAC to illustrate the power of proteomics in cannabis research, with respect to sequence coverage and post-translational modifications (PTMs).

1. Vincent, D.; Rochfort, S.; Spangenberg, G. Optimisation of protein extraction from medicinal cannabis mature buds for bottom-up proteomics. Molecules 2019, 24, 659 doi:10.3390/molecules24040659.
2. Vincent, D.; Ezernieks, V.; Rochfort, S.; Spangenberg, G. A multiple protease strategy to optimise shotgun proteomics of medicinal cannabis mature buds. Submitted.
3. Vincent, D.; Binos, S.; Rochfort, S.; Spangenberg, G. Top-down proteomics of medicinal cannabis. Proteomes 2019, 7, 33, doi:10.3390/proteomes7040033.

17

DIGESTIVE PHYSIOLOGY AND PROTEOMICS FOR CONTROLLING AN INVASIVE PEST

Sophia Escobar-Correas[1, 2], Omar Mendoza-Porras[1], Israel A Vega[2], Michelle Colgrave[1]

1. CSIRO, St.Lucia, QUEENSLAND, Australia

2. Laboratorio de Fisiología , IHEM-CONICET-UNCuyo, Mendoza, Argentina

This work focuses on the study of the invasive Golden Apple Snail (GAS), *Pomacea canaliculata*, identified as one of the "100 world's worst invasive alien species". We characterised the food digestion strategies of this pest through an analysis of the enzymes present along its digestive system and within its secreted fluids. We also explored the participation of an intracellular endosymbiont lodged in the epithelial cells of the digestive gland of the snail.

Adult snails were kept in aquariums under controlled conditions and with synchronized feeding for 48 hours before sampling. Tissues and endosymbionts were sampled. Liquid chromatography coupled to tandem mass spectrometry was applied to define the occurrence and origin of digestive enzymes along the digestive tract of *P. canaliculata.*

The proteome (proteins ≥ 95% confidence) was obtained from salivary glands (3,796 proteins), content of the crop (907 proteins), digestive gland (3,518 proteins), content of the style sac (792 proteins) and intestine (2,495 proteins). Later, 144 peptides from 55 peptidases, 21 peptides of 12 lipases and 250 peptides of 81 glycosidases were quantified using multiple reaction monitoring mass spectrometry. Moreover, we found 67 proteins and quantified 85 peptides of the endosymbiont. We identify two peptidases and two glycosidases of possible bacterial origin.

In synthesis, we describe a wide diversity of digestive enzymes along the digestive tract of *P. canaliculata*. These findings open a new field of study of these enzymes as targets for the control and eradication of this invasive species.

18

LC-MS/MS-based multi-omics approach to elucidate developmental adaptations in the parasitic nematode *Haemonchus*

Tao Wang[1], Guangxu Ma[1], Pasi Korhonen[1], Ching-Seng Ang [2], Shuai Nie[2], Rong Xu[3], Anson Koehler [1], Richard Simpson [3], David Greening [3], Gavin Reid [4, 5, 6], Nicholas Williamson [2], Robin Gasser [1]

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

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

3. Department of Biochemistry, La Trobe University, Bundoora, Victoria , 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

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

Parasitic nematodes of humans, animals and plants have a major adverse impact on global health and agricultural production worldwide. The ability of such worms to adapt to changing environments is essential for their survival inside and outside of hosts. Advanced nucleic acid sequencing and bioinformatic technologies have enabled an unprecedented number of worm genomes to be decoded, which is starting to improve our understanding of the parasites at the molecular level. However, there is still a lack of knowledge of host-parasite relationships and parasitism. Here, using *Haemonchus contortus* - one of the most important parasitic nematodes of livestock animals worldwide - as a model, we defined the somatic and excretory/secretory proteomes as well as the phosphoproteome and lipidome of this nematode to explore molecular alterations during its development via mass spectrometry-based tools. By integrating these -omic datasets, we have been able to elucidate adaptations, including nutrient acquisition and metabolism, that take place during this nematode's transition from the free-living stage in the environment to the parasitic stage inside the host animal. Understanding how worms orchestrate the process of parasitism at the molecular level could assist significantly in the development of molecular intervention strategies against parasitic nematodes.

19

Cellular glycoproteome and proteome during influenza infection

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

1. School of Chemistry and Molecular Biosciences, 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 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 during influenza infection. For the proteomic analyses, we implemented a SWATH approach to quantify changes in proteins from subcellular fractions. This revealed increased membrane lipid metabolism, clathrin-dependent endocytosis and cytokine signaling during viral infection. For the glycoproteomic analyses we enriched glycopeptides by HILIC from subcellular fractions and investigated the monosaccharide composition of the attached glycans in a site-specific manner. The work presented herein highlights site-specific changes in host protein glycosylation, including decreased sialylation and increased fucosylation in influenza infected cells.

20

"RNA regulons" as potential drivers of *Giardia duodenalis* stage transition

Balu Balan[1, 2], Samantha Emery-Corbin[1], Jarrod Sandow[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

Parasitic protists rapidly adapt to environment changes to survive and infect new hosts. These adaptations largely occur in absence of cell division and necessitate massive reconfigurations of transcriptional and expressional behaviors. These essential changes occur within minutes of invasion, such that they appear pre-programmed.

Pre-programming ('cell-fating') has been explored heavily in eukaryotic cells, including developmental biology and stem-cell differentiation. Post-transcriptional regulation (PTR), as mediated through translational repression, contributes to cell-fating by holding transcripts in stasis for later translation. RNA-binding proteins (RBPs) are essential in PTR, regulating multiple mRNAs through high order combinations known as "RNA regulons".

*Giardia duodenalis*is a prevalent gastrointestinal parasite causing over 200 million cases of diarrhoeal disease. Understanding PTR mechanisms in *Giardia* will illuminate new systems to inhibit critical aspects of its infection biology, host interaction and resistance to frontline drugs. Further, *Giardia* provides a model to 'close the gap' between higher eukaryotes and deep-branching species to ascertain evolutionary origins of eukaryotic PTR mechanisms.

Our multi-omics study explored multiple dimensions of RBP repertoires in *Giardia*. We bioinformatically curated the currently undefined *Giardia*'RBPome' and mapped its transcriptomic and proteomics kinetics across stage transition. Towards this, we performed the first quantitative deep proteome map of the entire *Giardia* stage transition. We complimented this with interactome capture proteomic characterization of poly-A mRNAs bound proteins *Giardia*through stage transition *in vitro*, the first characterization of this kind in any protist. From these analyses we have identified a eukaryote conserved RBP which contain intrinsic disordered regions responsible for phase separating behavior in model eukaryotes. We plan to further explore the role of this RBP as regulator of translational repression in *Giardia* infection biology. Together, we provide the first multi-platform, multi-omic reference map of RBPs in *Giardia* towards understanding the role and evolution of "RNA regulons" in eukaryotes.

21

Identification of novel *N*-glycosylation sites from the bacterium *Campylobacter jejuni*

Joel Cain[2, 1], Stuart Cordwell[2, 3, 1, 4]

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

Campylobacter jejuni is the predominant cause of acute gastroenteritis in the developed world. It's ability to perform N-glycosylation is a key trait which contributes to its pathogenicity as well as establishing a commensal relationship within avian guts (the primary vector for infection). Recently we established that loss of N-glycosylation was associated with an array of phenotypes including aberrant chemotactic responses, altered respiratory preferences, changes in cell morphology and increased

susceptibility towards temperature and osmotic stress. Given the lack of concrete connections between known targets of this PTM and the affected phenotypic traits, we looked to broaden the list of known N-glycoproteins within *C. jejuni* with the aim of identifying putative causative agents. Of the known 131 modification sites, we were able to unequivocally demonstrate occupation of ~80% of known N-glycopeptides in *C. jejuni*. We were also able to identify 54 novel modification sites, including 15 novel N-glycoproteins. We also looked to assess biological occupancy of novel and previously identified sites from relative quantitation of N-sequon containing peptides between wild-type and an N-glycosylation incompetent C. jejuni strain - an oligosaccharyltransferase mutant, Δ*pglB.*

22

Understanding glycopeptide antibiotic synthesis by mass spectrometric analysis of polycyclic peptides

Robert JA Goode[1, 2], Milda Kaniusaite[2, 3], Edward A Marschall[2, 3], Yongwei Zhao[2, 3], Anja Greule[2, 3], Julien Tailhades[2, 3], Max J Cryle[2, 3], Ralf B Schittenhelm[1, 2]

1. Monash Proteomics and Metabolomics Facility, Monash University, Clayton, Victoria, Australia

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

3. EMBL Australia, Monash University, Clayton, Victoria, Australia

Glycopeptide antibiotics are considered the last line of defence against several bacterial infections. A thorough understanding of their biosynthesis and the ability to interrogate and manipulate their structure is therefore critical to overcome antibiotic resistance and to further develop efficient antibiotics. Glycopeptide antibiotics typically consist of a core of seven non-standard amino acids generated through non-ribosomal peptide synthesis with up to four side chain crosslinks that are introduced by cytochrome P450 enzymes. Mass spectrometric analysis of these samples is challenging due to the lack of available software to handle the non-standard amino acids as well as the complex crosslinking. Here we present the results of different core peptides and enzyme combinations to understand the roles and specificities of these enzymes in the crosslinking process, and also highlight some of the challenges involved in analysing MSMS spectra from these samples.

23

Characterising insult-induced protein-protein crosslinks formed during food processing

Hannah McKerchar[1, 2, 3, 4], Renwick Dobson[1, 3, 4], Stefan Clerens[2, 3, 4], Jolon Dyer[2, 3], Juliet Gerrard[5, 3]

1. School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

2. AgResearch, Lincoln, New Zealand

3. Riddet Institute, Palmerston North, New Zealand

4. Biomolecular Interaction Centre, Christchurch, New Zealand

5. School of Biological Sciences and School of Chemical Sciences, University of Auckland, Auckland, New Zealand

During food processing insult-induced crosslinks form between proteins, which profoundly influence the nutritional value and properties of food. One crosslink that forms with heating and basic pH is lysinoalanine. Despite the adverse effects of lysinoalanine, its formation and biological fate are not well understood. Directly mapping insult-induced protein-protein crosslinks has many challenges.

Using mass spectrometry, we aim to identify a diagnostic fragmentation pattern of lysinoalanine to help to map its location in food-related samples.

Simple peptide models of a lysinoalanine, disulfide and lanthionine crosslinks have been isolated and characterised from MALDI-TOF/TOF fragmentation spectra. The model protein β-lactoglobulin was heated at pH 12 to form crosslinks and analysed using LC-QTOF-MS/MS. The locations of crosslinks in the protein can be identified by comparing acquired data to the theoretical masses of crosslinked peptides.

The crosslink models suggest the lysinoalanine preferentially fragments at the α-carbon and β-carbon. Trends in the model crosslinks show increasing the length of heating and pH both results in more lysinoalanine and lanthionine crosslinks forming and fewer disulfide bonds form. Increasing these conditions also results in lysinoalanine forming in favour of lanthionine. Identifying crosslinks in the milk protein, β-lactoglobulin, formed as a result of from thermal and alkaline treatment is been undertaken through LC-QTOF-MS. Characterisation of crosslinks in the protein will help validate the fragmentation pattern and trends identified in the peptide crosslink models.

We have identified a putative fragmentation pattern to help map the location of crosslinks within a milk protein. Knowing the location of the crosslinks gives the opportunity for these structural features to be included in digestion models. Varying reaction conditions has enabled us to investigate a proteins' susceptibility to from crosslinks. This guides us on how altering food preparation procedures can influences lysinoalanine formation and the opportunity to increase foods' nutritional value.

24

Multi-omic analysis reveals the pro-inflammatory cytokine IFNγ modulates the immunopeptidome of triple negative breast cancer cells.

Gabriel Goncalves[1], Kerry Mullan[1], Divya Duscharla[1], Rochelle Ayala [1], Ethan Passantino[1], Nicole Mifsud[1], Nathan Croft[1], Pouya Faridi[1], Anthony Purcell[1]

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

The heterogeneous nature of Triple-Negative Breast Cancer (TNBC) has highlighted the need for novel approaches in treatment. Recent studies demonstrate the feasibility of a peptide vaccination approach to induce T-cell mediated killing of cancerous cells. To identify potential T-cell targets we used the TNBC model cell line MDA-MB-231 and examined cell surface peptides (immunopeptidome) being presented by human leukocyte antigen (HLA) complexes after treatment with the pro-inflammatory cytokine interferon-γ (IFNγ) and related these findings to changes in the cellular transcriptome and proteome. Using high resolution mass spectrometry, we identified a total of 82,915 peptides from 9,553 source proteins were presented across HLA-I and HLA-II alleles and observed considerable remoulding of the immunopeptidome upon cytokine stimulation. We then examined the correlation between the 10,248 quantified transcripts and 6,784 quantified proteins following IFNγ treatment and found 229 differentially expressed proteins many of which were involved in the pathway of antigen presentation. These results highlighted a high degree of plasticity in the immunopeptidome caused by cytokine stimulation of TNBC. Of note, IFNγ increased the diversity and abundance of the peptide repertoire as a result of changes in the antigen presentation machinery rather than changes in transcript and protein abundance. This suggests that under pro-inflammatory conditions a greater variety of potential T cell targets are unveiled to the immune system. This has important implications for the development of personalised cancer vaccination strategies.

26

New tools to study protein ubiquitination

David Komander[1]

1. Ubiquitin Signalling Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne

Protein ubiquitination is ubiquitous, as its name implies, and is emerging as a key regulator of protein homeostasis and cell signalling. Deregulation of protein ubiquitination is involved in a variety of human diseases, including cancer and neurodegenerative and inflammatory conditions. The last decade has revealed a plethora of distinct ubiquitin signals. My lab tries to understand the complex 'ubiquitin code' by focussing on the different ubiquitin signals themselves, in order to eventually link the specificity in the system back to physiological functions.
Mass spectrometry has taken centre stage to access the complexity of ubiquitin signals. Tryptic digest of a ubiquitinated protein decorates ubiquitinated residues with a 114 Da GlyGly signature to aid site identification, however much of the complexity that originates from polyubiquitin architecture, or co-modifications, is being lost.
We recently reported Ub-clipping, a new methodology exploiting a viral leader protease that cleaves ubiquitin (and the ubiquitin-like modifier ISG15) from proteins while preserving the GlyGly modification on the intact protein (Swatek et al, Nature 2019). Ub-clipping enables new workflows, simplifies Ub site identification and offers unprecedented insights into the ubiquitin signals used in signalling pathways.

27

Deciphering irreversible cysteine redox post-translational modifications in myocardial ischemia / reperfusion injury

Stuart Cordwell[1], Alexander Rookyard[1]

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

Cysteine (Cys) is a crucial site for redox post-translational modification (PTM). Some Cys remain modified for the life of the protein (e.g. disulfides), while others vary in PTM in response to the redox environment. There are many different Cys redox PTM, including nitrosylation, glutathionylation and acylation, and these are considered 'reversible' as cells contain an array of enzymes that reduce these modifications (e.g. thioredoxins, glutaredoxin) back to their free thiol form. In conditions where the oxidant:antioxidant balance in the cell is in favour of oxidants, thiols can be over-oxidised via Cys-sulfenic acid (Cys-SOH) to either Cys sulfinic or sulfonic acids (Cys-SO2H/Cys-SO3H), which are considered enzymatically 'irreversible' and mark the protein for unfolding and ultimately, degradation. We describe a method for peptide-based, enrichment, identification and quantification of Cys-SO2H/Cys-SO3H based on negative selection by strong cation exchange (SCX) chromatography and positive selection by hydrophilic interaction liquid chromatography (HILIC), coupled to parallel reaction monitoring-mass spectrometry (PRM-MS). We identified >300 irreversibly modified Cys sites in an animal model of myocardial ischemia / reperfusion (I/R) injury, and quantified them in comparison to non-ischemic time controls and I/R in the presence of an antioxidant intervention mediated by N-mercaptopropionylglycine (MPG). Functional analysis showed that Cys sites oxidised during I/R and 'protected' by MPG were largely associated with metabolic processes, particularly the tricarboxylic acid cycle, and these data correlated with perturbations in TCA cycle flux as assayed by LC-MS/MS-based metabolomics. Cys sites identified here are targets of reactive oxygen species (ROS) that contribute to protein dysfunction during I/R.

28

HNF4A sumoylation is regulated by fasting.

Dylan J Harney[1], Luke Hatchwell[1], Michelle Cielesh[1], Kieren Young[1], Yen Chin C Koay[2], John F O'Sullivan[2, 3], Mark Larance[1]

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

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

3. Department of Cardiology, Royal Prince Alfred Hospital, Camperdown, NSW, Australia

Intermittent fasting (IF) increases lifespan and decreases metabolic disease phenotypes, but the mechanisms mediating these effects are not fully characterized. Our laboratory has previously identified that hepatic nuclear factor 4A (HNF4A) has reduced activity during IF using unbiased liver proteomics and gene set enrichment analysis. Alongside the change in HNF4A activity, EODF caused comprehensive induction of *de novo* lipogenesis (DNL) and cholesterol biosynthesis pathways. HNF4A is known to regulate cholesterol metabolism, VLDL-particle formation and bile acid biosynthetic pathways that are related to some intermittent fasting phenotypes. Recently, a study identified SUMO2/3 modification of HNF4A at K365 which led to increased transcriptional activity of HNF4A. In this study, we have demonstrated that HNF4A sumoylation is greatly decreased by acute fasting, but not by intermittent fasting in mouse liver extracts. Furthermore, we demonstrated that SUMO-HNF4A had altered sub-cellular localisation compared to unmodified HNF4A, with the sumoylated form moving from the nucleus to the cytosol. Together, this suggests sumoylation of HNF4a may provide a regulatory mechanism by which HNF4a contributes to the intermittent fasting phenotype. We have identified two other transcription factors, peroxisome proliferation factor a (PPARa) and sterol regulatory element binding protein 1c (SREBP1c), which were also significantly regulated during EODF. Both these proteins have known inhibitory interactions with HNF4A and we will test whether these interactions help to mediate HNF4A inhibition through sumoylation. We will also examine whether sumoylation of HNF4A is necessary and sufficient for HNF4A nuclear-cytoplasmic trafficking.

29

Protein TAILS tell remarkable tales: Positional proteomics reveals diverse N-Terminomes and proteolytic landscapes in disease

Chris Overall[1]

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

Identification of protein terminal peptides provides key information on protein stability and function. Our degradomics methods enrich and annotate terminomes—Terminal Amine Isotopic Labelling of Substrates (TAILS, *Kleifeld etal 2010)*, Amino Terminal Orientated Mass Spectrometry (ATOMS), Proteome-wide Identification of Protease Cleavage Specificity (PICS, *Schilling_&_Overall 2008*), LysargiNase, the new protease for proteomics we discovered (Huesgen *etal* 2015), and our N and C-termini database version 4.0 TopFIND (http://clipserve.clip.ubc.ca/topfind) utilised in our studies reveal widespread truncation and generation of termini in normal and diseased tissues.

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

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

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

By TAILS we previously reported that the innate immune cell MMPs orchestrate leukocyte chemotaxis by cleavage of most human chemokines. More recently IFNα and IFNγ, and complement proteins were discovered as new MMP substrates. MMP processing of these bioactive substrates dampens inflammation essential for terminating inflammatory responses and inspiring new potential novel precision therapies to modulate protease enzyme activities in diseases due to protease deregulation.

30

Challenges in Visualization and Analysis of Mass Spectrometry Imaging Experiments

Kylie Bemis[1]

1. Northeastern University, Boston, MASSACHUSSETS, United States

Mass spectrometry imaging (MSI) allows researchers to inspect the spatial distribution of molecular ions in a sample by repeatedly collecting mass spectra from spatial locations across its surface, producing hundreds or thousands of molecular ion images. While MSI promises exciting new insights in biomedical applications such as cancer research, the complexity of MSI datasets poses substantial challenges for visualization and statistical analysis. Furthermore, Improvements in instrumentation have led to rapid increases in mass and spatial resolution, producing larger datasets and larger file sizes, further compounding the analytic challenges.

These challenges must be met by an evolution in methods for visualization and statistical analysis of MSI experiments. However, despite the proliferation of machine learning algorithms and ad hoc data analysis and visualization tools for MSI, the development and adoption of appropriate statistical methods and reproducible experimental design has lagged behind. Many experiments with

otherwise high-quality data still suffer from inadequate sample sizes or flawed experimental design, and few methods for statistical analysis exist.

We will break down the major statistical challenges of MSI analysis in three categories of analytic goals: (1) segmentation, (2) classification, and (3) class comparison. For segmentation, we will demonstrate spatial shrunken centroids, which performs simultaneous segmentation and selection of important ions. For classification, we present our current work on using multiple-instance learning in the presence of uncertain class labels. For class comparison, we present a novel single-ion segmentation method that can be used as input to a statistical testing procedure. These methods are implemented in the open-source R package Cardinal, which provides a full workflow of data import, pre-processing, visualization, and statistical analysis for MSI experiments.

31

MassOmics: An R package of a cross-platform data processing pipeline for large-scale GC-MS untargeted metabolomics datasets.

George GUO[1], Elizabeth McKenzie[1], Beatrix Jones[1], Erica Zarate[1], Jamie V. de Seymour[2], Silas G. Villas-Bôas[1], Ting-Li Han[3]

1. University of auckland, Auckland, NEW ZEALAND, New Zealand

2. Massey University, Auckland, New Zealand

3. Department of Obstetrics and Gynaecology, the First Affiliated Hospital of Chongqing Medical University, CHONGQING, CHINA

Large-scale gas chromatography-mass spectrometry (GC-MS) based untargeted metabolomics, where hundreds or thousands of samples are analysed over a period of weeks or months, has specific challenges. These include variation in instrument performance, signal intensity loss due to column ageing, the build-up of contaminants in the ion source, and sample handling variability. In addition, the computational challenges of compound identification are intensified when dealing with a large number of samples. A data processing software package to address these problems is required. Our software, MassOmics, is designed to bring together R packages and scripts for GC-MS data processing to rapidly integrate and annotate peaks in large-scale datasets, all within a graphical user interface. This package also provides identification of background contaminants, data scaling and transformation, various batch effect removal methods, machine learning-powered grouping of metabolites, and metabolite importance analysis. With these functions, MassOmics can parse and summarize library batch search results from ChemStation and MassHunter, and produce an integration output of GC-MS dataset, which is compatible with various downstream statistical and metabolic pathway analysis tools. The module-based design and intermediate data transferring approach enable MassOmics to work with data integration platforms such as KNIME to generate an adaptive and customizable processing workflow. The MassOmics package is designed for researchers with little experience using R, and substantially improves GC-MS data extraction efficiency and accuracy, as well as reducing the time required for manual checking and re-integration.

32

MHCpLogics: a machine learning-based tool for unsupervised data visualisation and cluster analysis of immunopeptidomes

Mohammad Shahbazy[1], Pouya Faridi[1], Sri H Ramarathinam[1], Nathan P Croft[1], Anthony W Purcell[1]

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

Background The *major histocompatibility complex* encodes human leukocyte antigens (HLAs) in humans, which bind and present intracellular peptides that are then displayed on the cell surface for recognition by T cells. The repertoires of peptides presented by HLA are termed immunopeptidomes. The highly-polymorphic nature of HLA confers allele-specific differences in the sequence properties of bound ligands, designated as peptide-binding motifs. Herein, we developed MHCpLogics as a machine learning-based tool for clustering analysis, amino acid-based feature selection, and sequence motif visualisation of peptides to discover landscapes in human immunopeptidomes.

Methodology We used new experimental and previously published immunopeptidomics data from mono- and multi-allelic cell lines to cover a wide range of HLA alleles. Each peptide sequence was numerically encoded to allow subsequent machine learning analysis *via* code programming in MATLAB/Python.

Results The MHCpLogics tool provides dimensionality data reduction, exploratory analysis, visualisation, and clustering analysis of peptide data alongside exporting sequence motif logos. In the known HLA datasets, the tool showed clear deconvolution of motif clusters, highlighting the restricted nature of motifs from mono-allelic immunopeptidomes and yet readily segregable clusters from multi-allelic data. Across all data, contaminant sequences could be easily identified, allowing exclusion from further analysis. Visualisation modalities grant users the features/abilities to inspect clusters down to individual peptides and examine broader/higher-level patterns, as well as density visualisation and heatmap analysis. Additional statistical outputs provide information, e.g., the proportion of HLA-binders, hierarchical cluster analysis dendrograms, and amino acid frequencies.

Conclusion MHCpLogics can deconvolute large mass spectrometry-based immunopeptidome data, allowing interrogation of clusters/sub-clusters of peptide motifs, with the representation of data in a wide array of visualisation options and the ability to export peptide sequence lists. The tool will be an essential asset to the immunology community, allowing easy and rapid inspection of immunopeptidomes and, ultimately, the identification of HLA alleles present in unknown samples.

33

OmixLitMiner - Tool for fast evaluation of knowledge and importance of regulated individual proteins derived from differential proteomics

Hartmut Schlüter[1], Pascal Steffen[2], Jemma Wu[3], Vijay Raghunath[4], Hannah Voß[1], Mark P Molloy[5]

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. Department of Molecular Sciences, Macquarie University, Sydney, , New South Wales, Australia

4. Sydney Informatics Hub, The University of Sydney, Sydney, New South Wales, Australia

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

Differential proteomics studies today are usually often yielding hundreds of down- or up-regulated proteins which are associated with a defined perturbation. Because of the large number of identified regulated proteins, it is time consuming to answer the question, which proteins are the most important ones with respect to the scientific question and to estimate which of the regulated proteins are most promising for gaining new knowledge. For speeding up the process of evaluation of the importance of the proteins and for highlighting proteins yet unknown in the context of the scientific question, a text mining tool was developed, termed OmixLitMiner. Lists of accession numbers of the identified regulated proteins are uploaded into the tool (based on a script written in the computational language R), which automatically is searching for synonyms in UniProt by transferring them into PubMed and combining them with a keyword characterizing the scientific question. The PubMed search is repeated adding step by step several filters (filter “Title”: the synonyms must occur in the title; filter “Review”: Only hits are searched for representing review papers) and listing the hits after each step of search. After the searches are finished, the tool will summarize the results in a word cloud, statistical plots, MeSH-based clusters and assignment into one of the three categories: Category-1: “well studied”; Category-2: “not well studied”; Category-3: “Not known”. Category-1 proteins can help to validate the experiment since these proteins have been mentioned in the titles of reviews, which are also related to the scientific question of the study. Proteins, which are categorized as Category-2 and Category-3 proteins have not been reported in relationship to the topic of the study. Thus, these proteins may give access to new hypothesis and thereby for new knowledge.

34

CiteFuse enables multi-modal analysis of CITE-seq data

Hani Jieun Kim[1, 2], Yingxin Lin[1], Thomas A Geddes[1], Jean Yang[1], Pengyi Yang[1, 2]

1. The University of Sydney, Camperdown, NSW, Australia

2. Children's Medical Research Institute, Westmead, NSW, Australia

Multi-modality profiling of single cells represents one of the latest technological advancements in molecular biology. Among various single-cell multi-modality strategies, cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) allows simultaneous quantification of two distinct species: RNA and surface marker proteins (ADT). Here, we introduce CiteFuse, a streamlined package consisting of a suite of tools for the pre-processing, modality integration, clustering, ADT evaluation, RNA-ADT network construction, differential expression analysis, and interactive web-based visualization of CITE-seq data. We show the integrative capacity of CiteFuse to fuse the two data types and its relative advantage against data generated from single modality profiling. Furthermore, we illustrate the pre-processing steps in CiteFuse and in particular a novel doublet detection method based on a combined index of cell hashing and transcriptome data. Collectively, we demonstrate the utility and effectiveness of CiteFuse for the integrative analysis of transcriptome and epitope profiles from CITE-seq data.

35

Stimulomics: large-scale mapping of signalling networks downstream of major cell surface receptors

Emily S Humphrey[2, 1], Philipp E Geyer[2], Elise J Needham[3, 4], Florian Gnad[5], Matthias Mann[2], Sean J Humphrey[2, 3, 4]

1. School of Medicine, The University of Notre Dame, Sydney, NSW, Australia

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

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

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

5. Department of Bioinformatics and Computational Biology, Cell Signaling Technology Inc, Danvers, MA, USA

Publish consent withheld

36

Precision phosphoproteomics reveals kinase targets enhancing muscle insulin sensitivity

Elise J. Needham[1], Janne R. Hingst[2], Benjamin L. Parker[1, 3], Sean J. Humphrey[1], David E. James[1, 4], Jørgen F.P. Wojtaszewski[2]

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

2. Section of Molecular Physiology, Department of Nutrition, Exercise, and Sports, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

3. Department of Physiology, University of Melbourne, Melbourne, VIC, Australia

4. Charles Perkins Centre, Sydney Medical School, The University of Sydney, Sydney, NSW, Australia

Exercise promotes health by engaging complex signalling pathways to promote insulin sensitivity, however the mechanisms by which this occurs remain unclear. Here we performed high-throughput phosphoproteomics of human skeletal muscle from both exercised and non-exercised legs following exercise recovery and a subsequent insulin treatment, quantifying 12,619 Class 1 phosphopeptides. Typically, inter-subject variation is considered a limitation of large-scale omics measurements. Our experimental design enabled repeated-measures analyses, facilitating the classification of >1,500 phosphopeptides regulated by exercise and/or insulin in the study. Moreover, we exploited the heterogeneity in subject responses to exercise and insulin treatment, to statistically identify phosphorylation sites associated with improved glucose uptake. In concert with the phosphoproteomics, we measured glucose uptake into the specific muscles that were biopsied in response to the hyperinsulinemic-euglycemic clamp. By measuring temporal changes in the phosphoproteome along with simultaneous subject-matched measures of muscle glucose uptake, we correlated signalling changes with muscle metabolism on an individual level. Insulin and exercise strongly potentiated the glucose uptake of certain subjects, while other subjects responded less effectively. Strikingly, we observed phosphosites that displayed identical heterogeneity among the subjects. Correlating phosphosites with glucose uptake across the individuals highlighted phosphosites on proteins known to be involved in glucose uptake, as well as potential novel players. This analysis also identified a kinase that has not previously been reported to promote insulin sensitivity as regulating a large proportion of these glucose uptake-associated phosphosites. By considering individual responses and linking these to glucose uptake, these data reveal a landscape of how exercise promotes insulin signalling, providing novel targets to therapeutically promote muscle insulin sensitivity.

37

Analytical guidelines for co-fractionation mass spectrometry obtained through global profiling of gold standard *Saccharomyces cerevisiae* protein complexes

Chi Nam Ignatius Pang[1], Daniel Weissberger[1], Sara Ballouz[2], Loic M Thibaut[3], Joseph R Gillis[2], Marc R Wilkins[1], Gene Hart-Smith[4]

1. School of Biotechnology and Biomolecular Sciences, UNSW, Sydney, NSW, Australia

2. Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Woodbury, NY, USA

3. School of Mathematics and Statistics, UNSW, Sydney, NSW, Australia

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

Co-fractionation mass spectrometry (CF-MS) is a method by which endogenous and unmanipulated protein complexes can be analysed on a broad scale in single experiments. CF-MS involves extensive biochemical fractionation of protein complexes using one or more non-denaturing chromatographic techniques (e.g. size exclusion chromatography (SEC)), followed by quantitative proteomics of each fraction. Subunits from the same intact complex will have highly correlated fractionation profiles.

Despite its demonstrated utility (1-2), best practice approaches for CF-MS remain undefined. Here we gain insight into how to best collect and interpret CF-MS data by benchmarking CF-MS datasets against gold standard complexes in *Saccharomyces cerevisiae*, one of the few organisms for which high proteome-coverage reference libraries of gold standard complexes exist.

By benchmarking experimental and modelled CF-MS datasets, we find that co-analysis of data from complementary biochemical fractionation methods (e.g. using Fisher's combined probability test) identifies complexes with greater efficiency than stand-alone biochemical fractionation. Systematic identification of gold standard complexes using 17 correlation metrics indicates that some metrics (e.g. Spearman correlation) are more effective than others (e.g. Mutual Information).

Many fractionation profiles that were unable to be benchmarked were nonetheless highly correlated, and thus possibly derived from novel complexes. Principal component analysis of gold standard and putative novel complexes indicated that novel complexes frequently elute in later SEC fractions, and are therefore often small. To test the effects of using orthogonal data (e.g. Gene Ontology) to assist in the prediction of these novel complexes, the Extending 'Guilt-by-Association' by Degree R package (3) was used. These analyses found that identifications of gold standard complexes are likely to benefit from the integration of GO data, whereas predictions of novel complexes are not. This suggests that orthogonal experimental validation (e.g. cross-linking mass spectrometry) may be required to validate novel complexes in CF-MS datasets.

1. Wan et al. 2015 Nature 525: 339-44
2. Heusel et al. 2019 Molecular Systems Biology 15(1)
3. Ballouz et al. 2017 Bioinformatics 33: 612-24

38

Lipidomic Mapping, Localization and Characterization in Tissues by MALDI- Imaging Mass Spectrometry

Anjaneyaswamy Ravipati[1], Sydney Liu Lau[1], Tzong-Tyng Hung[2], Brendan Lee[2], Carl A Power[2], Nicholas Proschogo[3], Mark J Raftery[1]

1. Bioanalytical Mass Spectrometry Facility, University of New South Wales, SYDNEY, NSW, Australia

2. Biological Resource Imaging Laboratory, University of New South Wales, SYDNEY, NSW, Australia

3. School of Chemistry, Faculty of Science, The University of Sydney, SYDNEY, NSW, Australia

Biological tissues constitute a complex array of molecular components including carbohydrates, lipids, proteins, and nucleic acids. Matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry (IMS) has proven to be a versatile technique in biomedical research to reveal distribution of broad scale of compounds ranging from metabolites to proteins within biological systems. In this work, we optimized the sample preparation protocol MALDI-IMS: cryosectioning of small fresh-frozen tissues, obtaining histology reference images, homogenous matrix deposition on tissues slices using sublimation, and data processing for the comparison of phospholipids on different tissues including mouse kidney, liver and spleen.

We demonstrate the identification and localization of different classes of phospholipids at high spatial resolution of 25 µm in positive ion mode. The lipid species were identified based on accurate mass and the structures of major phospholipids were characterized on tissue using the MALDI-IMS LIFT method. Our results showed that the kidney was enriched with the most complex and abundant phospholipids including phosphatidylcholine (PC) (32:0), PC (38:6), PC (16:0/18:1), PC (36:4), PC (34:2). On the other hand, spleen was predominantly composed of PC (32:0), while the liver with PC (34:2) and PC (16:0/18:1).

High-dimensional IMS data were visualized in lower dimension through unsupervised multivariate analysis approaches (SCiLS Lab) including image segmentation and Principal component analysis (PCA): using this approach morphological features of all tissues unambiguously distinguished. Further analysis of spleen tissues at region of interests (ROI) level using supervised approach (ClinProTools) revealed heme B is uniquely localized in red pulp, supporting the fact that splenic sinusoids are engorged with blood.

Our study outcomes emphasise that the diversity and complexity in lipid composition of each tissue type indicate structural and functional similarities / variations at cellular level. It also underscores the high potential of MALDI - IMS to unravel pathophysiological changes associated with diseases, discovery of novel biomarkers, pharmaceutical research and drug development using mouse models.

39

Multi-Omics Analysis of the Intermittent Fasting Response

Luke Hatchwell[1], Dylan J Harney[1], Michelle Cielesh[1], Kieran Young[1], Yen Chin Koay[1, 2], John F O'Sullivan[1, 2, 3], 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

3. Department of Cardiology, Royal Prince Alfred Hospital, Camperdown, NSW

Every-other-day-fasting (EODF) is an effective intervention for treatment of metabolic disease including improvements in liver health. But how the liver proteome is reprogrammed by EODF is currently unknown. Here, we employed EODF in mice and multi-Omics analysis to identify key regulated pathways. Many changes in the liver proteome and metabolome were distinct in EODF animals and not present after a single fasting bout. Key among these was the simultaneous induction by EODF of de novo lipogenesis and fatty-acid oxidation enzymes. Together with activation of oxidative stress response pathways, these changes would contribute to the improvements in glucose tolerance and lifespan after EODF. Enrichment analysis showed unexpected downregulation of HNF4A targets by EODF and we confirmed inhibition of HNF4A function. Suppressed HNF4A targets include bile synthetic enzymes and secreted proteins such as a1-antitrypsin and inflammatory factors, which reflect known EODF phenotypes and were confirmed by plasma proteome analysis. Interactive online access is provided to this data resource (larancelab.com/eodf), which provides a global view of fasting-induced mechanisms in mice.

40

Using genomics and proteomics to understand the antibiotic resistance capabilities of a bacterial pathogen

Stephanie E L Town[1], Matt Padula[1], Steven 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

Next-generation genomic sequencing has shown potential as a predictor of phenotype and for understanding genes thought to confer antimicrobial resistance (AMR) in bacterial communities. Despite the availability of genome sequencing and improvements to proteomic workflows enabling robust, sensitive and comprehensive discovery and quantitation of biomolecules, there are still little to no experimental studies evaluating how the bacterial proteome responds to antimicrobial challenge in the case of AMR.

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 evaluate how effectively genome sequences predict the phenotype by connecting AMR genes to gene end-products on the proteoform-level, which has yet to be experimentally shown using a systems biology approach.

Long- and short-read genomic sequencing and assembly was conducted in-house on multi-drug resistant *E. coli* isolates to confirm the presence of AMR-related genes. A shotgun LC/MS/MS proteomics pipeline measured proteome changes with and without antibiotic challenge. PEAKS Studio X, UniProt, and STRING databases was used to analyse data.

Several proteins related to AMR were found to be upregulated, despite no antibiotic challenge, including multi-drug resistance proteins. Additionally, several proteins previously annotated as hypothetical were observed.

Our research findings are one of the first to experimentally link an AMR-related gene to the AMR-related protein using a systems biology approach, providing evidence that the genotype and the phenotype do differ. Results highlight that more proteoform-level evidence is required to validate the insights made by genomic sequencing projects, especially in cases which define the "resistance" status of an isolate based on the presence or absence of particular gene elements. Finally, this study supports genomic sequencing as having a strong potential to replace current clinical tests and provide more specificity in antibiotic selection in the clinic.

41

Protein coding potential of annotated non-coding RNAs encoded by the human genome

Hitesh Kore[1, 2], Keshava Datta[1], Shivashankar H Nagaraj[2, 3], Harsha Gowda[1, 2, 3]

1. Cancer Precision Medicine, QIMR Berghofer Medical Research Institute, Kelvin Groove, QLD, Australia

2. Faculty of Health, Queensland University of Technology, Kelvin Groove, QLD, Australia

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

There are ~20,500 annotated protein-coding genes in the human genome which constitutes nearly 2% of the entire genome. This gene set serves as reference dataset in biological studies. However, transcriptome profiling studies have revealed that a large fraction of the human genome is transcribed and most of it is considered to be non-coding. Various ribosome profiling studies in the past decade have identified significant ribosomal occupancy on lncRNAs. Translational efficiency of these transcripts is often comparable with mRNAs. Moreover, mass-spectrometry based studies have provided direct protein-level evidence for a subset of proteins encoded by annotated ncRNAs. In the last 5 years, small proteins encoded by these annotated non-coding regions have been shown to play an important role in various biological processes including development, muscle performance and DNA repair. This suggests that genome annotation pipelines have probably incorrectly annotated some of the protein-coding regions as non-coding. We have systematically analyzed publicly available transcriptome, Ribo-Seq and mass spectrometry datasets using comparative genomics and bioinformatics approaches to evaluate the protein-coding potential of annotated non-coding RNAs. We are developing a machine-learning algorithm that defines various aspects of RNA-protein translation paradigm to identify potential protein-coding candidates among annotated ncRNAs. Identification of novel proteins encoded by ncRNAs will enable researchers to further explore cellular functions regulated by these proteins and their role in various human diseases.

42

The Hitchhiker's Guide to the Yeast Interactome

Ignatius Pang[1], Aidan P. Tay[1, 2], Apurv Goel[1], Sara Ballouz[3], Daniel L. Winter[1], Daniel Weissberger[1], Loïc M. Thibaut[4], Joshua J. Hamey[1], Jesse Gillis[3], Gene Hart-Smith[5], Marc R. Wilkins[1]

1. School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia

2. Transformational Bioinformatics, The Commonwealth Scientific and Industrial Research Organization (CSIRO), Sydney, NSW, Australia

3. Stanley Center for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, The United States of America

4. Computational Genomics, Victor Chang Cardiac Research Institute, Sydney, NSW, Australia

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

The baker's yeast (*Saccharomyces cerevisiae*) is a well-characterised model organism with the most comprehensively mapped eukaryotic interactome networks to date. This talk will outline several studies in which the interactome was used as a scaffold for the co-analysis of -omics datasets (e.g. transcriptomics, proteomics) to gain insights on how these interactions are dynamically regulated within the cell. First, we developed a Cytoscape app, 'PTMOracle' [1, 2], that facilitates the co-visualization of PTMs in the context of PPI networks. Additional types of protein data, including protein sequence and sequence features such as domains, motifs and disordered region could be co-analysed with PPI networks, which enable users to assess the potential impacts of PTMs on protein-protein interactions. Second, the negative genetic interactions in yeast have been systematically screened to near-completeness, but the biological basis of these interactions remains poorly understood. To investigate this, we analysed negative genetic interactions within an integrated biological network, being the union of protein-protein, signalling and regulatory interactions. Network triplet motifs, which contain two genes / proteins that show negative genetic interaction and a third protein from the network, were analysed [3]. Strikingly, only six out of 15 possible triplet types were present in the cell, unlike random networks. Negative genetic interactions among the six triplet motifs showed strong dosage constraints and motifs containing multiple negative genetic interactions highlight regions of 'network vulnerability'; these could be targeted in fungal species for the regulation of cell growth. Third, protein correlation profiling (PCP) enables many intact protein complexes to be identified in single experiments. We co-analysed yeast PCP data with orthogonal gene co-expression data using EGAD [4] and found that the addition of gene co-expression to PCP data contributed mainly to confident identification of known complexes. In summary, providing that the -omics and interactome data are available, the analytical techniques described above are broadly applicable and could be used to analyse the interactomes of other eukaryotic model organisms including human.

1. Tay AP, Liang A, Wilkins MR, Pang CNI. (2019) Visualizing Post-Translational Modifications in Protein Interaction Networks Using PTMOracle. Curr Protoc Bioinformatics. 66(1):e71.

2. Tay AP, Pang CNI, Winter DL, Wilkins MR. (2017) PTMOracle: A Cytoscape App for Covisualizing and Coanalyzing Post-Translational Modifications in Protein Interaction Networks. J Proteome Res.16(5):1988-2003.
3. Pang CNI, Goel A, Wilkins MR. (2018) Investigating the Network Basis of Negative Genetic Interactions in Saccharomyces cerevisiae with Integrated Biological Networks and Triplet Motif Analysis. J Proteome Res. 17(3):1014-1030.
4. Ballouz S, Weber M, Pavlidis P, Gillis J. (2017) EGAD: ultra-fast functional analysis of gene networks. Bioinformatics. 33(4):612-614.

43

Identification of aggressive prostate cancer through in-depth proteomics of tissues and urines

Thomas Kislinger[1]

1. Princess Margaret Cancer Center, Toronto, ONTARIO, Canada

For the 135,000 American men diagnosed with low- or intermediate-risk prostate cancer (PC) each year, clinical outcomes are heterogeneous: 50-80% will be disease-free 10 years following curative-intent therapy while 10-20% will experience recurrence within 18 months, portending lethality. Conversely, 25% diagnosed with low-risk PC elect to enter Active Surveillance, where their disease is monitored by repeat PSA tests and ultrasound guided biopsy to rule out the presence of aggressive PC. Current clinical prognostic factors do not accurately predict disease aggression and clinical outcome for individual men resulting in undertreatment of occult aggressive disease and overtreatment of indolent disease. We have applied proteomics analyses of direct expressed prostatic secretions (dEPS), post-DRE-urines and tissues to identify proteomics signatures of aggressive disease. Combining comprehensive proteomics profiling of dEPS fluids with targeted proteomics and computational biology we discovered robust signatures for extracapsular prostate cancer (Kim et al. Nat Commun. 2016). We are extending on this discovery by developing novel approaches for proteomics profiling of prostate fluids stratified into low, intermediate and high-risk prostate cancer. Our goals are to develop biomarkers to follow patients on active surveillance. In parallel, we are performing proteomics analyses of prostate tissues that have already been extensively profiled by the Canadian Prostate Cancer Genome Network. We have integrated genomic, epigenomic, transcriptomic, and proteomic data generated from 76 intermediate-risk prostate cancer patients. We discovered that the prostate cancer proteome yields four subgroups that differ from previously published DNA-based subgroups and are associated with differential biochemical recurrence. Our data indicated that integration of complementary biomolecules led to the best predictive accuracy (Sinha et al. Cancer Cell 2019). Our results show that proteomics complements other -omics data in stratifying prostate cancer patients and is an underutilized approach for precision medicine. All data has been parsed into a relational database that currently contains quantitative data for over 10,000 proteins. Integration of these data with the rich clinical annotation will enable objective data mining and selection of candidate biomarkers for validation by targeted proteomics.

44

LFQ-Analyst, an interactive web-platform to analyse quantitative proteomics data

Anup Shah[1], Robert JA Goode[1], Cheng Huang[1], David R Powell[1], Ralf Schittenhelm[1]

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

Relative label-free quantification (LFQ) of shotgun proteomics data using precursor (MS1) signal intensities is one of the most commonly used applications to comprehensively and globally quantify proteins across biological samples and conditions. Due to the popularity of this technique, several software packages – such as the popular software suite MaxQuant – have been developed to extract, analyze and compare spectral features, and to report quantitative information of peptides, proteins and even post-translationally modified (PTM) sites. However, there is still a lack of accessible tools for the interpretation and downstream statistical analysis of these complex datasets, in particular for researchers and biologists with no or only limited experience in proteomics, bioinformatics and statistics.
We have therefore created LFQ-Analyst, which is an easy-to-use, interactive web application developed to perform differential expression analysis with “one click” and to visualize label-free quantitative proteomic datasets preprocessed with MaxQuant. LFQ-Analyst provides a wealth of user-analytic features and offers numerous publication-quality result graphics to facilitate statistical and exploratory analysis of label-free quantitative datasets. LFQ-Analyst, including an in-depth user manual, is freely available at https://bioinformatics.erc.monash.edu/apps/LFQ-Analyst.

45

Deciphering the stage-specific glycoprotein signatures of colorectal cancer using integrated mass spectrometry-based omics

Rebeca Kawahara Sakuma[1], Sayantani chatterjee[1], Nicolle Packer[1], Giuseppe Palmisano[2], Seong Beom Ahn[1], Morten Thaysen-Andersen[1]

1. Macquarie University, Sydney, NSW, Australia

2. Departament of Parasitology, University of Sao Paulo, Sao Paulo, SP, Brazil

Glycobiology plays central roles in colorectal cancer (CRC) as exemplified by the aberrant glycosylation decorating immune and CRC-derived glycoproteins. Previous efforts have explored the changes in protein glycosylation in CRC, but the site-specific information of the proteins carrying altered glycosylation as well as their cellular origin is crucially missing. In this study, we used integrated mass spectrometry-based quantitative glycomics and glycoproteomics to characterise the *N*-glycosylation of non-depleted plasma proteins and peripheral blood mononuclear cell (PBMC) obtained from CRC patients spanning four disease stages (Dukes I-IV, n=7/stage) and age-matched healthy donors (n=8). Porous graphitised carbon (PGC)–LC-MS/MS *N*-glycomics revealed a total of 74 and 44 biosynthetically-related *N*-glycans from plasma and PBMC, respectively, across the sample cohort. The plasma *N*-glycome showed 26 *N*-glycans including the abundant biantennary mono- and bi-α2,6-sialylated

glycans with and without α1,6-fucosylation and the related but less abundant tri- and tetra-antennary core-fucosylated sialoglycans. The *N*-glycome displayed stage-specific expression profiles that also differed quantitatively from the *N*-glycan profile of plasma proteins from healthy blood. The PBMC *N*-glycome data displayed a high abundance of paucimannosidic structures in particular in late stages of CRC. Integrated glycomics and glycoproteomics data analysis and adjustment for stage-specific protein level variations revealed that the CRC stage-specific plasma *N*-glycome changes were recapitulated at the individual glycosylation site level of the ten most abundant plasma glycoproteins including haptoglobin, Ig gamma-1 chain C region and fibrinogen. The glycoproteome of PBMC revealed that myeloperoxidase is an important contributor to the elevation of paucimannose observed in PBMCs from patients with advanced CRC. This study demonstrates the immense yet often under-utilised potential of employing integrated MS-based glycomics and glycoproteomics to deconstruct the regulation of the cell-, protein- and tumour-stage-specific glycosylation associated with CRC onset and development.

46

Interactive visualisation, and reproducibility in bioinformatics

David Powell[1]

1. Monash University, Clayton, VIC, Australia

Data visualisation is a key part of bioinformatics and data analysis, not just for communication of results but also for discovery of new insights. The ability to explore data interactively through visualisations is being rapidly adopted in our field, particularly through the use of R/Shiny. There is also a tension between interactive data exploration and reproducible analysis that needs to be addressed. In this talk, I'll discuss the importance of a fast feedback loop between idea, analysis and visualisation and how to improve our tools in this area. I'll also explore our recent work in interactive visualisation, and how we can take advantage of its benefits while also improving reproducibility of analysis.

47

Fitting big science on a small page

Martin Krzywinski[1]

1. BC Cancer Genome Sciences Center, Vancouver, BC, Canada

Without good practices in data exploration, we cannot have answers. Without good design practices, we cannot have good explanations of these answers. This is confounded by the fact that most answers are neither self-explanatory nor immediately obviously correct. Moreover, their reception and impact is directly related to the quality of the explanation.

There are many guidelines for visualization, backed by compelling perception studies that can help us make good choices in shapes, colour and layout. The principles that underpin these choices are now generally well accepted and implemented in many applications. Though it is now easy to make a pile of plots, it is generally very challenging to determine how to sort through the pile to select and order a set for an engaging and expository data presentation.

I will distill the core concepts of information design into practical guidelines for creating scientific figures, presentations and scientific storytelling. Topics include use of colour, information flow for data, process and concept figures. Using examples of published figures in the literature, I apply these guidelines to a redesign process that isolates and emphasizes the essential aspects of the figure without losing context.

This process of design, which is a kind of choreography for the page, can be of great help in assembling individual data visualizations into a cohesive explanation across many levels of detail. In the same way that visualizations are a way to organize data, design is a way to organize visualizations.

I will share with you my experiences in applying tools such as Circos and hive plots, among others, to combine science with visualization and design to create explanations, promote engagement, inspire imagination and, where possible, provide visual support in the often vexing process of research.

48

Trans-omic characterisation of stem cell networks

Pengyi Yang[1]

1. The University of Sydney, Glebe, NSW, Australia

Pluripotency in embryonic stem cells (ESCs) is a highly dynamic process and exists as a continuum of pluripotent states including naive, formative, and primed state. The transition of pluripotency between different states is a key process in ESC fate decisions and is driven by signalling and transcriptional networks established by key kinases and transcription factors. To understand the key steps in ESC fate decisions, we have recently dissected the phased progression of pluripotency from naive towards primed state using a trans-omic approach. Through the comprehensive mapping of the proteome, phosphoproteome, transcriptome, and epigenome of ESCs transitioning from naive towards primed pluripotency, we find that rapid, acute, and widespread changes to the phosphoproteome precede ordered changes to the epigenome, transcriptome, and proteome. Computational reconstruction of signalling and transcriptional networks identifies key kinases, substrates, transcriptional factors, chromatin remodellers, and their downstream targets that together govern ESC fates through the continuum of pluripotent states.

49

The Human Protein Atlas - Implications for Human Biology and Precision Medicine

Cecilia Lindskog[1]

1. Department of Immunology, Genetics and Pathology, Human Protein Atlas, Uppsala University, Uppsala, Sweden

In the evolving era of "big data", integration of datasets from different omics technologies have received increased attention, paving the way for further progress in molecular medicine and targeted treatment. The Human Protein Atlas database (www.proteinatlas.org) based on integration of transcriptomics, antibody-based imaging, mass spectrometry and systems biology constitutes the largest knowledge resource for spatial localization of proteins in organs, tissues, cells and organelles. Divided into six different parts, it covers a wide spectrum of protein localization at different levels. The Tissue Atlas[1] shows the distribution of proteins across all major human tissues and organs, and recent updates include a new classification of all human genes based on tissue specificity and distribution, using a combination of three different transcriptomics datasets. Other additions involve characterization of proteins selectively expressed in rare tissues[2], and single cell evaluation of >500 proteins elevated in testis[3]. The Cell Atlas[4] focuses on subcellular localization of proteins in single cells, and the Pathology Atlas[5] presents the consequence of all human genes on patient survival in cancer. Three new parts were added in September 2019; the Blood Atlas displaying transcriptomic profiles of human blood cells and concentration levels of proteins in blood; the Brain Atlas showing the distribution of proteins in human, mouse and pig brain; and the Metabolic Atlas summarizing the presence of metabolic pathways across human tissues. The Human Protein Atlas database has several potential implications for use in medicine, and constitutes an important starting point for identification of candidate proteins that may contribute to further understanding of disease mechanisms, aid in stratifying high-risk individuals, and guide treatment modalities. In summary, the Human Protein Atlas is a comprehensive stand-alone open-access resource available for researchers worldwide, and is believed to help accelerating efforts meeting future needs in personalized healthcare, and leading to products that will benefit humanity.

1. Uhlén M et al. Tissue-based Map of the Human Proteome. Science. 2015.
2. Sjostedt E et al. Integration of Transcriptomics and Antibody-Based Proteomics for Exploration of Proteins Expressed in Specialized Tissues. J Proteome Res. 2018.
3. Pineau C et al. Cell Type-Specific Expression of Testis Elevated Genes Based on Transcriptomics and Antibody-Based Proteomics. J Proteome Res. 2019.
4. Thul P et al. A Subcellular Map of the Human Proteome. Science. 2017
5. Uhlén M et al. A Pathology Atlas of the Human Cancer Transcriptome. Science. 2017.

50

Phosphomatics: A knowledge-based approach to investigating high-throughput phosphoproteomics data

Michael G Leeming[1], **Sean O'Callaghan**[2], **Ching-Seng Ang**[1], **Shuai Nie**[1], **Swati Varshney**[1], **Syeda Sadia Ameen**[1], **Heung-Chin Cheng**[1], **Nicholas A Williamson**[1]

1. University of Melbourne, Parkville, VIC, Australia

2. Nuritas Ltd, Dublin, Ireland

Mass spectrometry-based phosphoproteomics is undoubtedly one of the most powerful tools available for investigating the detailed molecular events that occur in response to cellular stimulus and, while these experiments can routinely detect and quantify thousands of phosphorylated peptides, interpreting this data remains challenging. Identification of the upstream kinases that affect the observed phosphorylations is key to understanding the factors responsible for differences in observed phosphorylation profiles however the astounding complexity of substrate-kinase relationships make this task difficult. For example, protein kinases frequently have broad but overlapping sets of substrates and protein function is often differentially regulated by selective phosphorylation at different sites. Making the most of our phosphoproteomics data clearly requires development of a knowledge base against which phosphoproteome data can be easily queried.

Here, we present '*Phosphomatics*' – a new web-based tool for interrogating possible upstream kinases for observed phosphopeptides. *Phosphomatics* allows users to upload the results of global phosphoproteomics experiments containing thousands of phosphorylated peptides and search each of these against databases of known substrate-kinase relationships. Users can then interactively explore substrate-kinase relationships constructed from their input data based upon low-throughput, manually curated literature sources and results can be interrogated from either substrate-centric, kinase-centric or pathway-centric perspectives. For specific substrate-kinase relationships, machine learning based tools, in addition to publications drawn from existing databases, are utilised to provide users as much evidence as possible to support the selected phosphorylation event. Users can download results files summarising possible substrate-kinase relationships as well as graphics providing a visual overview of phosphorylation networks and important kinases.

Phosphomatics is freely available via the internet at: www.phosphomatics.com

51

Delving deeper into the proteome with an improved and highly parallelised feature detection pipeline for the Bruker timsTOF Pro

Daryl Wilding-McBride[1], **Giuseppe Infusini**[1], **Andrew Webb**[1]

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

Mass spectrometry (MS)-based quantitative proteomics depends on accurate and precise feature detection. Reliable extraction of feature characteristics has a substantial impact on the number and quality of peptides and proteins identified and quantified.

Feature detection is inherently difficult because of the high complexity of MS proteomics data from biological samples, where the data is voluminous, complex and imbued with chemical and electronic noise. This combination of complexity and noise presents a daunting hurdle for improving algorithms when new technology arises, as any improvements in instrument sensitivity and speed also concurrently increase the data size, complexity and types of noise detected.

While Ion-Mobility (MS) is not a new technology, the introduction of the Bruker timsTOF Pro has made the technology commercially available. The instrument's dual TIMS stages provides improved sensitivity with respect to conventional drift tubes. The increased size and complexity of the four-dimensional data produced by the instrument, though, is a significant added challenge for existing feature detection algorithms.

In this work, we present a new data processing pipeline for feature detection, peptide and protein identification, and quantitation from timsTOF data. The algorithm focuses on performing feature detection on a patch of data where the instrument has performed a fragmentation event. The pipeline is highly parallelisable to reduce analysis times. We also introduce a step in the pipeline where we perform targeted feature detection based on the peptide features identified in the experiment, and use machine learning to propagate identifications across LC-MS runs and to reduce missing values.

To test this pipeline we used a set of samples from a HeLa:e.Coli mixture where the proteomes were mixed in 1:1 or 1:3 ratios and analysed using a 15-minute gradient. The analysis identifies 28,186 peptides (max q-value is 0.01). For comparison, MaxQuant identified 27,901 peptides. Our method also shows that we achieve fewer missing values.

52

Fidelity of protein synthesis is required for energy production

Aleksandra Filipovska[1]

1. The University of Western Australia, Nedlands, WESTERN AUSTRALIA, Australia

Mitochondria are composed of proteins encoded by both the nuclear and mitochondrial genomes and the coordinated expression of both genomes is essential for energy production. Impaired energy production leads to mitochondrial dysfunction that causes or contributes significantly to a variety of diseases including cardiovascular diseases. To unravel how mitochondrial function fails and to identify therapeutic targets it is necessary (i) to understand how gene expression is regulated between mitochondria and the nucleus and (ii) how this regulation is disrupted in disease. Mammalian mitochondrial ribosomes are unique molecular machines that translate 11 leaderless mRNAs. To date it is not clear how mitoribosomes recognize and initiate translation in the absence of untranslated regions in the mitochondrial mRNAs. Translation initiation in mitochondria shares similarities with prokaryotic systems, such as the formation of a ternary complex of fMet-tRNA^{Met}, mRNA and the 28S subunit, but differs in the requirements for initiation factors. Mitochondria have two initiation factors, MTIF2 that closes the decoding centre and stabilizes the binding of the fMet-tRNA^{Met} to the leaderless mRNAs, and MTIF3 whose role is not clear. We knocked out *Mtif3* in mice and show that this protein is essential for embryo development and heart- and skeletal muscle-specific loss of MTIF3 causes dilated cardiomyopathy. We identify increased but uncoordinated mitochondrial protein synthesis in mice lacking MTIF3 that results in loss of specific respiratory complexes by mass spectrometry. Therefore, we show that coordinated assembly of OXPHOS complexes requires stoichiometric levels of nuclear and mitochondrially-encoded protein subunits *in vivo*. Our ribosome profiling and transcriptomic analyses show that MTIF3 is required for recognition and regulation of translation initiation of mitochondrial mRNAs, but not dissociation of the ribosome subunits. To investigate translation fidelity we created yeast and mouse models with error-prone and hyper-accurate translation, which revealed that translation rate is more important than translational accuracy for cell function and energy production. Our proteomic and metabolomic analyses identified mammalian-specific signalling pathways that respond to changes in the fidelity of protein synthesis and regulate energy metabolism.

53

Addressing unmet clinical diagnostic needs with proteomics

Michelle M Hill[1]

1. QIMR Berghofer Medical Research Institute, Brisbane, Australia

Correct and timely diagnosis is the essential first step to achieving ideal health outcomes. Diagnostic error can occur due to wrong test being performed, delayed diagnosis, or lack of appropriate tests. Many common and rare diseases do not have specific diagnostic tests available to assist in clinical decision-making. To address these unmet diagnostic needs with proteomic technologies, my laboratory has worked with clinical collaborators over the past decade in various research projects. This talk will highlight two major areas of diagnostics: the mass spectrometry-based amyloidosis diagnostic test that has helped patients with this rare disorder to receive the right treatment, and the multi-stage development of surveillance blood test for oesophageal cancer.

54

Global proteomic and phosphoproteomic profiling to identify proteins and pathways that regulate cell survival in hypoxic conditions

Keshava K. Datta[1], **Rebekah Ziegman**[1], **Sonali Mohan**[1, 2], **Harsha Gowda**[1, 2, 3]

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

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

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

Hypoxia is a common feature in various solid tumors. Cancer cells in hypoxic environments are resistant to both chemotherapy and radiation. Hypoxia is also associated with immune suppression. Identification of proteins and pathways that regulate survival

of cancer cells in hypoxic environments can reveal potential vulnerabilities that can be exploited to improve efficacy of anti-cancer therapy. Gene expression profiling studies have identified several hypoxia-induced genes. This includes well-known transcription factor hypoxia-inducible factor 1-alpha (HIF-1α). We carried out global proteome profiling and phosphoproteome profiling in melanoma cell lines to identify proteins and pathways that are induced by hypoxia. We used Orbitrap Fusion Mass Spectrometer for analysis and employed TMT-based quantitation for global proteomic and phosphoproteomic comparison. As expected, several proteins that are known targets of hypoxia inducible factors (HIFs) were found to be overexpressed in the hypoxic models. In addition, several metabolic enzymes showed altered expression revealing metabolic reprogramming in hypoxic conditions. Phosphoproteomic profiling revealed kinase mediated signaling pathways that are induced in hypoxic conditions. Our data provides a comprehensive view of proteomic alterations in hypoxic conditions and reveals potential mechanisms that regulate cell survival in hypoxic environments. These mechanisms can be targeted to improve therapeutic outcomes in cancer treatment.

55

Analysis of FFPE tissues to investigate lymph node protein expression in patients with leukemia

Lauren A Thurgood[1], Lara M Escane[1], Ash N Rowland[1], Karen M Lower[1], Bryone Kuss[1]

1. Flinders University, Bedford Park, SA, Australia

Background: Chronic lymphocytic leukaemia (CLL) is one of the most prevalent leukaemias in Australia and is characterised by the proliferation of non-functional B-cells, in the microenvironments of lymph nodes (LN) and bone marrow. Circulating CLL cells are quiescent and their proteome may well be different from that of an active proliferating cell. In order to understand the drivers of CLL proliferation, it is critical to understand what is occurring in these proliferative niches. LN biopsies are rarely taken in CLL as they are not essential for diagnosis or disease staging, subsequently, this study used archived formalin fixed paraffin embedded (FFPE) LN biopsies for protein extraction.

Methods: We compared FFPE LN from CLL patients (n=14) to healthy LN sections (n=6). Proteins were isolated using heat and xylene to remove the paraffin, rehydration of the tissue followed by mechanical and chemical lysis. Protein identification and quantification was carried out using a SWATH MS/MS approach.

Results: Approximately 1000 proteins were identified, and pathway analysis identified lipid metabolism and degradation as the key differences between CLL and healthy LNs. Metabolic pathways were further explored using a variety of *in vitro* techniques including; FA uptake assays, microscopy, as well as proteomic/transcriptomic techniques to identify FA uptake receptors. We found that CLL is a lipolytic tumour, preferring the uptake of long-chain FAs over short-chain FAs or glucose and that the process is driven by endocytosis rather than receptor-mediated uptake. Once FAs are internalised, the CLL cells either store the excess fatty acids inside lipid droplets, which can be observed under electron microscopy, or increase beta-oxidation pathways to utilise the lipid energy store.

Significance: These findings have clinical implications in the development of diagnostic imaging techniques that improve understanding of disease activity (Lipid-based PET Scans) and in potential therapeutic approaches, whereby the metabolomics and lipid dependency of this leukaemia are exploited.

56

Simultaneous mass spectrometry imaging of multiple neuropeptides in the brain and alterations induced by experimental Parkinsonism and L-DOPA therapy

Per E. Andren[1], Heather Hulme[1], Elva Fridjonsdottir[1], Halla Gunnarsdottir[1], Theodosia Vallianatou[1], Xiaoqun Zhang[2], Henrik Wadensten[1], Reza Shariatgorji[1], Anna Nilsson[1], Erwan Bezard[3], Per Svenningsson[2]

1. Dept. of Pharmaceutical Biosciences, Medical Mass Spectrometry Imaging, Uppsala University, Uppsala, Sweden

2. Dept. of Neurology, Karolinska Institutet, Stockholm, Sweden

3. Université de Bordeaux, Institut des Maladies Neurodégénératives, Bordeaux, France

Neuropeptides are important signalling molecules in the brain and alterations in their expression levels have been linked to neurological disorders such as Parkinson's disease. It is challenging to map neuropeptide changes across and within brain regions because of their low *in vivo* concentrations and complex post-translational processing. Consequently, the role of neuropeptides in Parkinson's disease is not well understood. Thus, we have developed and validated a method to image multiple neuropeptides simultaneously in both rat and primate brain tissue sections by matrix-assisted laser desorption-ionisation mass spectrometry imaging at high lateral resolution. Using two neurotoxin models of Parkinson's disease, i.e., the unilateral 6-hydroxydopamine rat model and the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine primate model, we imaged changes in enkephalins, dynorphins, tachykinins and neurotensin associated with the dopaminergic denervation and L-DOPA treatment in multiple brain regions. L-DOPA administration significantly affected neuropeptides in the globus pallidus, while neuropeptides in the caudate-putamen were mostly affected by dopamine depletion. Using high lateral resolution imaging, we observed an increase of neurotensin in the dorsal sub-region of the globus pallidus after dopamine depletion. This study highlights the capacity of mass spectrometry imaging to elucidate the dynamics of neuropeptide signalling during Parkinson's disease and its treatment.

57

Reference proteome-based subtyping of liver cancer

J. Eugene Lee[1]

1. Korea Research Institute of Standards and Science, Daejeon, Korea

Classification of hepatocellular carcinoma (HCC) into clinical subtypes is in high demand for the selection of proper treatment. For patients with HCC who are not or no longer candidates for locoregional therapy, oral multi-kinase inhibitors are the only established systemic alternative. Yet, the response rate of these FDA-approved drugs against HCC are less than two percent, necessitating in-depth, personalized molecular characterization of the disease. Here, we adopted quantitative proteomic approaches in an attempt to characterize HCC subtypes in proteome level, and link these subtypes to drug response. We developed a liver reference proteome, a mixture of equal amounts of total proteins from seven liver cancer cell lines metabolically labeled with stable isotopes. This labeled reference proteome was utilized to cross-compare protein abundances of ten liver cancer cell lines that are widely used in research. Shotgun proteomic analysis on a hybrid quadrupole-orbitrap mass spectrometer with high mass accuracy at the MS and MS/MS levels yielded a liver proteome composed of 8,883 identified proteins, spanning 7 orders of magnitude in signal intensity. High-accuracy quantification allowed robust differentiation of HCC cell lines into three subtypes. Among these subtypes, one subtype displayed similar protein expression pattern as HCC tissues correlated with poor outcome, vascular invasion, high alpha-fetoprotein levels, and hepatitis B virus infection. The proteome profile of the other two subtypes did not mimic any of the catalogued proteome profile of HCC tissues. HCC characterization by protein expression may facilitate the establishment of clinical guidelines for subtype-specific systematic therapy.

58

Ciliary proteome from motility-defective multiciliated cells

Taejoon Kwon[2, 1], **Keun Yeong Kwon**[1], **Hyeongsun Jung**[1], **Byung Gyu Kim**[2], **Kyungjae Myung**[2, 1], **Tae Joo Park**[2, 1]

1. Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

2. Center for Genomic Integrity, IBS, Ulsan, Republic of Korea

Multiciliated cells (MCCs) on mucociliary epithelium have important roles to generate the directional fluid flow by coordinated movement of several hundreds of motile cilia, and their motility defect is linked to several clinical symptoms in human, such as chronic respiratory infections, increased risks of hydrocephalus and female infertility. Although there were several genes reported to regulate the motility of multi-cilia on MCCs, without perturbing the cilia formation, the protein composition changes inside the organelle by motility defect were not revealed yet.

Based on the mucociliary epithelium of the *Xenopus* embryos as a model system, we recently published the simple method to isolate almost intact whole ciliary structures for label-free shotgun proteomics and identified 1,009 vertebrate conserved ciliary proteins. Expanding this approach, here we reported the ciliary proteome of motility-defective multiciliated cells. By suppressing the expression of two different genes involved in cilia motility regulation, we identified about a hundred ciliary proteins differentially abundant (either by expression or localization change) compared to those in the normal cilia. Among them, 35 proteins including proteins for outer dynein arm complex, energy production, were commonly changed regardless of their genetic background, which may be specifically involved in cilia motility, not in ciliogenesis.

By providing the protein changes inside the organelle, this result will provide new insight to understand how cilia motility is regulated. The detailed validation results also will be presented.

59

Mini-malting and mini-mashing methods for early stage validation of new barley varieties and malt accreditation and quality control of beer brewing

Christopher H Caboche[1, 2], **Edward D Kerr**[1, 2], **Ben L Schulz**[1, 2, 3]

1. School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Queensland, Australia

2. Australian Infectious Diseases Research Centre, School of Chemistry & Molecular Biosciences, University of Queensland, St Lucia, Queensland, Australia

3. Centre for Biopharmaceutical Innovation, Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, St Lucia, Queensland, Australia

Beer brewing is a complex process which is well-optimized industrially yet poorly understood biochemically due to its molecular complexity. Accreditation for new malting barley varieties and quality control of existing malted barley for the brewing industry in Australia is a lengthy and expensive process that requires large amounts of barley grain, and relies on large, expensive, and specialized heavy equipment. To overcome these challenges, we have developed miniaturized methods for malting and mashing that are compatible with downstream analytics, and that requires only a single seed1. These methods can be easily implemented in industrial or scientific settings. We used DIA/SWATH-MS to measure the abundance and site-specific modifications of proteins throughout the mini-malting and mini-mashing processes, and benchmarked these against current market “micro” equipment and industry-scale equipment to compare performance. These mini-scale systems will allow cheaper, faster, and higher-throughput analyses to track the dynamic proteomes during the brewing processes of malting, mashing, boiling, and fermentation.

1. Kerr, E. D., Caboche, C. H., & Schulz, B. L. (2019). Posttranslational Modifications Drive Protein Stability to Control the Dynamic Beer Brewing Proteome. Molecular & Cellular Proteomics, 18(9), 1721–1731. https://doi.org/10.1074/mcp.ra119.001526

60

The role of upstream phosphorylation in the regulation of histone methylation

Ryan 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) using a multi-protease and multi-fragmentation type approach. With respect to phosphorylation, to date, we have identified modification sites on the histone methyltransferases Set2 (12 sites), Set5 (14 sites) and Dot1 (11 sites), and the demethylases Jhd1 (one site), Jhd2 (two sites) and Rph1 (19 sites). Of these 62 phosphorylation sites, 35 validate those observed previously in high throughput phosphoproteomic screens, and 27 sites are novel. To determine the upstream kinases responsible for the phosphorylation, and potential regulation of these enzymes, mass spectrometric analysis was employed to monitor levels of histone methylation in kinase knockout yeast strains. As a proof of concept, quantification of H3K79 methylation in the knockout cells established 43 kinases that are not responsible for the regulation of Dot1 methyltransferase activity. The screening of all other non-essential kinases is in progress. We plan to extend this methodology to the other yeast histone methyltransferases and demethylases in order to comprehensively integrate these enzymes into intracellular signalling pathways, and ultimately facilitate the assembly of the first regulatory network of histone methylation in eukaryotes.

61

Revealing the proteome of brain derived exosomes isolated from human amyotrophic lateral sclerosis post-mortem tissues

Natasha Vassileff[1], Kolin Harinda Rajapaksha[1], Mitch Shambrook[1], Amirmohammad Nasiri Kenari[1], Jacky Chan[2], Catriona Mclean[2, 3], Andrew Hill[1], Laura Vella[2], Lesley Cheng[1]

1. La Trobe Institute for Molecular Science, Bundoora, VICTORIA, Australia

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

3. Department of Anatomical Pathology, Alfred Health, Melbourne, Victoria, Australia

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease characterised by the deposition of misfolded proteins in the motor cortex and motor neurons. Although a multitude of ALS-associated mutated proteins have been identified, few have been linked to exosomes, a form of extracellular vesicles involved in inter-cellular communication. Additionally, the role of exosomes in ALS is undetermined, specifically in relation to pathogenic stress granule formation, a response to cellular stress involving aggregation of non-coding RNAs and their RNA binding proteins. Therefore, this study aimed to determine the proteome of brain derived exosomes (BDEs) isolated from ALS subjects and identify novel ALS-associated deregulated proteins and their potential contributions to pathogenic pathways in ALS. BDEs were isolated from human post-mortem ALS (n=10) and control (n=5) motor cortex brain tissues through a novel ultracentrifugation protocol (1). Following thorough characterisation, BDEs successfully met the minimum criteria required by The International Society for Extracellular Vesicles to be classified as exosomes (2). The BDEs' protein content subsequently underwent mass spectrometry analysis, allowing for a panel of novel ALS-associated proteins to be identified. This panel consisted of 16 statistically significant differentially packaged proteins identified in the ALS BDEs compared to the control BDEs. This included several up-regulated RNA binding proteins which were determined through pathway analysis to be associated with stress granule dynamics. The identification of these RNA binding proteins in the ALS BDEs suggests there may be a relationship between ALS-associated stress granules and ALS BDE packaging, highlighting a potential role for exosomes in the pathogenesis of ALS.

1. 1. Vella LJ, Scicluna BJ, Cheng L, Bawden EG, Masters CL, Ang C-S, Willamson N, McLean C, Barnham KJ, Hill AF (2017) A rigorous method to enrich for exosomes from brain tissue. Journal of Extracellular Vesicles 6: 1348885
2. 2. Lötvall J, Hill AF, Hochberg F, Buzás EI, Di Vizio D, Gardiner C, Gho YS, Kurochkin IV, Mathivanan S, Quesenberry P, Sahoo S, Tahara H, Wauben MH, Witwer KW, Théry C (2014) Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. Journal of Extracellular Vesicles 3: 26913

62

Development of a primary cell proficient phosphoproteomic workflow

Rune H Larsen[1, 2], Jarrod Sandow[1], Laura Dagley[1], Andrew Webb[1]

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

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

Phosphorylation is a critical post translational modification (PTM) that occurs on proteins to modulate the propagation of intracellular signals. Phosphoproteomics has established itself as the tool of choice to investigate the complexities of intracellular signalling cascades in an unbiased and highly sensitive manner. Phosphoproteomics is able to delineate how signals change when exposed to different stimuli or pathological conditions in time.

However, the utility of phosphoproteomics has been somewhat hampered by the relatively large amount of input material required to enable sufficient detection of protein phosphorylation abundance differences (~200 ug). This limitation has restricted the majority of phosphoproteomics studies to *in vitro*experiments or mixed tissues whereas the true power of phosphoproteomics is likely to be revealed from studies in specific subsets of primary cells. The advent of improved protein digestion and phosphopeptide enrichment methods using optimized buffers and cleaner digests have vastly decreased the sample amounts required. Secondly, the increases in speed and sensitivity of mass spectrometers such as the TimsTOF Pro have greatly aided the identification of phosphorylated peptides without the need for extensive fractionation. The TimsTOF pro with its PASEF acquisition strategy vastly increases the sensitivity and MS/MS acquisition speed while maintaining high quality spectra.

Here we present a combined workflow that utilizes our USP3 acid hydrolysis workflow coupled to an Fe-IMAC enrichment protocol analysed on a phosphoproteomics optimized timsTOF Pro. we show the capabilities of the workflow using low sample amounts derived from purified primary cells (£20 ug).

63

Studying uncharacterised non-structural proteins of Rabbit haemorrhagic disease virus

Elena Smertina[1, 2], Robyn Hall[2], Tanja Strive[3, 2], Michael Frese[1, 2]

1. Faculty of Science and Technology, University of Canberra , Canberra , ACT, Australia

2. Health & Biosecurity , CSIRO, Canberra, ACT, Australia

3. Invasive animals cooperative research centre, University of Canberra , Canberra, ACT, Australia

The Rabbit haemorrhagic disease virus (RHDV) is a highly virulent virus that cannot be grown in cell culture. The functions of several non-structural proteins of RHDV (p16, p23 and p29) are still unknown. It is generally assumed that these proteins counteract the host's antiviral defense, similar to the certain norovirus proteins, e.g., p48. Norovirus protein p48 affects the organisation of intracellular membranes and thus inhibits secretory pathways. The functions of RHDV p16, p23 and p29 could be deduced through the identification of their cellular and viral interaction partners. In addition, changes to whole cell proteome upon the expression of the proteins of interest can be tracked. To identify protein interaction partners, we utilised a pull-down assay with FLAG-tagged viral proteins and anti-FLAG resin. Several possible cellular interaction partners were identified after co-immunoprecipitation and quantified using mass spectrometry-based SILAC (stable isotope labelling in cell culture) approach. Preliminary results indicate that proteins p23 and p29 may take part in the disruption of intracellular membranes with subsequent formation of vesicular compartments for viral replication and interact with cellular stress-induced proteins.

64

Mapping the lysosomal transporter proteome of osteoclasts uncovers new regulators of bone homeostasis

Amy BP Ribet[1], Daniel Yagoub[2], Jamie Tan[1], Laila Abudulai[3, 4], Pei Ying Ng[1], Nathan J Pavlos[1]

1. School of Biomedical Science, UWA, CRAWLEY, WA, Australia

2. Medical school, UWA, CRAWLEY, WA, Australia

3. Centre for Microscopy Analysis and Characterisation, UWA, CRAWLEY, WA, Australia

4. School of Molecular Science, UWA, CRAWLEY, WA, Australia

The ruffled border membrane is a giant lysosome-related secretory organelle unique to bone-digesting osteoclasts. The fusion of secretory lysosomes with the ruffled border equips it with sets of lysosomal membrane proteins that are requisite for the cell's bone resorptive function. Despite its crucial importance, we still lack elementary information on the protein composition of the ruffled border membrane, including the numbers and identities of lysosomal membrane residents whose usual functions are to facilitate the exchange of molecules across its membrane (i.e. transporters). To extend the molecular inventory of membrane transport proteins operating at the ruffled border, we have combined biochemical methods with bottom-up proteomics. Through an in-gel tryptic digestion, and nanoflow liquid chromatography high-resolution tandem mass spectrometry (LC-MS/MS) we have unbiasedly surveyed the ruffled border lysosomal membrane proteome using isolated secretory lysosomes as a surrogate organelle. Our analysis identified 3804 master proteins (two or more unique peptides) of which 1100 are functionally assigned as membrane transport proteins. These transport proteins included established components of the osteoclasts' 'bone-resorbing machinery' including the entire V-ATPase proton pump complex and chloride ion channels thus validating our approach. It also uncovered several other membrane transport proteins predicted to reside on lysosomes but whose functions remain to be assigned, including several members of the secondary active transporter superfamily of solute carriers (Slc). By combining a suite of biochemical, cell biology and genetic studies, we demonstrated the robustness and utility of our proteomic screen using the Slc37a2 transporter as a prototype. This approach has allowed us, for the first time, to unmask the osteoclast lysosomal transport protein cache (termed the 'Transportome'), that will serve as a powerful resource for the future interrogation of 'orphan' lysosomal transport proteins operating at the osteoclast ruffled border and may account for a subset of human bone-sclerosing disorders for which the underlying molecular determinant remains unknown.

65

Defining the "mucinome": enzyme toolkit for enrichment and analysis of mucin-domain glycoproteins

Stacy Malaker[1]

1. Stanford University, Stanford, CA, United States

Mucin domains are densely O-glycosylated modular protein domains that are found in a wide variety of cell surface and secreted proteins. Mucin-domain glycoproteins are known to be key players in a host of human diseases, especially cancer, wherein mucin expression and glycosylation patterns are altered. Mucin biology has been difficult to study at the molecular level in part because methods to manipulate and structurally characterize mucin domains are lacking. One major issue is that these domains are resistant to degradation by trypsin, meaning the majority of their sequence space is often left unanalyzed. Selective mucin degradation or enrichment, especially in a sequence- and glycan-specific manner, can facilitate study of these proteins by mass spectrometry.

We first expressed and characterized a bacterial mucinase, StcE, and demonstrated that it selectively cleaves mucins in a glycan- and peptide- specific manner. We went on to use its unique properties to improve sequence coverage, glycosite mapping, and glycoform analysis of recombinant human mucins by mass spectrometry. To expand on this work, we expressed and characterized several other bacterial mucinases to generate a mucin-selective enzymatic toolkit. Their activities were confirmed using a panel of O-glycoproteins by mass spectrometry. We manually validated peptide sequences from MS/MS spectra to elucidate all cleaved peptides present in the mucinase-digested samples but not in the control samples, revealing that each enzyme has a slightly different cleavage motif. Interestingly, all of the enzymes rely on a combination of peptide sequence and glycosylation status. Together with StcE, we have characterized a total of five bacterial mucinases capable of digesting mucins into peptides amenable for mass spectrometric analysis.

Further, given the enzymes' selectivity for mucin-domain glycoproteins, we reasoned that they could be employed to purify mucins from protein mixtures. Thus, inactivated mucinases were conjugated to aldehyde beads using reductive amidation. Using the enzyme-conjugated beads, we demonstrate that we can selectively enrich for mucin-domain glycoproteins from lysate and crude cancer patient ascites fluid. We are thus attempting to define the "mucinome", as a comprehensive list of mucin-domain glycoproteins does not exist. Future experiments will be devoted to isolation, digestion, and characterization of mucins from human cancer patient ascites fluid, with the ultimate goal of identifying diagnostic and/or prognostic markers of disease states.

66

Tools and methods for probing the biology of tryptophan C-mannosylation

Ethan Goddard-Borger[2, 1]

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

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

Tryptophan C-mannosylation is an unusual metazoan co-translational modification found on many cell-surface receptors and extracellular proteins. It is the only type of protein glycosylation that involves the formation of a carbon-carbon bond between sugar and polypeptide. This unique chemistry is performed by the tryptophan C-mannosyltransferases, which are integral ER membrane proteins that utilise dolichol-phosphate mannose (Dol-P-Man) to glycosylate the C-2 position of tryptophan side chains within the WXXW consensus sequence.[1,2] Little else is known about tryptophan C-mannosylation or the enzymes that install this modification, largely due to limitations in the existing approaches to installing, detecting and blocking the modification. We have developed new approaches to tackling these problems to provide fresh insights into the biology of tryptophan C-mannosylation. Engineering C-mannosylation pathways into Pichia pastoris has provided a convenient microbial expression platform for the production of near-homogenous glycoforms on a multi-milligram scale. Efficient chemical syntheses of C-mannosyl tryptophan building blocks has enabled the use of solid-phase peptide synthesis for the production of pure glycoforms. Together, these methods for the production of C-mannosylated glycoforms have facilitated an examination of how C-mannosylation impacts protein stability, function and enzymatic activity. They have also led to the first monoclonal antibodies capable of detecting the modification.
In addition, we have established an in vitro assay of tryptophan C-mannosyltransferase activity, which have been used to explore the enzyme's substrate preference and perform mutagenesis studies. These experiments revealed that the tryptophan C-mannosyltransferases are more promiscuous than previously appreciated. This assay, our newfound knowledge of enzyme substrate preference, and new tools for detecting tryptophan mannosylation, have enabled us to develop the first C-mannosyltransferase inhibitor, which also happens to be active in cells. Collectively, this suite of tools provides a strong foundation for further exploration of tryptophan C-mannosylation biology.

1. F. F. Buettner, A. Ashikov, B. Tiemann, L. Lehle, H. Bakker, Mol. Cell, 2013, 50, 295-302.
2. A. Shcherbakova, B. Tiemann, F. F. Buettner, H. Bakker, Proc. Natl. Acad. Sci. USA, 2017, 114, 2574-2579.

67

Granule-Specific *N*-Glycosylation and Chlorination Activity of Neutrophil Myeloperoxidase

Harry C. Tjondro[1], Julian Ugonotti[1], Sayantani Chatterjee[1], Rebeca Sakuma[1], Ian Loke[1], Vignesh Venkatakrishnan[2], Siyun Chen[3], Weston Struwe[3], Benjamin L. Parker[4], Johan Bylund[2], Oliver C. Grant[5], Robert J. Woods[5], Anna Karlsson[2], Morten Thaysen-Andersen[1]

1. Biomolecular Discovery Research Centre, Macquarie University, Sydney, NSW, Australia

2. Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

3. Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, Oxford, United Kingdom

4. School of Biomedical Sciences, University of Melbourne, Melbourne, VIC, Australia

5. Complex Carbohydrate Research Center, University of Georgia, Athens, GA, United States

The heavily glycosylated myeloperoxidase (MPO), a key component of neutrophil-mediated innate immunity, produces anti-microbial hypochlorous acid to combat pathogens, but the functional importance of MPO glycosylation remains unexplored. Herein, we first performed deep structural characterisation of the *N*-glycosylation decorating human MPO extracted from whole neutrophils (nMPO) using mass spectrometry-driven glycomics, glycopeptide, and intact glycoprotein profiling. Unusual site-specific signatures were observed including paucimannosidic *N*-glycans of Asn323 (47%) and Asn483 (56%), oligomannosidic *N*-glycans of Asn355 (97%) and Asn391 (64%) while Asn729 was unoccupied (44%) or carried equally peculiar chitobiose core-type *N*-glycans (33%). Native mass spectrometry and a novel mass photometry technique revealed extreme glycoform heterogeneity and the expected stoichiometry of the common dimeric (141-148 kDa) and less-abundant monomeric (70-73 kDa) nMPO forms. Native gel electrophoresis and glycopeptide profiling demonstrated that monomeric and dimeric nMPO displayed glycan differences of Asn483 located in the dimerization interface. The solvent accessibility of the glycosylation sites of maturely folded monomeric and dimeric MPO correlated with the degree of early- (Golgi) and late-stage (granule) *N*-glycan processing providing support for an oligomerisation-dependent glycan processing at Asn483. Excitingly, glycoprofiling of subcellularly-fractionated MPO demonstrated distinct granule-specific Asn355- and Asn391-glycosylation across the neutrophil compartments. Notably, the *N*-glycans carried by specific/gelatinase granule-resident MPO (Sp/Ge-MPO) displayed extreme truncation to GlcNAcβ-Asn at Asn355 and Asn391 whereas the azurophilic granule-resident MPO (Az-MPO) displayed oligomannosidic glycans at those sites. Interestingly, the Sp/Ge-MPO exhibited a comparably higher chlorination activity than Az-MPO, an observation that could be recapitulated for endoglycosidase H-treated nMPO simulating the Sp/Ge-MPO glycoform. Molecular dynamics simulations of granule-relevant MPO glycoforms were used to support that Asn355- and Asn391-glycans are able to perturb the conformation of the heme-containing active site of human MPO. We are the first to report on the granule-specific *N*-glycosylation and chlorination activity of neutrophil MPO revealing a novel fascinating feature of neutrophil glycobiology.

68

Immunopeptidomic analysis reveals that deamidated HLA-bound peptides arise predominantly from deglycosylated precursors

Shutao Mei[1], Rochelle Ayala[1], Sri Ramarathinam[1], Patricia Illing[1], Pouya Faridi[1], Jiangning Song[1], Anthony Purcell[1], Nathan Croft[1]

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

Background:

The generation and functional recognition of post-translationally modified (PTM) peptides by T lymphocytes has attracted considerable interest in immunology. Disease associated antigen modification may generate novel PTM neo-peptides. These modified peptides presented by human leukocyte antigen (HLA) molecules have the potential to form excellent T cell targets. With the rapid advances in mass spectrometry, immunoproteomics studies routinely identify thousands of HLA-bound peptides from cell lines and tissue samples. However, except for S/T phosphorylation, in-depth analysis of the proportion and nature of peptides bearing one or more other PTMs has not been studied extensively.

Methodologies:

Here, we apply a mass spectrometry-based methodology to comprehensively profile peptides with PTMs bound to a variety of HLA allotypes.

Results:

Analysis of a diverse set of HLA class I immunopeptidomes revealed that methionine oxidation was the most commonly observed PTM, with deamidation of asparagine and glutamine the next most prevalent type of modification. Detailed analysis of HLA class I deamidation motifs revealed a strong prevalence of the known N-linked glycosylation motif (NX(S/T), where X is any amino acid except proline) for asparagine deamidated peptides. Notably no such motif was observed for glutamine deamidated peptides or for asparagine deamidated peptides isolated from HLA class II molecules. Subsequent blocking of PNGase activity confirmed the role of deglycosylation of NX(S/T)-bearing asparagine-deamidated peptides as a major source of HLA-bound ligands.

Conclusions:

These data indicate that the immunopeptidome is enriched for peptides derived from formerly glycosylated proteins and specifically those that have been retro-translocated from the ER and targeted for deglycosylation and degradation in the cytoplasm. The results not only highlight the link between glycosylation and asparagine deamidation but will help to train models to predict the presence of asparagine deamidated peptides in immunopeptidome, ultimately aiding novel vaccine design.

69

Synthetic Glycopeptides assisted large scale glycoproteomics from sample preparation to data analysis

Kathirvel Alagesan[1], Chi-Hung Lin[1], Arun Everest-Dass[1], Mark von Itzstein[1], Daniel Kolarich[1]

1. Griffith University, Gold Coast, QUEENSLAND, Australia

Intact glycopeptide characterization is an imperative yet challenging component of glycoprotein analysis. Elucidation of both glycan and peptide requires specific sample preparation workflows that, in combination with multiple tandem mass spectrometry (MS/MS) approaches, enable identification of single glycopeptide species (1). Here, we systematically optimized and evaluated the strengths and weaknesses of the glycoproteomics workflows using synthetic glycopeptides from sample preparation to its implication on data analysis.

A library of >100 synthetic *N*-glycopeptides representing human serum glycoproteins was synthesized. These synthetic glycopeptides fostered the systematic investigation to explore the advantages and limitations of the glycoproteomics workflow and comment on their suitability for high-throughput glycoproteomic studies. All experiments were performed on a quadrupole-Orbitrap-linear ion trap Tribrid mass spectrometer.

A simplified approach to purify and produce a panel of glycosylated amino acids carrying *N*-linked glycans with various structures was developed (2). These building blocks were used to synthesise a library containing >100 glycopeptides and their unglycosylated counterparts. Recently developed "Drop-HILIC" (3)(hydrophilic interaction chromatography) enrichment was further improved to allow for efficient enrichment of both *N*- and *O*-glycopeptides by using DMSO in solubilization buffer. Next, glycopeptide fragmentation characteristics were evaluated using different fragmentation techniques (HCD, SCE-HCD, ETciD, and EthcD) to assess the merit of each method in terms of (a) peptide backbone sequence coverage, (b) glycan composition, (c) proportion of signal in different fragment ion types (e.g., oxonium ions, Y-type ions, and peptide backbone fragment ions, and (d) unambiguous identification of the glycosylation site. We also evaluated and compared various software tools (Byonic, SugarQb) in their ability to identify these glycopeptides reliably. In the end, using the optimized workflow, we are in the process of establishing Human GlycoAtlas containing protein-specific, site-specific N- and O-glycosylation profile of plasma proteins.

To our knowledge, this constitutes the first broad and systematic analysis of the LC-MS/MS properties of glycopeptides using synthetic glycopeptides, allowing optimized bottom-up glycoproteomics experiments. The availability of these glycopeptides and spectra will facilitate the development and improvement of further experimental and computational strategies.

1. (1) Alagesan K, Everest-Dass A, Kolarich D. (2018). 'Isomeric Separation and Characterisation of Glycoconjugates' - Adv Exp Med Biol., 1104: 77-99. (2) Alagesan K, Kolarich D. (2019). 'Improved strategy for large scale isolation of sialylglycopeptide (SGP) from egg yolk powder' – MethodsX ., 9;6:773-778. (3) Alagesan K, Khilji SK, Kolarich D. (2017). 'It is all about the solvent: on the importance of mobile phase for ZIC-HILIC glycopeptide enrichment' - Anal Bioanal Chem., 409(2):529-538.

70

Proteomics Combined with Human Cardiac Organoids to Find New Mechanisms Driving Maturation and Regeneration

James Hudson[1]

1. QIMR Berghofer, Herston, QLD, Australia

The development of *in vitro* human model systems that more accurately represent *in vivo* biology may provide an additional model system for studying biological mechanisms. We have recently developed a high-throughput bioengineered human cardiac organoid (hCO) platform, which provides functional contractile tissue with biological properties similar to native heart tissue. We have developed functional and omics assays to tease apart mechanisms of action for a wide variety of perturbations. This approach may be useful to uncover holistic mechanisms of action of a wide variety of perturbations. in this presentation I will summarize how we have used these approaches to study maturation and regeneration of the heart, which has also opened up new questions to be answered.

71

Characterisation of a putative new metabolic hormone in human plasma

Michelle Cielesh[1], Sameer Kulkarni[2], Marcus M Seldin[3], Jason Low[4], Joel Mackay[4], Richard Payne[2], Samantha Hocking[5], Mark Larance[1]

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

2. School of Chemistry, University of Sydney, Sydney, NSW, Australia

3. School of Medicine University of California Irvine, Irvine, USA

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

5. Central Clinical School, Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia

We have developed protocol called SPEA to allow us to identify and quantify peptides derived from active low abundance small-protein hormones in plasma (Harney *et. al.* 2019. *Mol. Cell. Proteomics*). This has allowed us to explore all components of this fraction using unbiased mass spectrometry-based analysis. Using the SPEA protocol, we detected three peptides from the uncharacterised 8 kDa protein (erusiolin) in human plasma samples from a clinical trial examining the response to intermittent fasting. A highly conserved peptide from erusiolin was significantly increased in abundance after 8-weeks of intermittent fasting. From this, we hypothesised that the abundance of erusiolin was induced by fasting, which we have now tested using human

mixed meal tests and observed decreased abundance after food consumption. The mRNA encoding erusiolin is largely duodenum-specific in both humans and mice and immunohistochemistry analysis of both human and mouse tissues using an erusiolin-specific antibody has demonstrated a staining pattern consistent with expression in enteroendocrine cells. We have used the Quantitative Endocrine Network Interaction Estimation (QENIE) method in mice to identify potential target tissues for erusiolin, which showed the hypothalamus had many transcripts significantly linked to variation in the locus. Strikingly, a significant number of the transcripts were derived from the Prader-Willi Syndrome locus, which is a disease characterised by extreme hyperphagia (over-eating) and subsequent obesity. This leads us to hypothesise that the conserved peptide from erusiolin is secreted by the duodenum into blood plasma during fasting and acts on the hypothalamus to inhibit transcription of the PWS locus, leading to orexigenic (hunger) signals. We have now generated knock-out mice and are currently characterizing their phenotype compared to wildtype littermates for food-intake and related metabolic phenotypes.

72

Strategies to enrich low-molecular weight proteins for proteomic analysis

Parthiban Periasamy[1, 2], Keshava Datta[1], Harsha Gowda[1, 2, 3]

1. Genetics & Computational Biology Department, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

2. Faculty of Medicine, The University of Queensland, Brisbane, Queesland, Australia

3. School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

Global proteomic profiling studies enable identification and quantitation of thousands of proteins in biological samples. In these studies, high molecular weight proteins that are relatively abundant have better chance of being sampled by mass spectrometry than low molecular weight proteins that are less abundant. This is evident in most proteomics datasets where several biologically important protein classes including growth factors and cytokines are poorly represented. Extensive pre-fractionation of protein digests can improve proteome coverage and increase the likelihood of sampling low molecular weight/low abundant proteins. However, such strategies are often not feasible in large-scale proteomics projects as it would require more sample processing and instrument time.

In this study, we evaluated size exclusion chromatography and organic solvent-based protein precipitation methods for isolation/enrichment of low molecular weight/low abundant proteins. We present data from cancer cell lines and biofluids like plasma to demonstrate the utility of this approach to isolate/enrich low molecular weight proteins that are underrepresented in proteomics datasets. This strategy can be employed to investigate changes in expression levels of important classes of proteins including growth factors, cytokines and chemokines that are poorly represented in global proteomic profiling studies.

73

Rapid separation and identification of beer spoilage bacteria by inertial microfluidics and MALDI-TOF mass spectrometry

Mark R Condina[1], Brooke A Dilmetz[1], Sajad R Bazaz[2], Jon Meneses[3], Majid E Warkiani[2, 4], Peter Hoffmann[1]

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

2. School of Biomedical Engineering, University of Technology Sydney, Sydney, New South Wales, Australia

3. Coopers Brewery Ltd., Adelaide, South Australia, Australia

4. Institute of Molecular Medicine, Sechenov University, Moscow 119991, Russia

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS), in combination with Biotyper software, is a rapid, high-throughput, and accurate method for the identification of microbes. Microbial outbreaks in a brewery present a major risk for companies as it can lead to cost-intensive recalls and damage to the brand reputation. MALDI-TOF MS has been implemented into a brewery setting for quality control practices and the identification of beer spoilage microorganisms. However, the applicability of this approach is hindered by compatibility issues associated with mixed cultures, requiring the use of time-consuming selective cultivation techniques prior to identification. We propose a novel, low-cost approach based on the combination of inertial microfluidics and secondary flows in a spiral microchannel for high-throughput and efficient separation of yeasts (*Saccharomyces pastorianus* and *Saccharomyces cerevisiae*) from beer spoilage microorganisms (*Lactobacillus brevis* and *Pediococcus damnosus*). Flow rates were optimised using *S. pastorianus* and *L. brevis*, leading to separation of more than 90% of the *L. brevis* cells from yeast. The microorganisms were then identified to the species level using the MALDI-TOF MS platform using standard sample preparation protocols. This study shows the high-throughput and rapid separation of spoilage microorganisms (0.3–3 µm) from background yeast (5 µm) from beer, subsequent identification using MALDI Biotyper, and the potential applicability of the approach for biological control in the brewing industry.

74

Deglycosylation of glycoproteins with trifluoromethanesulfonic acid (TFMS): unravelling the chemistry allows extension of the protocol to the analysis of released glycan chains using standard proteomics methods

Paul D Veith[1], Richard AJ O'Hair[2], Michael G Leeming[3], Shuai Nie[3], Gavin E Reid[2, 4], Eric C Reynolds[1]

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

2. School of Chemistry, BIO21 Institute, University of Melbourne, Parkville, Victoria, Australia

3. Mass Spectrometry and Proteomics Facility, Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia

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

During the deglycosylation of glycoproteins with TFMS under anhydrous conditions, the glycan chains are cleaved but the protein remains intact. It is therefore used widely to analyse the deglycosylated protein and the site of glycosylation. There have been some reports of detecting N-acetylated sugars in the released glycans, but in general, analysis of the released glycans has not been reported, presumably due to a lack of understanding of the chemistry associated with the deglycosylation process. In our extensive MS analyses of glycoprotein virulence factors of the bacterial pathogen, Porphyromonas gingivalis we discovered that the toluene or other arene which is included in the TFMS reaction reacts with reducing sugars to form a 1,1-diaryl-1-deoxyalditol. The reaction goes to near completion allowing the expected products to be readily detected by MS at abundances much higher than intermediates and side-products. The di-aryl addition allows the modified glycans to bind to C18 columns and hence the reaction products can be washed and concentrated using C18 solid phase extraction kits and analysed by LC-MS. Furthermore, glycan chains that have the di-aryl addition produce a reliable diagnostic ion at m/z 209.13 (C16H17) which aids in their identification. While the chemistry was thought to be specific to the C-1 of reducing sugars, the arenes appear to also react with non-reducing sugars such as fructose and analyses are continuing to determine if the reaction products are reproducibly different. Sugars that are N-acetylated at C-2 are also cleaved and react with arenes producing additional diagnostic ions. The unravelling of this chemistry significantly extends the utility of the TFMS deglycosylation procedure by providing additional information on the released glycans.

75

An integrated mass spectrometry approach towards characterisation of snake venom proteins

Tara L Pukala[1], Chia-De R Wang[1], Emily R Bubner[1]

1. School of Physical Sciences, University of Adelaide, Adelaide, SA, Australia

Snake venoms contain a complex and widely underexplored cocktail of biologically active proteins. Consistent with their broad pharmacological activities and ecological divergence, the proteomic compositions of snake venoms can be highly variable, and efforts to characterise the primary sequences of such proteins are extensive and ongoing. In addition, a significant gap in knowledge exists in terms of higher-order interactions between proteins proposed to modulate venom potency. Here we apply a multifaceted mass spectrometry-based approach to characterise proteins and protein assemblies from selected phylogenetically diverse snake species.

Using a combination of proteomics and native mass spectrometry approaches we have contributed to ongoing efforts to catalogue the protein components of snake venoms, including identification of new higher order protein assemblies. In addition, we have employed other mass spectrometry-based methods including ion mobility-mass spectrometry, MALDI imaging and enzyme activity assays to further characterise the structure and function of venom proteins. This work emphasises the importance of understanding higher-order protein interactions in venoms and the utility of a combined mass spectrometric approach for this task.

76

Detection of intact insulin analogues in post-mortem vitreous humour using forensic toxicology and proteomic tools – application to overdose casework

Rebecca Tidy[1], Nicola Beckett[1], Bianca Douglas[1], Colin Priddis[1]

1. ChemCentre, Bentley, WESTERN AUSTRALIA, Australia

Proteomic applications in forensic science widens the scope of capabilities for forensic toxicology laboratories. However, currently these techniques are underutilised, one example being having the ability to conduct protein analyses for endogenous/synthetic insulins applicable to coronial investigations. Definitive confirmation of an insulin overdose by toxicological analysis of post-mortem biological matrices is rare and challenging. However, the ability to effectively identify and differentiate endogenous insulin and the common synthetic analogues (Apidra, Humalog, Lantus, Levemir and Novorapid) can assist coronial investigations pertaining to accidental or intentional overdoses in both diabetic and non-diabetic populations. This work was aimed at bridging the gap between current analytical capabilities of forensic toxicology laboratories and analysis of 'larger' molecules of forensic interest by utilising both forensic toxicology and proteomic techniques. This was achieved through employing mass spectrometry (MS) based insulin methodology paired with a semi-automated solid phase extraction (SPE) using a uElution plate. Data dependent ($ddMS_2$) acquisition was utilised to screen for intact insulins and parallel reaction monitoring (PRM) acquisition was performed to confirm the presence of insulins in the vitreous humour (VH) of suspected insulin overdose cases. This work reports the first Australian cases where synthetic insulins were confirmed by MS in the VH of Type 1 diabetics who intentionally or accidentally overdosed on their prescription medication Glargine and Aspart. Of particular interest was the proteomic characterisation of Glargine M1, the pharmacologically active metabolite of Glargine. This was performed on an in-house synthesised reference standard to verify the case findings in the absence of a certified reference standard. The results of this

work highlights advancements in forensic coronial toxicology and the promising potential of proteomic analysis in a forensic context.

77

functional-Mass Spectrometry Imaging - mapping the location of enzymes by their function

Brett Hamilton[1], David L Marshall[2], Nicholas R Casewell[3], Robert A Harrison[3], Stephen J Blanksby[2], Eivind A.B. Undheim[1]

1. The University of Queensland, Brisbane, QUEENSLAND, Australia

2. CARF, Queensland University of Technology, Brisbane, Queensland, Australia

3. Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Enzymes are the catalysts for virtually every physiological process and thus provide a key to unlocking aberrant metabolism across a wide range of pathologies; including cancers. High-resolution maps of enzyme activity within tissues would therefore represent powerful tools for elucidating enzymatic functions in health and disease, and have potentially novel diagnostic applications. To-date however, no single imaging modality is able to routinely and simultaneously map the distribution and efficacy of multiple enzymes *in situ*. Mass spectrometry imaging has contributed to the understanding of many biological systems on the basis of showing where the molecules exist, however, the dynamic range is affected by the chemical complexity of the samples in the tissue sections being analysed. Here, we present a new mass spectrometry imaging (MSI)-based method for assaying the spatial distribution of enzymatic activity directly from tissue sections - potentially enabling the visualisation of specific enzymes by their activity rather than their molecular weight, or proteotypic peptide fragment(s). By depositing enzyme substrates on tissue sections that are subsequently analysed by MSI, high-resolution maps of enzyme activity and substrate specificity can be generated. Using phospholipases as model enzyme targets, we demonstrate that functional mass spectrometry imaging (fMSI) represents a new and generalizable method for visualising biological activity across tissues, which is in-principle applicable to any enzyme class with soluble, low molecular weight products.

78

PromarkerD as an immunoaffinity mass spectrometry assay for diabetic kidney disease

Scott Bringans[1], Kirsten Peters[1], Tammy Casey[1], Sarah Thomas[1], Orla Coleman[2], Holger Ebhardt[2], Stephen R Pennington[2], Richard Lipscombe[1]

1. Proteomics International, Broadway, Nedlands, WA, Australia

2. Atturos, Dublin, Ireland

The PromarkerD assay for diabetic kidney disease was originally developed as a multi-protein targeted mass spectrometry assay directly from depleted (top 14 proteins) digested human plasma. This assay has been adapted to an ELISA format and more recently an immunoaffinity mass spectrometry assay. The immunoaffinity method utilises bead-based antibody binding for the specific PromarkerD protein biomarkers in single multiplex capture step. The captured protein biomarkers are reduced, alkylated and digested in situ on the beads with injection onto a microflow LCMS system for targeted mass spectrometry. The results obtained with the immunoaffinity method applied to a 100-person cohort were compared to the original direct plasma digestion method with correlation between the two methods confirmed by Bland and Altman plot analysis. To test the robustness of the process between laboratories the assay was also performed in laboratories in Australia and Ireland using the same samples. The advantages of the immunoaffinity technology developed for this assay are three-fold. Firstly, increased throughput of analysis with a 96 well based robotic handling capable system which also minimises human intervention and handling. Secondly, samples that are injected onto the LCMS system are in a much cleaner form than a crude plasma digest which enhances sensitivity and reduces machine down time due to less frequent source cleaning and minimal LC blockages (in micro flow format). Thirdly, the immunoaffinity method has been designed to be available as a simple technology transfer process to partner laboratories that have LCMS capability. These advantages may make such an immunoaffinity-MS technology approach a superior choice for multi-protein biomarker diagnostic assays.

79

Metabolomic/lipidomic DESI imaging of different cell cultures.

Emmanuelle Claude[1], Hadeer Mattar[2], Emrys Jones[1], Clare Mills[2]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Division of Infection, Immunity & Respiratory Medicine, Manchester Institute of Biotechnology, University of Manchester, Manchester, United Kingdom

Lipidomics can be considered as one of the important fields of metabolomics that provides a comprehensive structural and functional characterization of various lipids. DESI has been currently used as MSI technique for the analysis as well as the identification of lipids in the cell culture.

In this study, we aimed to investigate the distribution and the localisation of lipids among three different cell lines: two of the gut epithelial cell lines (Caco2 and HT29-MTX cells) and one of cancerous basophil cells RBL.

Cells were grown on cover slips and had the media removed for the MSI experiment by a wash in 150 mM ammonium acetate followed by drying step for 15-20 min. The cover slips were directly mounted onto microscope glass slides, using double sided tape and place onto the DESI stage, with no sample preparation. Data were acquired using a Prosolia DESI source which was mounted on a Waters Xevo G2-XS mass spectrometer.

Positive ion mode experiments, at 50 µm pixel size, generated strong signals directly from all three type of cell lines, identified to be mainly glycerophospholipids such as m/z 798.54 identified as PC (34:1), K+ as well as triglycerides. However, difference in

the lipid intensity profiles can be observed between the cell lines. In particular lipids such as m/z 820.54 putatively identified as PC (36:4),K+ was more expressed in the Caco2 cell line whereas m/z 756.53 (PC(P-32:0), K+ was more intense in HT29-MTX cell line. Furthermore, an experiment at 20 µm pixel size showed individual agglomerate of cell with different distribution within the same sample cell line.

Negative mode experiments also showed a very rich level of molecular information for the metabolites and lipidic mass range with the strongest signal at m/z 281.25, putatively identified as oleic acid which was the most abundant in Caco2 cell line.

80

Quantitative, targeted and high-throughput metabolomics workflow of small-volume plasma samples via blood cards.

Stephan Klatt[1, 2], Berin Boughton[3], Brunda Nijagal[4], Soumya Mukherjee[1, 2], Larissa Lago[1, 2], Anne Roberts[1, 2], Christopher Fowler[1, 2], Blaine Roberts[1, 2, 5]

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

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

3. Metabolomics Australia, School of Biosciences, The University of Melbourne, Parkville, Victoria 3052, Australia

4. Metabolomics Australia, Bio21 Institute, University of Melbourne, Parkville, Melbourne, Australia

5. Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

Here, we present a comparative targeted and high-throughput metabolomics workflow of small-volume plasma samples. In detail, we have extracted and analysed 3.5 ul of human plasma from 20 individuals, with the plasma derived from venipuncture and finger-prick blood. Finger-prick derived plasma was collected on specific plasma cards (Noviplex™ Plasma Prep Cards). When air-dried, card-applied plasma is stable and cards can be mailed to the next analytical laboratory, making them highly advantageous in rural areas compared to the standard venipuncture blood draw. Plasma samples were extracted with methanol, AQC-derivatised and analysed on a 6495 QQQ LC/MS instrument (Agilent Technologies) in dMRM mode. The dMRM method contains 77 metabolite targets, whereof most play important roles in neurodegenerative diseases like Alzheimer's disease and Parkinson's disease. Included metabolites are 20 amino acids, six hormones, eight metabolites of the kynurenine pathway, three polyamines and many more. In total, 74 metabolites were successfully detected in the 3.5 ul of starting material, with concentration differences between venipuncture and blood card-derived plasma. Method development was further adapted to a 96-well plate format, enabling sample processing in high-throughput and biomarker validation for the study of neurodegenerative diseases.

81

Complex-centric proteome profiling in one day with SEC-SWATH-MS

Ben Collins[1]

1. Queen's University Belfast, Belfast, ANTRIM, United Kingdom

Cellular functions are rarely attributable to a single molecule but rather to sets of molecules organized into modules such as protein complexes. Methods based on protein correlation profiling (PCP), such as size exclusion chromatography-SWATH mass spectrometry (SEC-SWATH-MS), provide rich information on the state of cellular protein complexes but remain somewhat impractical, as each biological sample requires ~weeks of measurement time.
By employing short (21 minute) gradient analysis we aimed to establish an SEC-SWATH-MS strategy operating at a rate of ~1 biological sample analysed per day while minimizing loss of information. Comparison with a prior study using 2 hour gradients showed we retain the majority of information at the protein and protein complex levels with >10-fold increase in speed. We have applied this method in two perturbed systems, (i) a comparison between the ccl2 and Kyoto variants of the HeLa cell line, and (ii) a comparison between THP-1 cells in monocytic, differentiated macrophage, or LPS stimulated states, revealing a variety of re-organized protein complexes.
The dramatic increase in throughput enables the unbiased characterization of proteome complex organization with minimal information loss, extended scope, and broad applicability.

82

What are we missing? The hidden impact of missing values in proteomics analysis and results

Melissa Davis[1]

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

Analysis of proteomics data requires us to deal with many sources of variability, including both biological and technical variability. However, a further source of variability is introduced by the choice of analytical methods, where different decisions can be made regarding methods for normalisation, imputation, and tests for differential expression. In this talk, I will present our analysis of the most common methods for imputation, and their accuracy and performance on different kinds of data. I will then explore the impact of imputation method selection on the detection of differentially expressed proteins, both using established benchmarking datasets as well as experimental data. These results highlight the importance of decisions regarding analysis methods, and the collective impact such decisions can have on the identification of biologically meaningful variation in an experiment.

83

Global and sequence-targeted purification of intact RBP-RNA complexes

Jeffrey Smith[1, 2], Stephen Wilcox[1], Laura Dagley[1], Jarrod Sandow[1], Rune Larsen[1, 2], Aaron Jex[1], Melissa Davis[1], Andrew Webb[1]

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

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

The lifespan of an RNA molecule from synthesis to degradation, and transcription to translation, is regulated by RNA binding proteins (RBPs). The identification and characterisation of RBPs must therefore define our first steps towards understanding how the post-transcriptional control of gene expression might impact homeostasis or disorder. Yet RNA Binding is a temporal phenomenon and, coupled with its potency of effect, represents a difficult model to study. Recent efforts to discover the corpus of an RNA-binding proteome have focused on the development of biochemical techniques to isolate protein-RNA complexes for mass-spectrometry (MS). To achieve this we developed a complete in-solution method to recover clean, purified RBP-RNA completely intact. Moreover, the stringency of these purifications is enough to permit qualitative confirmation of RNA-binding activity; while the efficiency of extraction can support complex, label-free experiments targeting quantitative changes in dynamic RBP-activity. Finally, we share preliminary results that integrate RNA-hybridisation probes to target RNA species based on common sequence elements and thence identify their protein binding partners.

84

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

Ashleigh L Dale[1, 2], Dylan J Harney[1, 2], Isobel Tenison-Collins[1, 2], Joel A Cain[1, 2], Mark Larance[1, 2], Stuart J 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, off-line peptide size exclusion chromatography and a hybrid MS2-MS3 fragmentation strategy was developed and employed to globally define PPIs in C. jejuni. This enabled the first non-binary and comprehensive analysis of the interactome of this organism. A total of 745 proteins were identified to partake in 1,133 unique and significant PPIs governed by 3,316 unique Lys-Lys residue contacts. Our XL-MS approach successfully covered 46% of the predicted proteome of C. jejuni and 57.3% of the proteome as previously identified by 'bottom-up' proteomics. Additionally, multiple proteases and enrichment strategies were compared and this allowed for increased depth and coverage of the proteome, as well as the identification of less abundant and harder to detect PPIs and cross-links, particularly membrane proteins. Interrogation of the XL-MS dataset returned 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.

85

Comprehensive Quantitative Stability Assay (CQSA) system for defining the molecular resolution of protein-compound interactions in *Mycobacterium tuberculosis*.

Robert L Moritz[1]

1. Institute For Systems Biology, Seattle, WA, United States

Tuberculosis is an old but re-emerging global health threat caused by mycobacteria belonging to the *Mycobacterium tuberculosis* family. With 1/3 (~2.5-billion) of the world's population infected with Mtb and at least 9M new cases, resulting in 1.7M deaths annually, Mtb is by far the largest burden of disease and a global health emergency affecting not only the developing world. Exacerbating this epidemic is the threat of increasing multidrug-resistant tuberculosis (MDR-TB) which will ultimately cause a further increase of TB cases amongst areas of the population currently unaffected creating a massive social impact. In more than over 50-years of research, Bedaquiline, is the only approved fast-tracked antibiotic, despite its severe complications, for the treatment of MDR-TB. These complications include nausea, arthralgias, hemoptysis, hyperuricemia, rash, extremity- and chest-pain, cardiac-QTc prolongation and more striking - unilateral-deafness. The ability to identify early-stage compounds with a better understanding of action at the molecular level rapidly provides information to drug development programs that can ultimately benefit patients in poor resource settings the most.

Alterations in physicochemical properties of proteins in cells, lysates, organelles, and living organisms by deliberate stability modulation probed by comprehensive-quantitative proteome-wide scale measurements is such a tool that can greatly inform

direct actions of molecules and positively define intended targets and possible off-targets in an unbiased format. We have developed the Comprehensive-Quantitative-Stability-Assay (CQSA) system that is a state-of-the-art, high-throughput proteomics method combining power, speed and molecular resolution to screen proteomes such as pathogens and their human hosts for protein-compound interactions over a wide range. We have used the CQSA system to specifically identify protein interactions with long-term known drugs in TB therapy and with promising lead compounds in TB drug development pipelines to provide information that will assist in the decision making of compound progression or its halt in drug development. The CQSA-system has confirmed target-protein engagement with high-specificity and has identified novel off-target effects where we have definitive x-ray crystallography structural data characterizing the binding and elucidating hitherto pleiotropy of these compounds.

91

Towards the generation of a comprehensive proteome map for *Phytophthora cinnamomi*, the causal agent of native dieback

Christina Andronis[1], Tammy Casey[2], Giles Hardy[3], James Hane[1], Richard Oliver[1], Richard Lipscombe[2], Kar-Chun Tan[1]

1. Curtin University, Perth, WA, Australia

2. Proteomics International, Perth, WA, Australia

3. Centre for phytophthora Science and Management, Murdoch University, Perth, WA, Australia

Phytophthora cinnamomi is a pathogenic oomycete that poses a significant threat to global biodiversity. It causes loss of vegetation/ crops and tree death by attaching to the fine-feeder roots and releasing apoplastic and cytosolic effectors. The destructive nature of these effectors during plant-pathogen interactions drives a complex array of defence responses and when successful, the pathogen can cause plant death. *Phytophthora cinnamomi* exhibits various life stages, which enable it to survive through harsh environments. This has contributed to its success as a pathogen as it takes advantage of weakened potential hosts

However, the molecular mechanism of *P. cinnamomi* pathogenicity is not well-understood. In order to better understand the pathogen, a proteomic approach was used to dissect sub-cellular proteomes of *P. cinnamomi*. We were able to extract high quality intracellular, secreted, zoospore and membrane proteomes of *P. cinnamomi* and perform analysis using gel-free label-free shotgun proteomics in addition to iTRAQ. We predicted functional annotations using gene ontologies for deferentially-abundant proteins, with a view to understand pathogen proteins production during host infection. A proteomic pipeline was also applied to non-quantitative data to identify gene models that were not predicted using traditional methods. This provided us with a second set of draft gene models that were subsequently defined via predicted gene ontology terms. This data is working toward a comprehensive reference proteome for the pathogen.

92

Interaction of Common Antibiotics with Human Serum Albumin as Determined by Hydrogen Deuterium Exchange and Chemical Cross Linking Mass Spectrometry

Ching-Seng Ang[1], Michael G Leeming[1], Shuai Nie[1], Swati Varshney[1], Caryn Hepburn[2], Nicholas A Williamson[1]

1. Bio21, University of Melbourne, Melbourne, VIC, Australia

2. Waters Australia, Rydalmere, NSW, Australia

Human serum albumin (HSA) is the most abundant protein in plasma and acts as one of the main carriers in the circulatory system. HSA participates in binding and transportation of a broad range of ligands such as fatty acids, drugs, hormones and other biologically important components that are essential for normal functioning of the body. HSA possesses multiple binding sites and is made up of three homologous domains (I, II and III) and each domain is made up of two subdomains (A and B) connected by random coil. There are a total of 9 binding pockets (FA1-FA9) with FA7, also named as Sudlow's binding site, being the major drug binding site.

The majority of studies on HAS-ligand interactions are conducted using theoretical techniques such as classical molecular dynamics (MD) simulations and molecular docking. We have utilized the power of hydrogen deuterium exchange in combination with chemical cross-linking mass spectrometry approaches to perform epitope mapping at the molecular level. This study provides an important insight by mapping exact binding location of the drug ligand to HSA at near native conditions. We have demonstrated the capability of this approach which could help in new antibiotic designs with effective pharmacological properties.

93

Investigating proteome changes between primary and metastatic cutaneous squamous cell carcinoma using SWATH mass spectrometry

Ali Azimi[1, 2], Kitty Li[3], Jennifer Kim[4], Pablo Fernández-Peñas[1, 2]

1. The Department of Dermatology, Westmead Hospital, The University of Sydney, Westmead, NSW 2145, Australia

2. Centre for Translational Skin Research, The University of Sydney, Westmead, NSW, Australia

3. School of Mathematics and Statistics, The University of Sydney, Camperdown, New South Wales, Australia

4. Department of Tissue Pathology and Diagnostic Oncology, institute of clinical pathology and medical research (ICPMR), Westmead Hospital, Westmead, New South Wales, Australia

Cutaneous squamous cell carcinoma (cSCC) is a common malignancy worldwide, responsible for 20% of all skin cancer cases, and the first as the cause of death from keratinocytic carcinomas. Around 5% of primary cSCCs metastasize to regional or distant body parts, leading to a 5-year survival rate of only 25-35%. Currently, clinical and histopathological assessment is used for the diagnosis of metastatic cSCC, and this is when the disease is already metastasised. Therapies for patients with metastatic cSCCs are lacking due to poor knowledge of the molecular alterations that drive this metastasis. In this paper, we used liquid

chromatography coupled with SWATH mass spectrometry workflow to analyse formalin-fixed and paraffin-embedded samples of primary (n=20) and metastatic cSCC (n=25) for the identification of protein biomarkers and molecular alterations discriminating between the lesion groups. We quantified 5037 proteins across all the samples studied of which 19 proteins including TBD2A, GCP60 and PDLI3 were increased and 11 proteins including CBPA3, SAMP and DMKN were decreased respectively in metastatic cSCC relative to the primary phenotype (adj. p-value < 0.05; fold change ± 1.8). The proteomics data also separated the two lesion groups on principal components analysis. Bioinformatics analysis revealed that Ras protein signalling, small GTPase mediated signalling and regulation of cytoskeleton organisation were associated with cSCC metastasis. Overall, results from this study showed excellent potential to discriminate between primary and metastatic cSCC subtypes, facilitating new diagnostic and therapeutic strategies.

94

Expression of *SCN2A* protein in different mouse models of *SCN2A* encephalopathy and after antisense oligonucleotide treatment

Todd Blackburn[1], Blaine Roberts[1], Anne Roberts[1], Melody Li[1], Nikola Jancovski[1], Alex Nemiroff[2], Kelley Dalby[2], Paayman Jafar-Najad[3], Frank Rigo[3], Nicholas Williamson[4], Snezana Maljevic[1], Steven Petrou[1, 2]

1. Florey Institute, Parkville, VIC, Australia

2. RogCon Biosciences, Miami, Florida, United States

3. Ionis Pharmaceuticals, Carlsbad, California, Unite States

4. Mass Spectrometry and Proteomics, Bio21 Institute, Parkville, VIC, Australia

OBJECTIVES:

Aim 1): To quantify A) total SCN2A protein and B) mutant SCN2A protein using selected reaction monitoring (SRM) in 3 heterozygous mouse models of SCN2A encephalopathy, each representing a unique clinical phenotype seen in the SCN2A disease population.

Aim 2): To quantify the change in total SCN2A protein, mutant SCN2A protein, and several other proteins in response to an antisense oligonucleotide (ASO) treatment in a gain-of-function (GoF) mouse model of SCN2A encephalopathy using SRM.

BACKGROUND: SCN2A encodes a voltage-gated sodium channel and *de novo* variants are the cause of autism and a spectrum of developmental and epileptic encephalopathies (DEEs). There is growing evidence to support that the different phenotypes correlate with the resulting functional consequence of the mutations, i.e. whether the protein is loss or gain-of-function.

DESIGN: Brain samples were collected from WT mice and mice heterozygous for SCN2A R854Q, S1759R, or R1883Q at postnatal (P) day 14. R1883Q mice were treated with an ASO that down-regulates SCN2A mRNA or a scrambled, non-specific ASO. Samples were digested with trypsin and analysed on the *Agilent 6495 triple quadrupole (QQQ)* mass spectrometer coupled to an *Agilent 1290 Infinity II* liquid chromatography unit.

RESULTS: No significant difference was measured in total SCN2A levels between R1883Q and WT mice injected with the control ASO. R1883Q mice treated with the SCN2A down-regulation ASO had a 3.83 fold reduction in total SCN2A. Tau was significantly increased in R1883Q mice, but was not stabilised to WT levels after ASO treatment.

CONCLUSIONS: Using the Agilent 6495 QQQ instrument we were able to detect protein concentration in a SCN2A disease model and quantify changes in response to ASO treatment. The ability to detect and quantify mutant protein using mass spectrometry shows immense potential as a tool for measuring target engagement for disease-modifying precision medicine treatments.

95

SWATH-MS proteomics reveals time-dependent acid stress response in *Campylobacter jejuni*

Utpal Bose[1], Stanley Chen[2], James Broadbent[1], Chawalit Kocharunchitt[3], John Bowman[3], Narelle Fegan[2], Lesley Duffy[2], Michelle Colgrave[1]

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

2. Agriculture and Food, CSIRO, Coopers Plains, QLD, Australia

3. Tasmanian Institute of Agriculture, University of Tasmania, Hobart, Tasmania, Australia

Campylobacter is the leading cause of food-borne gastrointestinal disease worldwide due to its high prevalence and main reservoir in chicken. Antimicrobial agents have been used during chicken processing in Australia, including chlorine and citric acid. Interaction between chlorine and organic matter during immersion chilling produces by-products that are implicated as human carcinogens. Peracetic acid (PAA) has been considered as an alternative to chlorine in the poultry industry. However, the time-dependent effectiveness of PAA in the Australian poultry industry is uncertain and little is known how *Campylobacter* survives and tolerates PAA treatment. In this study a *C. jejuni* poultry strain was investigated to identify the stress response, adaptive changes and survival mechanisms upon the sub-lethal PAA treatment (80 ppm) across four treatment time points (0, 15, 30, 45 min) utilising sequential window acquisition of all theoretical fragment-ion spectra-mass spectrometry (SWATH-MS) with a SCIEX TripleTOF 6600 MS.

A total of 2,146 peptides mapping to 591 proteins were quantified at a <1% false discovery rate. Pairwise comparisons between the 0- and the 15-, 30- and 45-min timepoints revealed proteome-wide alterations in the order of ~11%, ~16% and ~17% proteins, respectively. Between 0 and the 45 min treatment yielded 41 up-regulated proteins (~7%), while 63 (~11%) proteins were down-regulated after 45 min. Gene Ontology (molecular function) revealed that ATP-binding and oxidoreductase activities were up-regulated, whilst hydrolase and transferase activities were down-regulated. The 30- and 45-min treatment showed ~28% overlap in proteins, whilst 2, 5 and 8% of proteins were uniquely produced during 15-, 30- and 45-min treatments, respectively.

Overall time-dependent PAA treatment on *C. jejuni* delivers an informative portrait of its survival mechanism and presents opportunities to develop better PAA-based poultry processing strategies for commercial application in the future.

96

A High Throughput approach to identifying compound inhibitors of mRNA processing and export

Kirstyn Carey[1], Tobias Williams[1], Vihandha Wickramasinghe[1]

1. Peter Maccallum cancer centre, Parkville, VICTORIA, Australia

Once considered to be a constitutive step of the gene expression pathway, work from our laboratory has demonstrated that export of mRNA from the nucleus to the cytoplasm can be highly selective, giving priority to some mRNAs over others. Perturbations in this pathway result in dysregulated cellular proliferation and genome instability, demonstrating that selective mRNA export regulates processes known to be drivers of malignant transformation. There is accumulating evidence that mRNA export dysregulation contributes to cancer development. Furthermore, transcriptionally driven cancers are heavily reliant on the gene expression pathway to ensure production of oncogenic protein isoforms and to sustain proliferative growth. This creates novel vulnerabilities in cancer cells that can be therapeutically exploited using compounds that affect the gene expression pathway, such as nuclear export of mRNA. Compounds inhibiting all other steps of the gene expression pathway are promising therapeutic candidates, and many have reached clinical trials. Thus, development of RNA export inhibitors offers a unique opportunity to develop innovative, targeted cancer therapeutics.

We have developed and optimised a sensitive, high-throughput assay to screen 25,000 natural, “drug-like” compounds examining their effects on cell death and RNA localisation within breast cancer cells. Coupled with advanced image analysis, it is possible to reproducibly detect and quantify perturbations of mRNA processing and export. We have demonstrated the first ever compounds inhibiting mRNA export. Immunofluorescence analysis demonstrates compounds disrupt normal localisation of mRNA export machinery components. Furthermore, many compounds show selective toxicity to cancer cells over non-transformed cells *in vitro*.

Future work focuses on mechanistic characterisation of compounds and target identification, integrating chemical biology and proteomic strategies. Comprehensive characterisation of these compounds will facilitate our understanding of mRNA processing and export and will hopefully lead to development of pharmacologically relevant compounds which can eventually advance to clinical trials as cancer therapeutics.

97

The E3 ubiquitin ligase SCF$^{(cyclin\text{-}F)}$ complex regulates caspase-3, a key executor of the apoptotic cell death program

Flora Cheng[1], Alana De Luca[2], Stephanie Rayner[1], Jennilee Davidson[1], Marco Morsh[1], Ian Blair[1], Roger Chung[1], Albert Lee[1]

1. Macquarie University, Macquarie Park, NSW, Australia

2. University of Tasmania, Tasmania, Australia

CCNF is a causative gene in familial and sporadic cases of FTD/ALS and encodes cyclin F, a component of an SCF$^{(cyclin\ F)}$ E3 ligase complex involved in ubiquitin-mediated proteolysis. Previous studies have shown that expression of the cyclin F^{S621G} mutation leads to defective protein degradation, motor axonopathy, and signature features of ALS pathogenesis in vitro and in vivo [1-3]. In this study, we used an unbiased label-free quantification proteomics strategy to uncover changes to biological processes and cellular pathways of other mutations (K97R, S195R, S509P and R574Q)[4]. Focusing on the apoptosis pathways that were predicted to be activated in a few mutations, we investigated the effect of cyclin F^{S621G} on its E3 ligase Lys48-specific ubiquitylation of caspase-3, a key mediator of neuronal apoptosis.

Six mutations were selected for the label free proteomics analysis, four fALS (K97R, S195R, S509P, and S621G), and one database SNP (R574Q) were expressed in human cell lines. Three mutants (K97R, S621G and S195R) were predicted to show activation of the apoptosis signalling pathways compared to wild-type. Proximity-based labelling and immunoprecipitation of cyclin F wild-type and cyclin F^{S621G} demonstrated direct interaction with caspase-3, and in-vitro ubiquitylation of recombinant caspase-3 by cyclin F revealed modulation of caspase-3 activity that is influenced by the mutational status of the cyclin F disease variants. The relationship between cyclin F disease variants and caspase 3 indicate a potential mechanism by which ALS mutations may contribute to increased neuronal death leading to disease pathogenesis. Our data presents interesting insight into survival pathway changes caused by ALS-associated mutations in *CCNF* that may contribute to our new understanding of cyclin F in neuronal proteostasis.

1. Hogan, A.L et al., Expression of ALS/FTD-linked mutant CCNF in zebrafish leads to increased cell death in the spinal cord and an aberrant motor phenotype. Hum Mol Genet, 2017. 26(14): p. 2616-2626 DOI: 10.1093/hmg/ddx136.
2. Lee, A et al., Casein kinase II phosphorylation of cyclin F at serine 621 regulates the Lys48-ubiquitylation E3 ligase activity of the SCF(cyclin F) complex. Open Biol, 2017. 7(10) DOI: 10.1098/rsob.170058.
3. Lee, A et al., Pathogenic mutation in the ALS/FTD gene, CCNF, causes elevated Lys48-linked ubiquitylation and defective autophagy. Cell Mol Life Sci, 2017 DOI: 10.1007/s00018-017-2632-8.
4. Williams, K.L et al., CCNF mutations in amyotrophic lateral sclerosis and frontotemporal dementia. Nature Communications, 2016. 7: p. 1-8

98

MASS SPECTROMETRY IMAGING STUDY OF LIPID METABOLITES IN THE ADULT MOUSE TESTIS

Sheba Jarvis[1], Charlotte Bevan[1], Lee Gethings[2], Emmanuelle Claude[2]

1. Imperial College London, Hammersmith hospital, London, UK

2. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

The testis is a complex and heterogeneous organ with endocrine and exocrine functions. The most abundant testicular lipids, seminolipids, (accounting for 3% of total lipids) are selective to maturing germ cells and crucial for male fertility[1,2]. Furthermore, phosphatidylcholines (PCs) play a major role on the structure and function of testes[3]. A comprehensive understanding of the overall lipid metabolites localisation is critical to understand testicular function and to this day, studies have been limited to specific lipid classes. With the technical improvement of ionization sources such as matrix-assisted laser/desorption ionization (MALDI) and desorption electrospray ionization (DESI-MSI), it is now possible to achieve pixel sizes that specifically distinguish the complex testicular anatomy.

Here we used a multi-modal mass spectrometry imaging (MSI) approach with MALDI-MSI and DESI-MSI as complementary approaches to improve the characterization of testicular lipid metabolites.

In both modes of ionisation using MALDI or DESI, distinct lipids profiles could be distinguished between Leydig/blood vessels, more mature germ cells and Sertoli cells/early germ cells.

In negative mode, the major lipid detected was a seminolipid, *m/z* 795.53 identified as C16:0-alkyl-C16:0-acyl and localised in the germ cells. Other seminolipids were also detected at lower intensities such as *m/z* 809.51 which was more intense depending on the maturity of the germ cells. The second most intense phospholipid detected was *m/z* 885.55 (PI (38:4, H^-) that is highly localised in the Leydig and blood vessels.

In positive mode, mainly glycerophospholipids were detected. PC (36:1) and PC (38:5) have both been detected with a sodium and potassium cation, strongly localised in the Leydig/blood vessels cell types. PC (34:1) and PC (36:4) (Na^+ and K^+) ions were mainly localised in the Sertoli cells/early germ cell type and PC (38:5) and PC 38:6) were localised in the more mature germ cells.

1. 1:Rodemer C; et al. Human Molecular Genetics; 12: 1881-1885 2:Vos JP; et al. Biochim, Biophys Acta 1211:1250149 3:Siangcham T; et al. PLoS ONE 10(3): e0120412. doi: 10.1371/journal. Pone.0120412

99

Chronic obstructive pulmonary disease: is autoimmunity a contributing factor?

Penelope V Dalla[2, 1], David Chapman[2, 1], Brian Oliver[2, 1], Matthew Padula[1]

1. University of Technology Sydney, Sydney

2. Woolcock Institute of Medical Research, Glebe, NSW, Australia

Introduction: Recently there have been indications that autoimmunity plays a role in the development and progression of chronic obstructive pulmonary disease (COPD). Previous studies have demonstrated that COPD patients have higher concentrations of autoantibodies which target a more diverse panel of auto-antigens. The aim of this project is to identify the antigenic targets of the autoantibodies in the serum of COPD patients using a proteomics approach to better understand disease progression.

Methods: Serum and tissue from COPD and donor (non-COPD) patients were obtained pre lung transplant at St Vincent's Hospital, Sydney, Thoracic Department. Serum antibodies were isolated using immunoprecipitation and samples prepared for liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis using tryptic shotgun preparation methods.

Results: Using the mass spectrometry analysis, we identified proteins associated with the antibodies in the serum derived from all COPD and donor samples. These proteins are found in different abundance in the serum derived from COPD compared to healthy donor samples and have very diverse functions.

Conclusion: Different antigens were found in the patient cohorts studied in this project. Further examination is required to determine whether any of the autoantigens contribute to disease development and progression.

100

Measurement of protein losses during sample preparation

Chris Desire[1], Clifford Young[1], Jan Degenhardt[1], Brooke Dilmetz[1], Mark Condina[1], Peter Hoffmann[1]

1. University of South Australia, Mawson Lakes, South Australia, Australia

Sample preparation is a critical part of the proteomic analysis workflow. Such protocols invariably involve protein solubilisation and extraction, followed by chemical reduction and alkylation, enzymatic digestion, and peptide clean-up steps. Due to the multistage nature of proteomic sample preparation, sample losses incurred after each step accumulate to diminish protein recovery and analytical sensitivity, generating biases regarding the presence of specific protein groups.

The severity of protein loss can be difficult to measure as protein assays suffer from interference caused by reagent use. It is known that the bicinchoninic acid assay is affected by reducing agents, while the Bradford assay is sensitive to commonly used surfactants. Additionally, protein losses are generally evaluated after the clean-up process, offering no insight as to the extent of loss occurring in each individual step.

Here we present an overview of the use of a reported fluorescent tryptophan assay for the investigation of different aspects of the sample preparation process for bovine serum albumin (BSA), as a model protein, with the goal of achieving an improved understanding of protein losses across concentration ranges, and to allow for further optimisation of each individual step in the workflow. Additionally, the influence of common cell lysis conditions (e.g. SDC, SDS, RIPA) on protein loss is investigated, and their impact on later stages of sample preparation will be a particular focus of the work.

101

High-confidence identification of protein methylation in *Giardia duodenalis*, a neglected gastrointestinal parasite

Samantha J Emery-Corbin[1], Joshua J Hamey[2], Marc R Wilkins[2], Aaron R Jex[3, 1]

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

2. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW, Australia

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

Protein methylation coordinates epigenetic gene expression and drives cell development. For many human parasites, limited drug classes and resistance to frontline drugs mean new avenues for inhibiting parasite life-cycles are urgently required. Methyltransferase enzymes represent prime drug targets in *Giardia*— the most common human gastrointestinal parasite causing diarrheal disease in 200 million people worldwide. *Giardia* is a unique model for protein methylation; this deep-branching protist has a diverged lysine methylome (K-Me) and is the first eukaryote with no arginine methylome.

We acknowledge large-scale methyl-site identifications using LC-MS/MS are subject to high false discovery rates (FDR) in methylpeptide spectrum matches (methyl-PSMs)[1]. Ideally, 'true' methyl-sites are identified by orthogonal methyl-peptide identification through isotopic labelling of methylation sites via heavy-methyl SILAC. These are not feasible for non-model organisms and parasites which have limited *in vitro* cultures and/or undefined media. Therefore, we implemented new methyl-site confidence-filtering pipelines in *Giardia*. Filtering is implemented post-database searching with Maxquant[2], complimented by avoiding methanol- and SDS-based sample preparation (which introduce methyl-artefacts[1]), and high-mass accuracy peptide fragmentation with higher-energy collisional dissociation (HCD). Filtering removes common sources of false positive methyl-PSMs, and retains sites with robust MS/MS evidence. We validated methyl-site filters with mouse controls, demonstrating filtering enriches known ('true') mouse methyl-sites.

These confidence filters identified over 200 K-Me sites in 160 proteins in *Giardia*. These *Giardia* K-Me proteins are mostly species- and lineage-specific with only 58/160 and 40/160 shared orthologs in *H. sapiens* and *S. cerevisiae*, respectively. Many K-Me proteins in *Giardia* specific gene-families were associated with cytoskeletal regulation. Significant enrichment for coiled-coil features demonstrated many K-Me sites specifically occurred within coiled-coil heptads in solvent-facing positions, with intriguing implication for regulation of this versatile protein folding motif.

1. Hart-Smith G, Yagoub D, Tay AP, Pickford R, Wilkins MR. Large Scale Mass Spectrometry-based Identifications of Enzyme-mediated Protein Methylation Are Subject to High False Discovery Rates. Mol Cell Proteomics 15, 989-1006 (2016).
2. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11, 2301-2319 (2016).

102

OMIC Investigation of Chondrocytes' Response to varying mechanical load – Assessment of aging and impact stress

Scott Higgins[1], Christopher J Hughes[2], Lee A Gethings[2], Paul Townsend[1]

1. Faculty of Biology, Medicine & Health, University of Manchester, Manchester, United Kingdom

2. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

Osteoarthritis from a clinical standpoint is defined as the progressive loss of articular cartilage accompanied by subarticular bone remodelling and is a highly prevalent disease which affects as many as 27 million people in the US. Currently, there are no effective treatments available and often involve surgical intervention. Therefore, a systems approach is required to analyse the response of resident cartilage cells to physiological and pathophysiological load. In order to gain additional understanding of the underlying biology, an extensive proteomic profile of the chondrocyte cell response was investigated.

Porcine cartilage tissue was used as a model organism to assess various loading regimes for introducing impact stress prior to the analysis of human derived samples. Proteins were then extracted, and data acquired using an ion mobility, data independent approach (HDMS^E). Samples were acquired in a random fashion and as technical triplicates. The data were processed using Progenesis QI for Proteomics and searched against a Uniprot *Porcine specific* database, with 1% FDR. Curated results were interrogated using Ingenuity to derive pathways of biological significance. The protein markers identified provided additional insight into the characteristics of cartilage and the underlying biology. These markers are shown to be indicative of impact stress and potentially markers of osteoarthritis. Future work will focus on transitioning these markers into a high throughput assay for potential clinical utility.

103

Insights into the proteomics of abiotic stresses in rice

Fatemeh Habibpourmehraban[1], Sara Hamzelou[1], Brian Atwell[2], Paul Haynes[1]

1. Department of Molecular Science, Macquarie university, Sydney, NSW, Australia

2. Department of Biology, Macquarie university, Sydney, NSW, Australia

Rice as a cereal crop species is a significant part of the staple diet for half of the world's population and is grown in every continent apart from Antarctica. Reduced rice crop productivity is mostly attributed to various abiotic stresses, which are a major area of concern when we are faced with increasing food requirements. The major abiotic stresses include drought, oscillating temperature and high salinity, all of which negatively influence the yield of crops.

Plants respond to multiple stresses differently from how they respond to individual stresses, activating a specific program of gene and protein regulation relating to the exact environmental stress. Rather than being additive, the presence of an abiotic stress can have the effect of reducing or enhancing the susceptibility to other abiotic stresses. In this project, we are examining the combined effects of drought, temperature and salt stress in different permutations.

As a baseline study, we have investigated the proteomic response to drought stress in eight different rice varieties; Nipponbare, Doongara, IAC1131, Mahsuri, Reiziq, N22, IR2006-P12 and IDSA77. Plants were grown in a temperature controlled greenhouse to late vegetative stage and then exposed to drought stress, with leaf samples collected at the point of severe stress, and also following recovery. For initial studies of combination of multiple abiotic stresses, Nipponbare plants were subjected to individual stress condition including salt, drought and temperature oscillation. After that, the proteome response of the same genotype was studied under multiple abiotic stress condition. Proteins were extracted from 3-week leaf tissue, with trypsin in-solution digested peptides separated and identified using nanoflow reversed-phase liquid chromatography – tandem orbitrap mass spectrometry on a Thermo Q-exactive. Peptides and proteins were identified using GPM software. Our results will be useful for the design of agronomically relevant strategies for the development of broad spectrum stress tolerant crops.

104

Crosstalk between protein methylation and phosphorylation in yeast and human

Joshua Hamey[1], Daniela-Lee Smith[1], Ryan J Separovich[1], Marc R Wilkins[1]

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

Post-translational modifications (PTMs) are essential controllers of protein function within the cell. Phosphorylation and methylation, two of the most common PTMs, both regulate many cellular processes including cellular signaling, transcription, translation and chromatin remodeling. However, the extent of crosstalk between these PTMs remains largely unexplored. We have focused on two crucial aspects of PTM crosstalk: discovery of modifications pairs and characterisation of modifying enzyme specificity. First, through PTM enrichment, heavy methyl SILAC and mass spectrometric analysis in Saccharomyces cerevisiae, we uncovered new potential instances of crosstalk between methylation and phosphorylation. Phospho-methyl modification pairs mostly occurred on proteins associated with transcription and translation – these are highly connected proteins within the interactome. In particular, disordered glycine-arginine rich regions of proteins were subject to arginine methylation and serine phosphorylation. We showed that one such protein, yeast fibrillarin, is subject to crosstalk at serine-arginine-glycine-glycine (SRGG) motifs - methylation by the Hmt1 methyltransferase and phosphorylation by the Sky1 kinase were shown to be mutually exclusive. Phospho-methyl modification pairs were also discovered on the highly methylated elongation factor 1A (eEF1A) protein. This protein is targeted by five, highly specific methyltransferases, and novel phosphorylation events were discovered in close proximity to known lysine methylation sites. Moreover, deletion of eEF1A methyltransferases revealed the degree of crosstalk between these modifications. We also explored phospho-methyl crosstalk by investigating the mechanisms that underpin methyltransferase substrate recognition. To this end we developed and implemented methyltransferase motif analysis by mass spectrometry (MT-MAMS), a method to systematically characterise the amino acid sequences targeted by methyltransferases. Application of MT-MAMS to five yeast and human methyltransferases revealed known and likely crosstalk. In particular, we showed that Hmt1 is intolerant of acidic residues immediately upstream of the target arginine, reflective of its crosstalk with phosphorylation at these positions. As the roles of methylation and phosphorylation in the cell are further uncovered, it will be essential to determine the role that their crosstalk plays in determining their functions.

105

Proteome response to drought stress of three species of rice from different geographic regions

Sara Hamzelou[1], Karthik Kamath[2], Farhad Masoomi-Aladizgeh[3], Matthew M. Johnsen[1], Brian J. Atwell[3], Paul A. Haynes[1]

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

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

3. Department of Biological Sciences , Macquarie University, Sydney, NSW, Australia

Drought is a serious threat to global food security. The Asian rice cultivar *O. sativa* cv. Nipponbare is a domesticated *sativa* cultivar with a low level of drought-tolerance. The wild species *O. australiensis* and the cultivated *O. glaberrima* are considered to contain rich untapped reservoirs of valuable genes which have great potential for use in successful marker assisted breeding approaches. In this study, we analysed the proteome of leaves of three species of rice grown principally in Asia, Africa and Oceania. Plants were exposed to drought stress, and leaves from stressed and untreated control ones were harvested for protein extraction, followed by label-free quantitative shotgun proteomics

Proteins were extracted from leaf tissue using trichloroacetic acid – acetone extraction and precipitation. In-solution digested peptides were separated and identified using nanoflow reversed-phase liquid chromatography – tandem orbitrap mass spectrometry on a Thermo Q-Exactive. Peptides and proteins were identified and quantified using MaxQuant.

Oryza australiensis is tolerant to variable water supply in nature and in our controlled experiments. Our results showed no significant difference in water potential of *O. australiensis* leaves under drought stress, which indicates it is able to retain more water in leaf cells under drought conditions. However, this response was not observed in *O. glaberrima* which is also known as a stress tolerant species. Our analysis has shown that the identified differentially expressed proteins were assigned to various biochemical pathways, illustrating the different strategies of each species to modulate molecular responses to overcome water stress. A majority of proteins increased in abundance in stress conditions in *Oryza australiensis* were associated with photosynthesis and carbohydrate biosynthesis. Proteins participating in leaf starch biosynthesis in *O. australiensis* can be used in selective breeding to supplement the bio-engineering of improved photosynthetic efficiency, thus potentially contributing to improved crop yield in cultivated rice.

106

A draft map of the Rice plant proteome

Yunqi Wu[1, 2], Karthik Kamath[3, 1], Matthew McKay[1, 2], Ardeshir Amirkhani[1, 2], David Cantor[1, 2], Gene Hart-Smith[1, 2], Dana Pascovici[1, 2], Mohsen Rahiminejad[4], Armin Soleymaniniya[4], Paul Haynes[2], Ghasem Hosseini Salekdeh[4], Mehdi Mirzaei[1, 2]

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

2. Faculty of Science and Engineering, Sydney, NSW, Australia

3. Faculty of Science, Sydney, NSW, Australia

4. Department of Systems Biology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research, Education, and Extension Organization (AREEO), Karaj, Iran

Rice (*Oryza sativa*) is one of the most widely cultivated crops on the planet. It has the smallest genome and greatest biodiversity of all cereals and similarities in gene order (synteny), gene structure, gene sequence, and gene function, making it an ideal model cereal. The rice plant genome contains 35,678 predicted protein-encoding genes; however, over 89% of these genes lack confidently mapped protein level evidence. To address this gap and to expand on knowledge of the rice proteome, we performed high-resolution mass spectrometry based proteomics of rice plant. In-depth proteomic analysis of 30 distinct anatomical parts of the rice plant, including regenerative, reproductive and undifferentiated organs in different developmental stages, resulted in identification of 13,371 protein-coding genes (FDR 1%) with at least two unique peptides with the length of <30 & >9 amino acids (gold identifiable proteins referred to HPP guideline v.3). Error-tolerant sequence database searches, which search for all entries in the Unimod database in a systematic manner, revealed particularly widespread evidence for the identification of canonical proteins alongside sequence variants. These results extend the identified proteome coverage of rice to about 40% of the total annotated protein-coding genes in rice plants. This compendium of the rice proteome (to be available as an interactive web-based resource soon), as a community resource, can now drive disruptive innovation in rice cultivar improvement and propel rice research and application through improved understanding of rice biology.

107

Quantitative proteomic analyses of molecular response to the RNA polymerase I transcription inhibitor CX-5461 in olaparib-resistant ID8 mouse ovarian cancer cells

Jian Kang[1], Huy Nguyen[2], Michael Hendley[2], Natalie Brajanovski[1], Ching-Seng Ang[3], Richard Pearson[1], Elaine Sanij[1]

1. Division of Research, Peter Maccullum Cancer Centre, East Melbourne, VIC, Australia

2. Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, VIC, Australia

3. Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Melbourne, VIC, Australia

The poly (ADP-ribose) polymerase inhibitor (PARPi) olaparib have been approved by FDA for the treatment of homologous recombination (HR) DNA repair–deficient ovarian cancer. However, acquired resistance to PARPi is common and develops through multiple mechanisms including restoration of homologous recombination status. We have developed a selective inhibitor of ribosome biogenesis, CX-5461, which targets RNA Polymerase I (Pol I) transcription of ribosomal RNA (rRNA) synthesis. We have recently demonstrated that CX-5461 activate DNA damage response at rRNA genes, is synthetic lethal with HR deficiency and enhances synthetic lethal interactions of PARPi with HR deficiency in high-grade serous ovarian cancer (HGSOC) patient-derived xenograft (PDX) model. Importantly, CX-5461 has single-agent efficacy in PARPi-resistant HGSOC-PDX. To investigate the mechanisms of PARPi resistance and the mechanisms of efficacy of CX-5461 in PARPi-resistant ovarian cancer cells, we generated olaparib resistant-ID8 mouse ovarian cancer cell lines. Using this cell model, we performed quantitative global proteomic and phosphoproteomic analyses and identified altered metabolism to be associated with olaparib resistance in ID8 cells. We also identified key biological processes to be implicated in cellular response to CX-5461 including ribosome biogenesis, mRNA translation and mRNA processing and mitotic cell cycle progression. Further biochemical investigations will elucidate the molecular mechanisms of PARPi resistance and define the signalling pathways that mediate the actions of CX-5461, which will facilitate the rational design of more effective regimens for ovarian cancer patients.

108

Cell wall and whole cell proteomes define flocculation and fermentation behaviour of yeast

Edward EDK Kerr[1, 2, 3], Duin McDiarmid [1, 3], James Fraser[1, 3], Ben BLS Schulz[1, 3, 4]

1. University of Queensland, Brisbane, QLD, Australia

2. Newstead Brewing Company, Brisbane, QLD, Australia

3. Australian Infectious Diseases Research Centre, University of Queensland, Brisbane, Queensland, Australia

4. Centre for Biopharmaceutical Innovation, Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, Queensland, Australia

Flocculation is one of the most important characteristics of brewing yeast as it allows for the easy and cheap removal of cells after fermentation. The genes responsible for both the Flo1 and NewFlo flocculation phenotypes are well characterized. However, the relationship between Flo protein abundance and flocculation efficiency is poorly understood. In this present study, we used mass spectrometry proteomics to compare the cell wall and whole cell proteomes of commercial yeast strains with diverse flocculation behaviours. We found that the relative abundance of Flo1/5 or Flo10 in the cell wall was correlated with the ability of these yeast strains to flocculate. Analysis of whole cell proteomes identified differences in the proteomes of yeast strains and identified the potential for high metabolic diversity. Characterization of the cell wall and whole cell proteomes during fermentation showed high levels of Flo10 in cells that settled early during fermentation. Our data reveal the diversity of the cell wall and global proteomes of brewing yeast, highlighting the potential biochemical diversity present in yeast that can be utilized in the production of fermented beverages.

109

New Stemsation: Proteomic Characterisation of Human Adipose Stem Cells

Naomi Koh Belic[1], Jerran Santos[1], Bruce Milthorpe[1], Matthew Padula[1]

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

Human adipose stem cells are widely used as treatments for a plethora of disorders, despite minimal scientific evidence of their safety let alone efficacy. These unproven stem cell treatments are being offered by predatory clinics around the globe, more than 60 of which are practicing in Australia. Our understanding of human adipose stem cells is limited, and this study is the first to characterise the proteome of freshly isolated human adipose stem cells.

Adipose stem cells were isolated from abdominal lipoaspirates of 8 healthy patients. Their proteome was characterised through a comprehensive analysis of the whole cell lysate, membrane bound fraction, extracellular vesicles and secreted cytokines.

An average of 4379 proteins were detected in the whole cell lysate of each patient, with 1854 of these proteins conserved across all 8 patients, demonstrating patient diversity. In addition, an average of 1583 proteins were detected in the membrane bound fraction of each patient with 577 proteins conserved across all patients, providing an extensive catalogue of cell surface markers that are useful for antibody-based assay development to address the paucity of reliable markers.

Stem cells secrete extracellular vesicles in substantial quantities, and they are known to play a significant role in cancer, injury healing and immune suppression. An average of 2582 proteins were detected in the extracellular vesicles of each patient, 1312 of which were conserved across all patients, 15.37% of these detected proteins were unique to the extracellular vesicles. An additional 27 cytokines were quantified and 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 these cells. This is an invaluable tool as it will assist in developing much needed understanding of cells that are already being used in the clinic.

110

Redox modifications of cysteine in the liver of type 2 diabetes mellitus

Desmond K Li[1, 2], Alexander W Rookyard[3, 2], Lauren E Smith[1, 2], Yen Chin Koay[2, 4], Holly McEwan[2, 5], Anthony Don[2, 5], John F O'Sullivan[2, 4], Stuart J Cordwell[3, 1, 2, 6], Melanie Y White[3, 1, 2]

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

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

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

4. Heart Research Institute, Sydney, NSW, Australia

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

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

Reactive oxygen species (ROS) and associated redox modifications have normal physiological roles in signalling but are also implicated in a number of pathologies. Oxidative stress associated with type 2 diabetes mellitus (T2DM) has been attributed to increased generation/reduced clearance of ROS in insulin sensitive tissue including the liver. During oxidative stress, proteins undergo redox modifications through the thiols of cysteine, which may alter structure, functionality and signalling pathways. To determine proteins of the liver which undergo redox modifications with T2DM, we investigated a number of associated metabolic disorders using redox proteomics. Rats were subjected to a high fat diet (HFD) to induce peripheral insulin resistance, as well as streptozotocin (STZ) to create *β* cell dysfunction, thereby generating the T2DM pathology. Rats treated with only HFD or STZ were utilised to analyse the pre-diabetic state. To enrich low abundance redox modified peptides, thiol disulfide exchange was utilised for reversibly modified cysteines with quantitation by iTRAQ and mass spectrometry (MS). Strong cation exchange was used to select for irreversibly modified cysteines, combined with MS based label-free parallel reaction monitoring (PRM) quantitation. We identified 7,635 peptides with reversibly oxidised cysteines, with over 3,000 changing abundance in response

to disease progression and quantified 123 PRM amenable peptides with irreversibly oxidised cysteine. Oxidative modifications to proteins which maintain redox balance suggests dysregulation, contributing to oxidative stress in T2DM. Observed changes in enzymes involved in glucose metabolism suggest a role of redox modifications in glucose utilisation/production. Modifications to enzymes in fatty acid metabolism correlate with the hyperlipidemia associated with T2DM. Many proteins showed increased levels of irreversible oxidative modifications contributing to protein damage. Changes in the redox status of these proteins could be indicative of these protein species prone to oxidative damage, contributing to pathogenesis or a level of cellular regulation in response to the metabolic disturbance.

111

Association of post-translational modifications in the evolution of diabetic nephropathy in a rodent model of type 2 diabetes mellitus

Shivanjali J. Lingam[1], Desmond K. Li[1], Stuart J. Cordwell[2], Melanie Y. White[2]

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

2. Discipline of Pathology, School of Medical Sciences, School of Life and Environmental Sciences, Charles Perkins Centre, University of Sydney, Sydney, New South Wales, Australia

Type 2 diabetes mellitus (T2DM) mediated renal dysfunction (diabetic nephropathy (DN)) is a progressive chronic complication which further increases the risk of T2DM mortality. 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 DN leading to increased oxidative stress. The increased capacity for this T2DM-induced oxidative stress modulates protein function by accelerating post-translational modifications (PTMs) of protein residues. Here, we examined the cysteine redox and phosphorylated-modified renal proteome in T2DM. Rats were fed a standard CHOW (C) (12% fat) or high fat (HFD) (42% fat) diet for 8 weeks, with T2DM induced in 50% of the animals after 4 weeks utilising a low dose of streptozotocin (STZ; 35mg/kg); a pancreatic β-cell toxin. The remaining 50% were injected with a buffer vehicle (Cit). At the cessation of the feeding protocol, 9 animals per treatment group (CHOW/Cit; CHOW/STZ; HFD/Cit; HFD/STZ) were euthanised and kidneys excised. Tissue lysates were used to perform histological analysis to confirm the presence of DN pathological features of glomerulosclerosis and hyalinisation more pronounced in the HFD/STZ group. Tissue lysates were also 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 phosphorylation and redox PTMs, we performed isobaric tagging (TMT) and enriched for modified peptides using thiol-affinity chromatography, with subsequent analysis by LC-MS/MS. Significantly modified proteins, changing in abundance between one or more treatment group in comparison to CHOW/Cit, were mapped to enriched regulatory pathways. Our data suggests that the rat renal redoxome and phosphoproteome are sensitive to diet and T2DM, indicating a possible physiological remodelling role of oxidation and phosphoproteome in DN.

112

PromarkerD: A novel test for predicting rapid decline in renal function in type 2 diabetes

Scott Bringans[1], Richard Lipscombe[1], Kirsten Peters[1, 2], Tammy Casey[1], Jason Ito[1], Wendy Davis[2], Timothy Davis[2]

1. Proteomics International, Broadway, Nedlands, WA, Australia

2. Medical School, University of Western Australia, Perth, WA, Australia

Chronic kidney disease (CKD) affects one in three adults with diabetes, accounting for 40,000 deaths and $100 billion (USD) in healthcare spending annually. The current tests for CKD, urinary albumin:creatinine ratio (ACR) and estimated glomerular filtration rate (eGFR), have limited accuracy to predict CKD progression.

PromarkerD is a blood test that measures three plasma protein biomarkers (CD5L, ApoA4, and IBP3) combined with three clinical factors (age, HDL-cholesterol and eGFR) to predict risk of renal decline in patients with type 2 diabetes (T2D) over the ensuing 4 years. PromarkerD was developed using a proteomics workflow in patients with T2D drawn from the longitudinal observational Fremantle Diabetes Study Phase II (FDS2). Two versions of the test were developed, a targeted mass spectrometry based assay and an ELISA assay, both of which were used to measure the PromarkerD plasma biomarkers. Risk predictions were compared between the two platforms using Bland and Altman plot analysis.

The plasma biomarkers add significant incremental benefit to conventional clinical risk factors for predicting rapid decline in renal function in T2D. The biomarkers outperform both eGFR and ACR for predicting future renal decline and provide physicians with a more informed approach to managing diabetic kidney disease and patient care. PromarkerD may also be useful for risk stratification in future clinical trials.

113

Molecular profiling of fermented dairy products using mass spectral fingerprinting technologies

Evelyne Maes[1], Fionnuala Murphy[1], Ines Weissenbacher[1], Erin Lee[1], Hanh Nguyen[2], Raise Ahmad[2], Julie Dalziel[2], Anita Grosvenor[1], Alastair Ross[1], Li Day[2], Stefan Clerens[1]

1. AgResearch Limited, Lincoln, New Zealand

2. AgResearch Limited, Palmerston North, New Zealand

The proteins in milk and dairy products are an important nutrient source for humans. Fermented milk products have long been a common way to consume milk across many cultures, and add flavour, functionality and nutrition. Fermentation of milk releases small peptides that lead to enhanced sensory properties and may also have health promoting bioactivities. The peptide and small

molecule profiles of milk and its fermented products can reveal the different metabolic actions of diverse starter cultures and aid in the selection of optimal microbes for enhanced fermented milks.

In this study, we compared the molecular profiles of four dairy products: bovine milk (negative control), milk hydrolysate (positive control) and two milk products fermented using two different commercial starter cultures. Several mass spectrometry-based approaches were used to evaluate the difference in molecular expression between the four samples. Rapid evaporative ionisation mass spectrometry (REIMS) screening was used to determine the small molecule diversity of the samples. Next, the peptides and amino acids present in the samples were profiled using a fingerprint obtained by matrix-assisted laser desorption ionisation (MALDI) – time-of-flight (TOF). Combined, the small molecule and peptide fingerprints of the products function as a 'molecular signature' that can rapidly reveal 1) the molecular similarity of fermented products, 2) the effects of proteolysis and 3) the small molecule composition of the products. To identify the sequence of specific peptides released, and allow the prediction of bioactivity, a peptidomics approach was then used. Taken together, molecular fingerprinting was able to distinguish the effects of different starter cultures on proteolysis and subsequent peptide release in different milk products. With this suite of techniques, a rapid molecular profile of the influence of different starter cultures in the fermentation process can be obtained which can be applied to tailor new products and cultures.

114

Kinetic mass spectrometry imaging to study cytosolic ribosomes remodeling in plants

Federico Martinez-Seidel[1], Olga Beine-Golovchuk[2], Pipob Suwanchaikasem[3], Joachim Kopka[4], Ute Roessner[3], Berin Boughton[3]

1. The University of Melbourne; Max-Planck-Institute of Molecular Plant Physiology, West Melbourne, VIC, Australia

2. Biochemie Zentrum (BZH), Universität Heidelberg, D-69120 Heidelberg, Germany

3. School of BioSciences, The University of Melbourne, Parkville , Victoria 3010, Australia

4. Molecular Physiology, Max-Planck-Institute of Molecular Plant Physiology, D-14476 Potsdam-Golm, Germany

Translational feedback on the cellular proteome has been seen historically as a combination of post-transcriptional modifications and ribosome associated factors influencing the translational status of a transcript. Thus, leaving the ribosome itself as a static non-tuneable machine (Genuth and Barna, 2018b). Consequently, specialized ribosomal sub-populations that would be capable of selective translation remain a controversial premise (Haag and Dinman, 2019). Functional specialization of plant ribosomes is supported by current research (Genuth and Barna, 2018a), and could partially rely on the usage of non-canonical ribosomal RNAs and proteins (RPs) (Martinez-Seidel, et al. 2019 - manuscript in interactive review). Many RP paralogs in plants appear to have neo or subfunctionalized. Hence, the usage of specific RPs, or paralogs, to assemble new ribosomes could ultimately influence the functional features that newly synthesized ribosomes use to constrain translation. We have determined that during cold acclimation, Arabidopsis is able to reprogram its RP composition independently from growth, confirming that a non-canonical RP composition characterizes stress-induced heterogeneous and possibly specialized ribosomes. Furthermore, the reprogramming is not random but significantly constrained to specific ribosomal regions. The RPs populating the significantly remodelled regions are of major importance to understand the functional aspects of this phenomenon. A combination of spatio-temporal mass spectrometry, i.e., stable isotope label-assisted mass spectrometry imaging will allow us to determine the synthesis of cytosolic ribosomal proteins in root meristems. The meristematic zones in plant roots represent the main regions responsible for plant growth and as such are a hotspot for the appearance of new and potentially specialized ribosomal populations. To access this information we have developed and optimized a mass spectrometry imaging method from sample embedding steps up until data analysis. Results show that we are able to discriminate root meristematic tissue as the largest variance influencing factor based on a K-means clustering partition of extracted ion counts.

1. Genuth, N. R., and Barna, M. (2018a). Heterogeneity and specialized functions of translation machinery: from genes to organisms. Nat. Rev. Genet. 2018, 1. doi:10.1038/s41576-018-0008-z.
2. Genuth, N. R., and Barna, M. (2018b). The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life. Mol. Cell 71, 364–374. doi:10.1016/j.molcel.2018.07.018.
3. Haag, E. S., and Dinman, J. D. (2019). Still Searching for Specialized Ribosomes. Dev. Cell. doi:10.1016/j.devcel.2019.03.005.

115

Cotton reproduction: SWATH-MS reveals stage-specific proteins involved in cotton pollen development

Farhad Masoomi-Aladizgeh[1], Matthew J McKay[2], Mehdi Mirzaei[2], Brian J Atwell[1]

1. Department of Biological Sciences, Macquarie University, North Ryde, NSW, Australia

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

Sexual reproductive success is critical for the production of seeds and is therefore essential for agricultural productivity. However, pollen development is highly sensitive to environmental events such as heat and frost, which can decimate crops.[1] The overarching aim of this study was to investigate cotton (*Gossypium hirsutum*) pollen throughout the reproductive cycle by analysing proteomes at four distinct stages of development. Pollen grains were isolated from the gametophytic phases of development including tetrads, uninucleate and binucleate microspores and mature pollen. Using a phenol extraction method, in-solution digestion and LC-MS analysis, a library consisting of 5257 *G. hirsutum* proteins was constructed from all four developmental stages. SWATH-MS was used to quantify 4501 proteins and indicated that some proteins are stage specific and are associated with the development processes including those involved in the biosynthesis of secondary metabolites, carbon metabolism, biosynthesis of amino acids, protein processing in the endoplasmic reticulum, and glycolysis. We conclude that developing pollen cells have qualitatively distinct protein profiles at each stage. Proteins identified at the tetrad stage are

particularly unique, and indicate that the earliest stage of cotton pollen development requires the synthesis of cellular components that enable subsequent mitoses and structural changes to the haploid germ cells.

1.Hatfield, J. L., Prueger J. H. (2015), Temperature extremes: Effect on plant growth and development, *Weather and Climate Extremes*, 10 (A), 4-10.

116

Prevalence and mechanisms of diastaticus-like phenotypes and genotypes in *Saccharomyces cerevisiae* brewing strains

Pippa E Mckinstry[1], Ben Schulz[1], Edward Kerr[1]

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

S. cerevisiae var. diastaticus contamination is a growing concern for craft breweries worldwide. This *S. cerevisiae* variant can efficiently ferment residual starch due to production of extracellular glucoamylases. The presence of *S. cerevisiae var. diastaticus* in the brewing process and after packaging can cause a variety of problems due to the fermentation of residual sugars, including altered flavour, increased ethanol, and increased CO_2. As the source of this contaminating yeast is not known, the only control measures focus on cleaning and sanitation. However, little is known about how contamination of beer with *S. cerevisiae var. diastaticus* occurs, both in terms of the evolutionary history of *S. cerevisiae var. diastaticus*, and in the physical source of contamination in breweries. We detected the presence of *S. cerevisiae var. diastaticus*-like phenotypes in a selection of commercially available yeasts. We then used DIA/SWATH-MS to identify and measure secreted proteins from these yeasts that are associated with the digestion of residual carbohydrates in finished beer. Together, this provides additional insights to the environmental impact on gene expression and growth patterns of *S. cerevisiae var. diastaticus*.

131

Proteomic Characterisation of poor sperm chromatin compaction suggests nuclear retention

Jacob Netherton[1], Mark Baker[1]

1. The University of Newcastle, Callaghan, NSW, Australia

Spermatogenesis is an extremely specialised process that generates a cell capable of the protection and delivery of the paternal genome to the oocyte. During the development of a spermatozoon, the basic chromatin structure of DNA bound to histones is drastically altered, and nuclear volume is greatly reduced. Importantly poor chromatin compaction has been associated with interruption of the histone/protamine 1/protamine 2 ratios in infertile men, however it has been shown these ratios are subject to change over time and to environmental stresses.

To better understand the process of sperm nuclear condensation, we isolated the sperm nuclei from cells with markers of good and poor compaction from an ejaculate. Comparative proteomics was performed on the nuclear proteins, using the quantitative SWATH platform on the Sciex 6600 Triple ToF. We confidently identified 342 proteins, and of these proteins 20 were found to be more abundant in the sperm possessing poor chromatin compaction, many of which are associated with nucleoplasm. Immunoblots using an antibody against TOP2A and PDIA3 confirmed the proteomic analysis. Unexpectedly, no changes were observed in any of the identified histone peptides (H4, H3.3, H1T, H2A/B), nor for protamine 2. Our data suggests an alternate explanation for poor chromatin compaction. Rather than changes in histone or protamine content, it appears that retained or excess nucleoplasm is more prevalent in poorly compacted nuclei.

132

Bacterial death phase initiation by targeted membrane lipid modification

Yu-Yen Chen[1], Shuai Nie[2], Jacqueline Heath[1], Caroline Moore[1], Thusitha W. Rupasinghe[3], Nada Slakeski[1], Eric C. Reynolds[1]

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

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

3. Metabolomics Australia, Bio21 Institute, The University of Melbourne, Parkville, Victoria, Australia

Here we report mechanistic insights into bacterial death phase initiation that occurs as part of the bacterial life cycle. *Porphyromonas gingivalis* is a keystone pathogen in chronic periodontitis. A *P. gingivalis* mutant (Δ*PG1879*), lacking the predicted inner membrane (IM)-associated patatin/PLA2 superfamily member PG1879, did not exhibit a death phase, but instead a prolonged stationary phase. The surviving cells exhibited intact cell membranes and 8-day-old culture regrowth. Contrastingly, the wild type entered death phase, showed a compromised IM and did not regrow under identical conditions. Lipidomic analysis revealed 29 lipid classes, with 16 being novel, and common membrane lipid remodelling that enabled survival during stationary phase. The remodelling involved upregulation of dihydroceramide-containing sphingolipids, in addition to serineglycine dipeptide-containing acyl-oxyacyl lipids (SGL). A PG1879-dependent mass production of Lyso SGL 17:0_OH initiated the death phase by O-deacylation of its acylated counterpart SGL 17:0_OH [15:0].

133

Defining the interactome of the immunity protein GccH - what can it tell us about glycocin F induced bacteriostasis?

Gillian E Norris[2, 1], Sean W Bisset[2, 1], Marc A Bailie[2], Trevor S Loo[2], Mark L Patchett[2]

1. Maurice Wilkins Centre for Biodiscovery, Auckland, New Zealand

2. School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Glycocin F (GccF) is a diglycosylated bacteriocin produced by *Lactobacillus plantarum* KW30, which has two N-acetylglucosamine (GlcNAc) moieties, one of which (linked to a loop serine) is essential for activity, while the other, a more unusually cysteine linked GlcNAc substantially enhances activity (50 fold). Nanomolar concentrations of GccF induce rapid bacteriostasis (within minutes) of susceptible pathogenic bacteria such as *Enterococcus faecalis* and *E. faecium*. Previous research has implicated the transmembrane domain of the GlcNAc-specific phosphotransferase system (PTS) transporter, PTS18CBA, in bacteriostasis, but exactly how it is involved is not yet known. Recent work in our lab identified the immunity protein used to protect producer cells from the effects of GccF. GccH, a 118 amino acid protein that has no homology to any other protein in the data base apart from one whose gene sequence lies immediately adjacent to that encoding the PTS (38 % identity over 42 amino acids). To help us understand the mechanism of action of GccF, we investigated the potential binding partners of this immunity protein. To do this, we expressed HA-tagged GccH in *L. plantarum* cells and carried out anti-HA pulldowns on GccF-treated and untreated cells. Proteins were extracted from the cells in the presence of detergent to ensure that membrane proteins were solubilised On-bead digest of the immunoprecipitated proteins was carried out, and the peptides separated using nano LCMSMS. Peptides were identified with Proteome Discoverer2.2 and proteins interacting with GccH included the GlcNAc-specific PTS transporter and GlcNAc-6-phosphate deaminase, the first enzyme that interacts with imported GlcNAc. The significance of these and related results will be discussed.

134

Using open source data for peptide spectral library generation with data independent analysis label free quantification (DIA-LFQ) for plasma biomarker discovery

Matthew O'Rourke[1], Mark Molloy[1]

1. University of Sydney, Camperdown, NSW, Australia

Introduction: High throughput plasma proteomics using data independent acquisition (DIA)is an important approach for biomarker discovery. However, plasma is a formidable matrix to analyse as it is dominated by several high abundance proteins (e.g. albumin, immunoglobulins) that make detection and quantitation of less abundant proteins a challenge. DIA is best used with reference libraries of peptide MS/MS spectra and there are several approach to generate them. In this report we explored the use of reference plasma libraries from public data which had been produced using Orbitrap instruments.

Method: We have adopted a new approach that uses previously acquired MS data from public repositories for peptide MS/MS library generation, effectively negating the time and cost outlay needed for in-house generation. A dataset was downloaded from proteome exchange (Mann Et al PDX00284) and was combined with a single set of high pH reverse phase LCMS data generated on QE-HF instrument, created from an undepleted pool of individual patient samples containing iRT peptides. The final library contained 1344 protein groups, 9966 peptide entries and 2277 proteins. This “enhanced” library was then used as the reference for a pilot study of undepleted plasma from 15 diabetes patients taking fenofibrate. Data was generated on a QE-HF using capillary flow over 60min, with 35 windows over the 350-1200 m/z range.

Results: Fenofibrate is known to cause reduction in inflammatory process associated with diabetes mellitus and increases in lipid metabolism associated with weight reduction. Using a longitudinal analysis of a total of ~250 quantitated proteins, we found that there was a significant up regulation of proteins associated with lipid metabolism and downregulation in proteins associated with inflammation post treatment with fenofibrate. These results confirmed that open source data can be used for library generation for conducting plasma DIA LFQ experiments.

135

Unravelling the biological mechanisms of Acute Lymphoblastic Leukemia (ALL) drug-resistance with integrated multi-glycoconjugate analysis

Tiago Oliveira[1], Andreia Filipa Ferreira de Almeida[1], Kathirvel Alagesan[1], Arun Everest-Dass[1], Francis Jacob[2], Maria Kavallaris[3], Nicolle Packer[1, 4, 5], Nora Heisterkamp[6], Mark von Itzstein[1], Daniel Kolarich[1, 5]

1. Institute for Glycomics - Griffith University, Southport, QLD, Australia

2. Glyco-Oncology, Ovarian Cancer Research, Department of Biomedicine, University Hospital Basel and University of Basel, Basel, Switzerland

3. Children's Cancer Institute, Lowy Cancer Research Centre, University of New South Wales, Sydney, NSW, Australia

4. Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, NSW, Australia

5. ARC Centre for Nanoscale BioPhotonics, Macquarie University and Griffith University, North Ryde and Gold Coast, NSW and QLD, Australia

6. Department of Systems Biology, Beckman Research Institute City of Hope, Monrovia, CA, USA

Acute Lymphoblastic Leukemias (ALL) are the most common type of children cancer diagnosed worldwide. Although ALL’s long-term survival rates have increased over the past 40 years, approximately 15% of infant and more than 50% of adult patients eventually relapse after a short remission period[1]. The biological mechanisms resulting in ALL's drug resistance, however, are still not fully understood. Changes in glycosylation have previously been linked to ALL drug resistance[2], and the pharmacological

inhibition of glycan-binding proteins such as Galectin-1 sensitized ALL cells to the action of chemotherapeutical drugs[3]. In summary, glycosylation appears to play a relevant role in ALL pathogenesis and drug resistance.

In our quest to capture a comprehensive picture of the cell surface glycosylation, we devised a strategy to determine the glycolipidome, the *N*- and *O*-glycome as well as the proteome and glycoproteome, ideally from a single batch of starting material. After triphasic extraction using a Methanol:Chloroform:Water mixture, the glycan component of glycolipids and glycoproteins are enzymatically and chemically released from the respective fractions and analysed using porous graphitized carbon (PGC) nanoLC-Electrospray tandem mass spectrometry (ESI-MS/MS). The glycoproteome is determined after tryptic digestion and Hydrophilic Interaction Chromatography (HILIC) fractionation on microcrystalline cellulose prior analysis by reverse phased nanoLC-ESI-MS/MS on an Orbitrap Fusion.

The current workflow was used to analyse CCRF-CEM T-ALL cell lines with different levels of 13-desoxyepothilone B (dEpoB) resistance, as well as primary ALL patient samples. Preliminary analysis allowed the identification of more 130 different proteins exclusively present on the dEpoB resistant CCRF-CEM cell lines, a tenth of which are glycoproteins. Moreover, 300 different proteins were present in ALL patient samples but not present in any of the normal bone marrow (BM) control cells. Also, ALL patients exhibited a distinct *N*- and *O*-glycan fingerprint when comparing to normal BM samples. Ultimately, the aim is to provide novel opportunities to tackle therapy-resistance in ALL by mining unique molecular markers and identify drug resistance-associated pathways.

1. Terwilliger, T. & Abdul-Hay, M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. Blood Cancer J 7, e577 (2017).
2. Nakano, M., Saldanha, R., Gobel, A., Kavallaris, M. & Packer, N. H. Identification of Glycan Structure Alterations on Cell Membrane Proteins in Desoxyepothilone B Resistant Leukemia Cells. Molecular & Cellular Proteomics 10 (2011).
3. Paz, H. et al. Treatment of B-cell precursor acute lymphoblastic leukemia with the Galectin-1 inhibitor PTX008. J Exp Clin Cancer Res 37, 67 (2018).

136

TANDEM ION MOBILITY COUPLED WITH MASS SPECTROMETRY FOR GAS PHASE UNFOLDING STUDIES

Dale Cooper-Shepherd[1], Martin Palmer[1], Heather Patsiouras[2], Lee Gethings[1], James I Langridge[1]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Sydney, NSW, Australia

Over more than a decade ion mobility spectrometry (IMS) has been employed to great effect in the field of native mass spectrometry-based structural biology. In this field IMS has been used for separating protein oligomeric states, co-populated conformations and for measuring experimental collision cross-sections for use in determining protein complex architecture. In addition, from an early stage the potential of pre-IMS activation in probing gas phase protein unfolding pathways and stabilities was pursued. This experiment, called collision-induced unfolding (CIU) has spawned a whole field of research into probing protein domain architecture, protein-ligand stabilization and therapeutic antibody comparisons. Here we describe an IMS-based instrument platform with IMSn functionality which allows novel protein CIU studies to be performed in which protein sub-populations can be mobility-selected for further rounds of IMS. Furthermore, by adding an activation step in between IMS experiments, sequential rounds of unfolding can be performed on the same ion populations allowing greater insight into protein unfolding pathways.

Human TTR (Sigma Aldrich) was prepared at a concentration of 4 micromolar in 200 mM ammonium acetate. Native ion mobility experiments were performed on a cyclic ion mobility-enabled Q-Tof (ESI-Q-cIM-Tof) system fitted with an extended time-of-flight mass analyser, a segmented quadrupole transfer ion guide and dual gain ADC. In addition to unfolding studies the ability to mobility-select after quadrupole isolation and activation allows high mobility-resolution interrogation of ligands released from protein-ligand complexes.

137

Technical Reproducibility and Analytical Precision for a Multi-omic Study utilising Data Independent Strategies with CCS libraries

Christopher J Hughes[1], Sarah Lennon[1], Lee A Gethings[1], Robert S Plumb[2]

1. Waters Corporation, Wilmslow, CHESHIRE, United Kingdom

2. Waters Corporation, Milford, MA, USA

OMIC based studies often consist of large scale cohorts in order to allow for more confidence in statistically relevant findings. The LC-MS system and associated methodologies used to perform these experiments need to demonstrate robustness, reproducibility and stability over the entire analysis. These characteristics then allow for the identification of statistically relevant markers and reliable quantification. In this study, we demonstrate these features for a multi-omic study using collision cross section (CCS) libraries with a data independent approach over multiple platforms /laboratories.

Appropriate sample preparation protocols relating to human plasma were utilised to provided proteomic, lipidomic and metabolomic sample sets. LC-MS data were collected using standard flow chromatographic conditions coupled to a Synapt XS mass spectrometer, operating in DIA ($HDMS^E$). For ion mobility (IMS) based acquisitions, collision cross section (CCS) values were generated. The IMS was calibrated using a mixture of compounds, covering a range of m/z values and selected for use in either positive or negative ESI. Data were processed using either Progenesis QI for Proteomics or QI for proteomic and lipidomic/metabolomic data respectively. Identifications resulting from Uniprot (proteomics), LipidMaps (lipids) and HMDB (metabolites) were appended to the processed data. Multivariate statistical analysis was conducted using SIMCA P. In order to assess the reproducibility aspect, data was collected for the same sample sets using replica chromatographic conditions over multiple platforms in different laboratories.

Assessment of the data across all OMIC experiments from the multiple platforms, resulted in low %CV's for normalised abundance and hence provided high quantitative accuracy. High mass accuracy was also maintained across the various datasets with a significant amount of the collected data being less than 2 ppm. CCS measurements were also shown to be highly reproducibility for both day-to-day (based on the same platform) and laboratory-to-laboratory.

138

Turnover optimized short nanoLC gradients on a tims equipped QTOF for high throughput and deep proteome measurements

Thomas Kosinski[1], Scarlet Koch[1], Thorsten Ledertheil[1], Adam Rainczuk[2], Christian Meier-Credo[1], Christoph Gebhardt[1], Heiner Koch[1]

1. Bruker Daltonik GmbH, Bremen, Germany

2. Bruker Pty Ltd, Preston VIC 3072, VIC, Australia

High sample throughput in proteomics, like that used in genomics, is highly desirable. Moreover, the highest analytical depth in proteomics is only achieved on fractionated samples, requiring subsequent analysis with short gradients to achieve reasonable overall measurement times per sample. The timsTOF Pro with trapped ion mobility spectrometry (TIMS) offers additional separation power and increased peak capacity. The powerful Parallel Accumulation Serial Fragmentation (PASEF) method (Meier et al., JPR 2015) for very high sequencing speed is perfectly suited for proteome analysis on short gradients. We have optimized MS conditions, column lengths and LC overhead times to obtain runs of 28.8 min injection to injection (50 samples/day) on the nanoElute (Bruker Daltonik) and demonstrate applicability for high throughput application.

139

Identification of Epidermal Growth Factor Receptor Variant III (EGFR VIII) Expression in Meningioma Indian Patients

Rashmi Rana[1], Kirti Chauhan[1], Rajesh Acharya[1], Sunila Jain[1]

1. Sir Ganga Ram Hospital, Delhi, DELHI, India

Background: Meningiomas are the most common type of intracranial tumors of central nervous system and account for approximately 36% of all primary tumors of CNS. According to the WHO grading system and histological studies it is classified into different grades: grade (I) benign, grade (II) atypical, grade (III) anaplastic or malignant. There are certain mutations and alterations of tumor suppressor genes and chromosomal aberrations leads to the tumor formation and its progression in meningiomas. Epidermal growth factor receptor (EGFR) plays an important role in the regulation of cell division and tumor growth and it's over expression responsible for the stimulation of angiogenesis, cell survival and metastatic proliferation. EGFR variant III is the most common mutation has been already reported in glioblastoma but not yet reported in meningiomas and other type of tumors.

Methods: The main aim of this study is to determine expression of EGFR variant III in different grades of meningioma using primarily immunohistochemistry and also via other techniques such as FACS (fluorescence activated cell sorting) and western blotting.

Results: Results shows the expression of EGFR variant III is highly expressed in benign meningioma (grade I) as compare to atypical (grade II) meningioma. These results could be considered in large number on cohort as possible diagnostic /prognostic biomarker and somehow influence the treatment.

Conclusion: EGFR expression is greatest in benign meningiomas and may serve a potential target for therapeutic intervention with selective EGFR inhibitors.

141

A statistical approach to rank differentially expressed proteins using confidence bound on effect size.

Anup D Shah[1], Paul F Harrison[2], Ralf B Schittenhelm[1], David R Powell[2]

1. Monash Proteomics and Metabolomics Facility, Monash University, Clayton, VIC, Australia

2. Monash Bioinformatics Platform, Monash University, Clayton, VIC, Australia

High-throughput quantitative "omics" experiments often involve characterisation of a large number of biomolecules. Over the years, prioritising this list for subsequent follow-up analysis is generally carried out by applying arbitrary *p* value and/or log fold change thresholds. Recently, statisticians worldwide have expressed their concerns over the use of p value as a surrogate for effect size (1). Additionally, filtering by fold changes also requires an expert insight to determine the threshold. As an alternative, a method of ranking biomolecules by confidence bounds on the effect size, 'topconfects', has been proposed recently (2). It emphasises on biggest effect size in contrast to consistent but small changes in case of *p* value based filtering. This method has shown promising results with 'transcriptomics' data highlighting the biologically significance. Here, we applied 'topconfects' to a label-free proteomics dataset and found marked differences in top-ranked protein list generated by confidence bounds and *p* values.

1. Ronald L. Wasserstein & Nicole A. Lazar (2016) The ASA Statement on p-Values: Context, Process, and Purpose, The American Statistician, 70:2, 129-133
2. Harrison, P.F., Pattison, A.D., Powell, D.R. Beilharz T.H. (2019) Topconfects: a package for confident effect sizes in differential expression analysis provides a more biologically useful ranked gene list. Genome Biol 20, 67

142

Molecular signature associated with lymph node metastasis identified from minimally invasive FFPE colon cancer

Pascal Steffen[1], Jun Li[1], Mahsa Seyed Ahadi[2], Anthony Gill[2], Alexander Engel[3], Mark Molloy[1]

1. Department of Medicine and Health, The University of Sydney, Kolling Institute, Sydney, New South Wales, Australia

2. Department of Anatomical Pathology, Royal North Shore Hospital, Sydney, New South Wales, Australia

3. Department of Colorectal Surgery, Royal North Shore Hospital, Sydney, New South Wales, Australia

Background:

Proteomic analysis of formalin-fixed paraffin embedded tissues is challenging due to low sample recovery and poor LC-MS workflow compatibility. We developed an improved sample preparation workflow and data-independent acquisition strategy for quantitative analysis of colorectal cancer (CRC) FFPE specimens. We applied this approach to investigate molecular features associated with lymph node metastasis in minimally invasive tumours (i.e. T1-2 depth of invasion).

Methods:

Twenty Stage1 and 3a CRC FFPE sections (5x5 µm per sample) were macro-dissected, lysed, and digested according to the SP3 protocol. A pooled sample was used to acquire a DIA gas-phase fractionated chromatogram library. Wide window DIA data were acquired using an overlapping window scheme on a QExactive HFX mass spectrometer with 140min LC time using a 40cm x 75µm self-packed pulled column. For gene expression analysis, samples were analysed by Nanostring using the Cancer Progression kit profiling 770 genes.

Results:

We identified over 6000 proteins and quantified over 4000 proteins across all samples. Using unsupervised clustering methods, we were able to clearly distinguish the cancer Stages based on their proteomic profile. 237 proteins were significantly regulated (q-value <0.05, FC>2). GO-Enrichment analysis showed that in Stage 3a samples (glycoso)aminoglycan metabolic processes (e.g. Biglycan) as well proteins required for Epithelial to Mesenchymal transition were upregulated. Gene expression data for EMT related genes supported the proteomic findings. Two genes (TPM2 and CNN1) that were found significant in both proteomic and Nanostring data and showed upregulation in Stage 3a were also found to be significant for overall survival for patients when queried against the TCGA data for CRC.

Conclusion:

Protein expression phenotype from FFPE tissues of minimally invasive CRC tumours reveals features of EMT associated with lymph node metastasis.

143

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

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

1. CSIRO, St Lucia, QLD, Australia

2. CSIRO, Black Mountain, ACT, Australia

Background

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

Methodologies

A diverse panel of Australian wheat varieties were grown under short (8h) or long day length (16h) conditions to facilitate variation in flowering time. A total of 632 unique wheat samples were processed and measured by variable window SWATH acquisition along with 112 control measurements. Peptide responses were extracted from raw SWATH data along with peak group false discovery rate estimates. These data were processed and analysed using a suite of scripts within the R statistical computing environment. The resulting matrix of pan-wheat measurements was subject to multivariate and machine learning analysis to assess variance and quantify the ability of protein abundances to predict flowering time.

Findings

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

Concluding

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

144

Coordinated post-transcriptional and post-translational regulation through stage transition of *Giardia duodenalis*

Qiao Su[1, 2], Samantha Emery-Corbin[1], Balu Balan[1, 3], Staffan G Svärd[4], Ivo Mueller[1, 5], Aaron Jex[1, 3]

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

2. Faculty of Medicine, Dentistry and Health Science, University of Melbourne, Melbourne, VIC, Australia

3. Faculty of Veterinary and Agricultural Science, University of Melbourne, MELBOURNE, VIC, Australia

4. Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

5. Malaria Parasites and Hosts Unit, Pasteur Institute, Paris, France

Giardia duodenalis is a gastrointestinal parasite responsible for 200-300 million cases of diarrheal disease worldwide, with direct transmission via fecal-oral route. Transmission depends stage transition from binucleate, flagellated trophozoite to tetranucleate cysts, a process considered one of simplest developmental transition within eukaryotes. Studies show parasites induced *in vitro* are irreversibly committed to the encystation process by 3-6 hours, undergoing dramatic changes to gene transcription and translation toward shifting energy metabolism, protein production and requisite epigenetic remodeling. Independent transcriptomic and proteomic studies have demonstrated that encystation is defined by temporally regulated cascades of required encystation genes and pathways, and we hypothesize these are coordinated through post-transcriptional and post-translational regulation.

In this study, we have generated same pellet RNA and protein for deep-sequencing from the predominant human-infective *Giardia duodenalis* sub-species (isolate GS). Our six-point time-series across encystation includes steady-growth, pre-encystation priming, three early to mid-encystation timepoints and the mature cyst. Quantitation will allow us to dynamically follow the transcript-protein correlation throughout developmental transitions to identify post-transcriptional or post-translational regulatory events. Towards this, we will specifically curate RNA-binding proteins of this human-infective sub-species to monitor their kinetics, which can be followed by interactome capture at significant timepoints. Further, we are quantitating H3 and H4 acetylysine and methyllysine marks through priming and early-mid encystation via immunoblotting, covering the critical 'point-of-no-return' during which the parasite commits to stage transition.

Combining quantitative transcriptomics and proteomics for the first time, this study will provide the resolution to understand the regulation of stage transition in *Giardia*, which will provide novel insights into processes critical to interrupt its transmission towards new interventions and drug development.

145

Mass spectrometry-based proteome screening of two narrow-leafed lupin cultivars

Arineh Tahmasian[1], Mitchell G Nye-Wood[1], Angela Juhasz[1], Thao T Le[1], James A Broadbent[2], Michelle L Colgrave[1, 2]

1. Edith Cowan University, Joondalup, WA, Australia

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

Lupin, a member of Fabaceae family is the largest pulse crop grown in Australia. Lupin seeds have favourable nutritional composition including high protein and fibre content and low levels of carbohydrate and fat. This prospective superfood has many health-promoting effects and can be used as a novel plant-based protein source to complement traditional sources.

Narrow-leafed lupin (NLL; *Lupinus angustifolius*) is the dominant lupin species grown in Australia, domesticated in the 1960s. For this study liquid chromatography-tandem mass spectrometry (LC-MS/MS) was employed for the comprehensive proteome analysis of mature seeds from the two Australian NLL cultivars: Tanjil and Unicrop.

Workflows employing data dependent acquisition (DDA) and quantitative data independent acquisition (DIA, using SWATH-MS) were undertaken on a on a TripleTOF 6600 (SCIEX) mass spectrometer. Using a bottom-up proteomics approach exploring different protein extraction methods, the urea-based buffer yielded the highest number of protein identifications from defatted seeds of cultivars Tanjil and Unicrop (1443 and 1430 proteins respectively). The protein functional classes and protein abundances were contrasted in the solvent-extracted proteomes of these two lupin cultivars.

Proteomic studies of genetically diverse lupin varieties under different environmental conditions has the potential to impact lupin crop improvement programs. These programs aim to increase the nutritional value of lupin, while adapting to climatic and environmental challenges.

146

Detecting Proteomic Differences Between Nevi And Early Melanoma Using SWATH-MS

Rachel Teh[1, 2], Ali Azimi[1, 2], Marina Ali[1, 2], Graham Mann[3, 4], Pablo Fernandez-Penas[1, 2]

1. The Department of Dermatology, Westmead Hospital, University of Sydney, Westmead , NSW, 2145, Australia

2. Centre for Translational Skin Research, University of Sydney, Westmead, NSW, Australia

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

4. Melanoma Institute Australia, University of Sydney, Wollstonecraft, NSW, Australia

Background: Australia has one of the highest incidences of melanoma skin cancer in the world. If caught early, malignancy can be successfully prevented. Currently, there are no molecular markers that can differentiate between the common mole or nevi, and melanoma. Therefore, resulting in an increase in unnecessary biopsies in order to remove the offending lesion.

Aim & Objectives: To investigate potential differences in protein content between nevi and early invasive melanoma.

Method: The epidermis layer of the skin was extracted via laser capture microdissection from archived de-identified formalin fixed paraffin embedded (FFPE) human skin tissues. Samples consisted of normal non-lesional (n=5), benign junctional nevi (n=5), dysplastic nevi (n=5), melanoma in situ (n=4), and minimally invasive melanomas of less than 1 mm deep (n=5). All samples were collected from non-chronic sun exposed areas of the torso and proximal limbs. Following preparation and proteolytic digestion, the samples' proteome was analysed in a TripleTOF mass spectrometer using Sequential Acquisition Window Theoretical Fragment Ion Spectra (SWATH MS) workflow. Differential abundance and bioinformatics analysis were then performed on the proteomic data using ingenuity pathway analysis (IPA) and Human Protein Atlas (HPA) database.

Results: We identified 3119 proteins across the 24 samples studied. Our differential abundance analysis subsequently identified a number of proteins that are significantly changed between the normal skin and the lesions as well as between the lesion groups. Bioinformatics analysis predicted a majority of the altered proteins are involved in cell cycle progression, cellular proliferation, apoptosis and differentiation as well as inflammation and oxidative stress.

Conclusion: This pilot study demonstrates the successful detection of proteins in FFPE samples of benign nevi, dysplastic nevi, melanoma in situ, and early melanoma using SWATH-MS methodology. Exploration of this technique further will allow the identification of novel biomarkers for the detection of melanoma in its early stages.

147

Protein-based human identification

Rebecca Tidy[1], Glendon Parker[2], Zachary Goecker[2], Brett Chapman[3], Nicola Beckett[1], Thao Le[1, 3], Colin Priddis[1]

1. ChemCentre, Perth, WA, Australia

2. Department of Environmental Toxicology, University of California Davis, Davis, CA, USA

3. School of Science, Edith Cowan University, Joondalup, WA 6027, Australia, Joondalup, WA , Australia

Protein-based human identification (PBHI) is a novel and innovative tool that has potential to be used by forensic investigators within Australia to associate evidence with a suspect. DNA-based human identification is more familiar, but both methods are complementary and take advantage of an individual's biological variation. Proteomic genotyping of the human hair shaft detects genetically variant peptides (GVPs) that result from single amino acid polymorphisms. These GVPs can be used to complement established DNA-based methods or be used when these methods fail to provide a forensically useful answer. A core focus will be the development of this technology to the point where they can be used to infer the profile of matching genetic variation and to use this in a legal context. Firstly, mirroring established methodology, an investigation of the feasibility to perform human hair shaft proteomic genotyping within a forensic toxicology setting was conducted. Extractions of genetically variant peptides (GVPs) from human hair shaft test samples were performed prior to analysis on a QE Plus Orbitrap instrument equipped with ultra-high performance liquid chromatography, a typical forensic toxicology platform. Data analysis was performed using Proteome Discover and evaluated against published results for test samples under investigation. The combined total of GVPs identified from the test extracts was 80. Comparing the data to a pre-established GVP panel consisting on 24 (12 variants), 11/12 GVPs were identified in multiple test samples. Additional GVP variants were also identified. Verification of two test extracts GVPs to published data; 11/11 and 14/15 GVPs were identified correctly. These results are promising and indicate huge potential of this information for human identification and forensic intelligence. The approach detailed here exploits new advances in protein science and forensic toxicology.

148

Mediator Lipidomics: Towards Comprehensive Metabolic Profiling of Eicosanoids and Related Fatty Acids

John Hewetson[1], Atsuhiko 'Ash' Toyama[2], Masaki Yamada[2], Tatsuro Nakamura[3], Takahisa Murata[3]

1. Shimadzu Scientific Instruments, Ermington, NSW, Australia

2. Analytical and Measuring Instruments Division, Shimadzu Corporation, Kyoto

3. Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan

Eicosanoids, such as prostaglandins, omega-3 fatty acids and their metabolites, are a major family of bioactive lipids that mediate autocrine, paracrine and endocrine signaling in diverse pathophysiological systems. Quantitative analysis of eicosanoids is an attractive approach to increase insight into their signaling roles to better characterize a disease state and for potential contribution to biomarker development.

Using the UHPLC-MS/MS platform with established ease of compound identification, robustness and high-throughput, we sought to maximally increase the coverage of lipid mediators to develop a widely-targeted method for comprehensive metabolic profiling of eicosanoids and related fatty acids. The method was developed consisting of 326 MRM transitions acquired within 20 minutes of chromatographic separation for measuring 196 fatty acid metabolites and 18 deuterium-labeled analogs as internal standard. Identification of target signals was systematized by combining the reference ion ratio criteria with the tool-assisted retention time matching criteria at an unprecedented precision of 3 second time window. The method sensitivity evaluated as on-column LLOQ ranged from 0.1 to 1 pg for the majority of targets, which enabled quantitative detection of 67 targets from 5 µL equivalent volume of control human plasma. We used this method for a preliminary investigation of model mice undergoing inflammatory responses. Multivariate analysis of around 100 targets detected from mice serum samples resulted in characteristic profiles with notable implications.

149

Unravelling the molecular contributions to collagen higher order structure

Danielle R Visser[1], Gillian E Norris[1]

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

Collagens I (α1 and α2) and collagen III are the dominant structural molecules in skin, and their higher order structures give it both strength and flexibility. Two amino acids, hydroxylysine and hydroxyproline are unique to all collagens and are intimately involved in fibre formation. Cross-linking, a vital step in collagen fibrillogenesis, contributes greatly to the structural integrity of collagenous tissues with defects in cross-link formation leading to pathogenesis. Because these cross-links form between hydroxylysine residues and/or lysine and histidine residues, we wondered if differences in the pattern of collagen I-hydroxylysine glycosylation would affect the number, location and types of crosslinks formed, and as a result differences in the physiochemical properties of collagen structures. While there are many reports in the literature on the structure and properties of collagen, much of this research was reported last century using less sensitive techniques than those available today (Neuman, 1950). Using different extraction and purification methods to prevent the loss of specific features of the collagen molecule, we showed, using LCMS, that two cross-links recently proposed to be artefacts of extraction (Eyre, 2019), were in fact present in all skins analysed. Mass spectral analysis also revealed that there are interesting differences in the hydroxylysine glycosylation patterns of the collagen α1(I) and collagen α2(I) chains in all skins tested. Furthermore, there were differences in the hydroxylysine glycosylation patterns of the collagens purified from the skins of different animals, and the ratios of the different types of collagen which are extracted from each animal. In addition, amino acid analysis combined with mass spectral analysis showed that on average 70% of proline residues were hydroxylated in collagen I, a figure much higher than previously thought. While these findings need to be confirmed, they challenge some long held beliefs about the collagen molecule and provide a firm foundation for future work.

1. Eyre, D. R., Weis, M., & Rai, J. (2019). Analyses of lysine aldehyde cross-linking in collagen reveal that the mature cross-link histidinohydroxylysinonorleucine is an artifact. Journal of Biological Chemistry, 294(16), 6578-6590.
2. Neuman, R. E., & Logan, M. A. (1950). The determination of hydroxyproline. J Biol Chem, 184(1), 299-306.

150

Phosphomatics: A free web based tool for the analysis of phosphoproteomic peptide lists

Michael G Leeming[1], Ching-Seng Ang[1], Shuai Nie[1], Swati Varshney[1], Syeda Sadia Ameen[1], Heung-Chin Cheng[1], Nicholas A Williamson[1]

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

Here, we present '*Phosphomatics*' – a new web-based tool for interrogating possible upstream kinases for phosphorylated substrates observed in phosphoproteomics experiments. *Phosphomatics* takes an input list of phosphopeptides, which can include fold change information, and automatically searches databases of known substrate-kinase relationships for information on any known upstream kinase/s that are recorded as being responsible for specific phosphorylation events. Thus *Phosphomatics* can rapidly identify possible upstream kinases that may be active in an experimental dataset. *Phosphomatics* additionally maps the phosphoproteins and the possible kinases on to KEGG signalling pathways, connects the user to the primary literature sources, can identify similar substrate peptides in the dataset, and has an inbuilt text-mining feature for rapid literature searches. *Phosphomatics* primarily utilises data from the SIGNOR database of causal relationships and then connects other databases and web tools such as UNIPROT, PubMed, Protein Atlas, and iTextMine.

Phosphomatics is a free and user-friendly web resource for biologists to rapidly interrogate their phosphoproteomics datasets.

Phosphomatics is freely available via the internet at: www.phosphomatics.com

151

Surviving ER stress in a mouse model of ulcerative colitis

Richard Wilson[1], Rohit Gundamaraju[2], Nuri Gueven[3], Raj D Eri[2]

1. Central Science Laboratory, University of Tasmania, Hobart, Tas, Australia

2. School of Health Sciences, University of Tasmania, Launceston, Tasmania, Australia

3. Pharmacy, School of Medicine, Faculty of Health, University of Tasmania, Hobart, TAS, Australia

Background: Accumulating evidence suggests that the goblet cell-derived mucin-2 (Muc2) is a major component of the immune system and that perturbations in Muc2 lead to an ulcerative colitis-like phenotype. The animal model *Winnie* carries a missense mutation in Muc2 that causes Muc2 misfolding, accumulation in goblet cells and ER stress. Excessive ER stress is a hallmark of many diseases, including ulcerative colitis, cancer, diabetes and Parkinson's disease. However, rather than committing to cell death, the typical outcome of unresolved ER stress, *Winnie* goblet cells are characterised by hyper proliferation, suggesting additional regulation of this cellular stress response.

Methods: To elucidate the molecular mechanisms underlying ulcerative colitis in the *Winnie* model we isolated goblet cells from *Winnie* and wild type mice and used label-free quantitative proteomics and bioinformatics to understand the functional consequences of Muc2 misfolding and accumulation.

Results: A large number of changes were identified that highlight a dramatic reprogramming of energy production, including enhanced utilization of butyrate, a key energy source of colonic cells. A major finding was the marked up-regulation of the coiled-coil-helix-coiled-coil-helix domain proteins Chchd2, Chchd3 and Chchd6. In particular, we identified and confirmed the upregulation and nuclear translocation of Chchd2, a protein known to inhibit oxidative stress induced apoptosis.

Conclusions: This study is the first to apply proteome-level analysis to the pre-clinical *Winnie* model of ulcerative colitis. Identification of proteins and pathways affected in isolated *Winnie* goblet cells provides evidence for novel adaptive mechanisms underlying cell survival under conditions of chronic ER stress.

152

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

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

1. Bruker, Melbourne, VIC, Australia

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

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

4. Bruker Daltonic GmBH, Bremen, Germany

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

6. University of Zurich, Zurich, Switzerland

7. University of Copenhagen, Copenhagen, Denmark

Introduction

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

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

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

Conclusion

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

153

Evaluation of a tims-Q-TOF instrument for targeted proteomics

Steve Wilson[1], Antoine Lesure[2], Pierre-Oliver Schmit[3], Joseph Longworth[2], Francois Bernardin[2], Gary Kruppa[3], Gunnar Dittmar[2]

1. Bruker, Melbourne, VIC, Australia

2. Luxembourg Institute of Health, Strassen, Luxembourg

3. Bruker Daltonic GmBH, Bremen, Germany

Background

Targeted proteomics approaches are now commonly used, either to perform targeted biomarker candidate discovery or to validate candidate biomarkers that had been pointed out by untargeted discovery proteomics approaches. In this study, we are performing a first evaluation of the potential of the recently introduced nano-ESI tims Q-TOF architecture for targeted proteomics approaches. More specifically, we want to check if the time and space concentration of ions that results of the Trapped Ion Mobility Spectrometry (tims) separation process can translate in to increased sensitivity and specificity performances for targeted proteomics approaches.

Methods

An equimolar mixture of 259 quantified synthetic peptides labelled with stable isotopes (AQUA) was diluted in a 100ng/µl human cell line digest. The dilution series covered 6 concentration levels ranging from 31.25 amol/µl to 25 fmol/µl. All samples were separated by nano-HPLC with a 60 min gradient and analysed on a high-resolution timsTOF Pro instrument (Bruker Daltonics) operated in data dependent PASEF mode or an exploratory targeted TIMS-PRM acquisition modes. The global sensitivity, selectivity and detectability of the different acquisition modes was evaluated with the latest version of the Skyline Ô software.

Findings

The PASEF acquisition of cell lysate digest spiked with 25fmol, 6,25fmol, 1562.5amol, 500amol, 125amol and 31.25 amol of the AQUA mixture allowed to identify 253/235/205/104/48 and 4 of the original 259 AQUA peptides, respectively. Using an exploratory tims-PRM approach with a 100 ms tims trapping time, and prior to any collision energy optimization, 110/168/205/213 and all of the AQUA peptides could be quantified at a the 31/125/500/1562/6250 amol level, respectively. Increasing the tims trapping time allowed to increase the detected S/N ratio. The results obtained after a more complete optimization will also be presented.

Conclusions
We have demonstrated a real (yet) unexploited potential of the tims-Q-TOF architecture for targeted proteomics approaches.

154

Achieving optimal study designs for large-scale proteomics

Rohan Shah[1], Merridee A Wouters[1], Sean Peters[1], Rebecca Poulos[1], Qing Zhong[1]

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

Mass spectrometry-based proteomics is a valuable method that can be used to interrogate healthy and diseased tissue samples at high depth. Recent technological advances now enable high-throughput proteomics, allowing the technique to operate at a level of resolution previously only possible in genomics. The Australian Cancer Research Foundation (ACRF) International Centre for the Proteome of Human Cancer (ProCan®) houses six SCIEX 6600 TripleTOF instruments in a single facility capable of processing approximately 10,000 tumour samples per year, enabling large-scale cancer tissue proteomics. To efficiently process these samples, cohorts must be analysed in multiple batches, collected on multiple instruments and over extended periods of time. Analysing these kinds of proteomic datasets requires appropriate normalization and batch correction. However, it is still unclear how to best design such experiments given the inevitable practical constraints of a laboratory setting.

One desirable criterion for an experiment is that it should be minimally affected by any inter-batch variation introduced in the laboratory. This is particularly difficult to achieve in clinically relevant studies involving human tissue, where samples are limited and multiple influential variables are involved, such as gender, age and cancer subtype. We have developed a Monte Carlo algorithm for producing approximately blocked study designs based on investigator-nominated variables. The method is implemented as a C++ program with a Python wrapper and will be available as a package upon publication. Output of the program is an Excel spreadsheet, providing interpretable batching and run orders for a laboratory user environment. The effectiveness of the blocking process can be visualized by the user with several quality control worksheets. Our method reduces the risk of conducting experiments that are confounded between technical and biological variation and should be of significant value for the design of large-scale proteomic studies.

156

Improved middle-down characterization of antibodies using multiple ion activation techniques and Proton Transfer Charge Reduction on an Orbitrap Eclipse mass spectrometer

Romain Huguet[1], Steve Binos[2], Kristina Srzentic[1], John E.P Syka[1], Christopher Mullen[1], Joshua Silveira[1], Jennifer Sutton[1], Luca Fornelli[3]

1. Thermo Scientific, San Jose, CA, USA

2. Thermo Fisher Scientific, Parkville, VIC, Australia

3. University of Oklahoma, Normal, OK, United States

Relative to peptide-based mass spectrometry approaches, middle-down (MD) strategies offer the advantage of higher molecular sequence integrity in the characterization of biotherapeutics. However, reaching the protein sequence coverage required to confidently verify the complete mAb primary sequence or map post-translational modifications in MD experiments can be challenging. We apply advanced ion activation techniques, including ultraviolet photodissociation (UVPD) and electron transfer dissociation (ETD) in combination with Proton Transfer Charge Reduction (PTCR) reactions to improve IgG1 mAb characterization via liquid chromatography (LC)-MS. We have demonstrated that rapid (8 min long) gradient middle-down LC-MS/MS analyses of digested NIST mAb collectively yielded sequence coverage in excess of 80%. 90% sequence coverage was observed for Fc/2 and LC and around 80% for Fd mAb subunits. ETD and UVPD spectra of polypeptide precursors over 15 kDa are extremely complex and therefore are challenging to process with current m/z to mass spectral conversion software. The high spectral density of these product ion peaks is such that isotopic peak clusters often overlap and are frequently not sufficiently resolved to be differentiated from noise, nor to accurately assign a charge state and monisotopic peak, and thus not properly converted to neutral masses. To achieve more complete sequence coverage, product ion peak clusters associated with fragmentation in the middle region of the mAb subunits must be observed and properly mass deconvoluted. We extend our LC-MS/MS analyses of mAb subunits by utilizing PTCR subsequent to ETD and UVPD to enhance sequence coverage. In a single LC run, we obtained for Fc/2 and Fd over 50% and around 60% for LC mAb subunits. All the Complementarity Determining Regions (CDRs) were partially to fully sequenced. In case of Fc/2, this strategy unraveled an extensive series of large z-ions, unambiguously confirming the glycosylation site of the most abundant glycan variant.

An Orbitrap Eclipse Tribrid mass spectrometer with real time search enhances multiplexed proteome coverage and quantitation accuracy.

Aaron M Robitaille[1], Romain Huguet[1], Ryan Bomgarden[2], Derek Bailey[1], Graeme McAlister[1], Arne Kreutzmann[3], Daniel Mourad[3], John Rogers[2], Daniel Lopez-Ferrer[1], Andreas Huhmer[1], Vlad Zabrouskov[1], Bruno Madio[4]

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

2. Thermo Fisher Scientific, Rockford, IL, USA

3. Thermo Fisher Scientific, Bremen, Germany

4. Thermo Fisher Scientific ANZ, Scoresby, Victoria, Australia

Quantitative proteomics strategies using Tandem Mass Tags™ (TMT™) enable precise measurement of peptides or proteins from samples multiplexed into a single high-resolution LC/MS experiment. Interference can suppress ratio quantitation and thereby mask true differences in abundance. Here we evaluate if an Orbitrap Eclipse Tribrid mass spectrometer including real time search (RTS), advanced spectral processing algorithms, and modified hardware can enhance TMT quantification accuracy and proteome coverage. Synchronous precursor selection (SPS) based methods provided higher accuracy compared to MS2 methods for TMT quantitation. However, depending on which fragments are selected for MS3 quantitation, accuracy can still be distorted. To improve upon this, we implemented RTS between MS2 and MS3 scans. Using this approach, MS3 scans are only triggered if a peptide is identified from the preceding MS2. This increased the number of peptides identified with RTS by 30%. Secondly, RTS selects fragment ions for MS3 quantitation that are generated from the identified peptide on the fly. Thus, quantitation can be improved to be 95% interference free. The Orbitrap Eclipse Tribrid mass spectrometer has an optimized quadrupole that improves ion transmission, enabling narrower isolation widths to improve TMT quantitation accuracy. Additionally, we evaluated how next generation isobaric mass tags could increase multiplexing capacity on the new instrumentation. Overall, the Orbitrap Eclipse™ Tribrid mass spectrometer includes features such as TurboTMT and Precursor Fit which facilitate intelligent acquisition methods that improve TMT quantitation accuracy, precision, and proteome coverage.