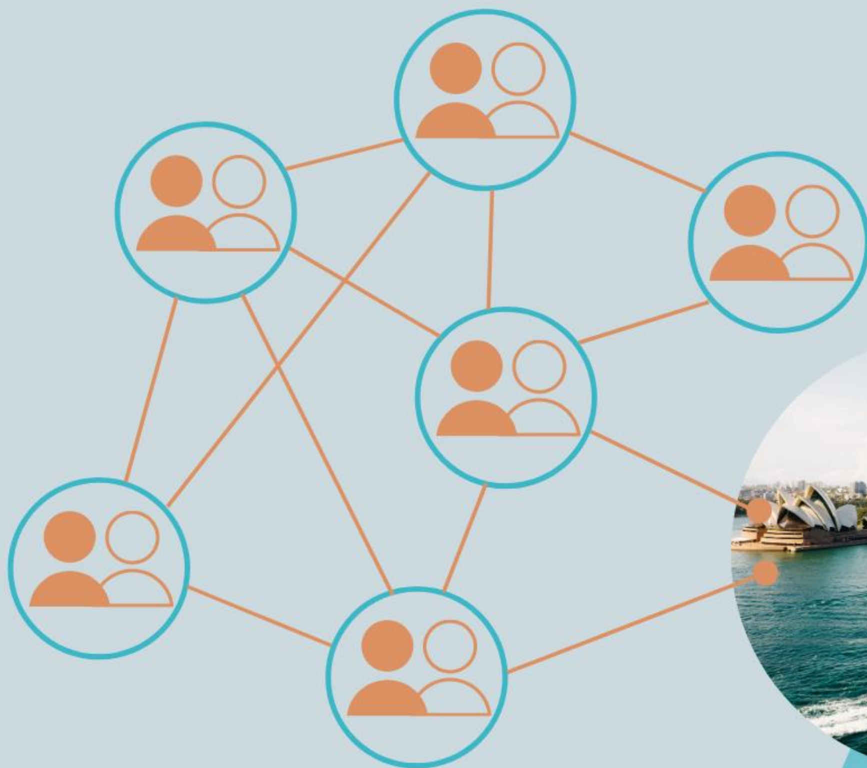

26TH ANNUAL LORNE PROTEOMICS SYMPOSIUM

4 - 5 February 2021

Melbourne, Sydney, and Brisbane Hubs | Virtual

ABSTRACT BOOKLET

1

Multi-omics to explore bacterial virulence and vaccine design

Joel A. Cain[1], William P. Klare[1], Lok Man[1], Ashleigh L. Dale[1], Stuart J. Cordwell[1]

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

The use of proteomics to inform subsequent biological validation studies requires substantial stringency in the analytical approach to ensure that the most important leads are followed. Our laboratory explores virulence determinants including an *N*-linked glycosylation (*pgl*) system and nutrient transporters in the gastrointestinal pathogen, *Campylobacter jejuni*; as well as those involved in *Pseudomonas aeruginosa* infection of the lungs in cystic fibrosis patients. Target identification is based on the response of the proteome to environmental conditions that mimic the host, including gut bile salts and lung mucous, low iron, growth temperatures and various inhibitors. Our proteomics workflow includes parallel label-based liquid chromatography / tandem mass spectrometry (LC-MS/MS) and system-wide validation using data independent analysis (DIA). Here, we discuss the correlation between large-scale datasets and how they facilitate subsequent studies, as well as highlight poorly or non-correlating data. In each case, we show how validated changes in the proteome reflect 'functional reality' that can be determined by integrated 'omics including transcriptomics, metabolomics, lipidomics and lipid A analysis combined with molecular genetics and cellular and *in vivo* virulence assays. Furthermore, we show how a proteomics-informed 'reverse vaccinology' approach has led to the identification of 6 vaccine antigens, two of which provide >99% reduction in bacterial load in the lungs of *P. aeruginosa*-infected mice.

2

Methylation, phosphorylation and their crosstalk on translation-associated proteins

Joshua J Hamey[1], Amy Nguyen[1], Daniela-Lee Smith[1], Marc R Wilkins[1]

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

Protein translation is a central process in the cell. It is currently emerging that protein translation is extensively regulated. This is mediated in large part by post-translational modifications (PTMs), in particular methylation and phosphorylation. However, there remains much to be understood about the prevalence and crosstalk of these two PTMs on translation-associated proteins. Through PTM enrichment, heavy methyl SILAC and mass spectrometric analysis in *Saccharomyces cerevisiae*, we uncovered dozens of novel methylation and phosphorylation sites on translation-associated proteins. We also discovered several instances where these PTMs co-occur on the same protein molecule, indicating co-regulation and potential crosstalk. Arginine methylation, in particular, was found to be pervasive across arginine-glycine repeat regions of proteins associated with translation and ribosome biogenesis. We showed that for one protein, yeast fibrillarin, phospho-methyl crosstalk occurs at serine-arginine-glycine-glycine (SRGG) motifs, where phosphorylation strongly inhibits the deposition of methylation (1). We also discovered several phospho-methyl co-occurrences on the highly conserved and abundant eukaryotic elongation factor 1A (eEF1A) (2). Through *in vivo* mutagenesis and mass spectrometry, we showed that phosphorylation inhibits nearby methylation, but that methylation does not affect nearby phosphorylation. Overall, our results indicate that phospho-methyl crosstalk is more widespread on translation-associated proteins than previously appreciated, and is likely to be central to the regulation of protein translation.

1. Smith, D. L., Erce, M. A., Lai, Y. W., Tomasetig, F., Hart-Smith, G., Hamey, J. J., & Wilkins, M. R. (2020). Crosstalk of phosphorylation and arginine methylation in disordered SRGG repeats of Saccharomyces cerevisiae fibrillarin and its association with nucleolar localization. Journal of molecular biology, 432(2), 448-466.
2. Hamey, J. J., & Wilkins, M. R. (2018). Methylation of elongation factor 1A: where, who, and why?. Trends in biochemical sciences, 43(3), 211-223.

3

Cathelicidin-3 enriched extracellular vesicles as preclinical biomarker for Tasmanian Devil Facial Tumour Disease

Camila Espejo[1], Richard Wilson[1], Greg Woods[1], Eduard Willms[2], Andrew Hill[2], Bruce Lyons[1]

1. University of Tasmania, Hobart, TAS, Australia

2. College of Science, Health and Engineering , La Trobe University, Melbourne, Victoria, Australia

Wild Tasmanian devil populations have declined precipitously due to the spread of a transmissible cancer, Devil Facial Tumor Disease (DFTD). DFTD causes tumours on the face, neck, and oral region that results in death from metastases, and inability to compete and feed. Despite a latent period that can be greater than one year, the diagnosis of DFTD is only possible by visual inspection once tumours have appeared. Confirmation of DFTD requires pathological examination of tumour biopsies. In this context, a pre-clinical test to detect infection before tumours appear is critically needed to control the advance of the epidemic and prevent extinction of devils. Major advances have been made in early detection of human cancer from bodily fluids, based on molecular analysis of extracellular vesicles (EVs). EVs are cell-derived, membrane-enclosed structures of 50-300 nanometers that are present in bodily fluids such as blood serum. We used data-independent acquisition mass spectrometry techniques to analyse the proteome of extracellular vesicles derived from blood serum of a discovery cohort of devils (12 DFTD+ and 10 captive healthy control). This analysis revealed that cathelicidin-3 (CATH3) could classify devils with DFTD in late stages from healthy controls with 100% sensitivity and 100% specificity. To validate these results, we expanded this sample to 65 devils in different stages of DFTD (33 DFTD+; 15 latent, sampled 3-6 months before diagnosis; 17 control) and found that CATH3 predicted DFTD status with 87.8% sensitivity and 94.1% specificity. Further, it classified latent devils as DFTD+ with 93% sensitivity and 94% specificity. Cathelicidins are antimicrobial peptides, and while humans possess only one cathelicidin gene, there is evidence that the human CAP18 gene plays a role in cancer regulation. On this basis, we hypothesize that cathelicidin- 3 (CATH3) is involved in the pathogenesis of

SCIEX
The Power of Precision

DFTD and is a very promising candidate biomarker for early DFTD diagnosis. A preclinical biomarker for DFTD will greatly improve the capabilities of management and conservation actions including allowing healthy devil selection for insurance population, enhancing epidemiological monitoring, and improving outcomes of eventual therapeutic or prophylactics deployments.

4

Integration of immunopeptidomics and functional immunology for the identification of drug-induced T cell epitopes

Johanna EE Tuomisto[1], Julian P Vivian[1], James McCluskey[2], Jamie Rossjohn[1], Nicole A Mifsud[1], Patricia T Illing[1], Anthony W Purcell[1]

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

2. Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Melbourne, Victoria, Australia

Abacavir is an antiretroviral drug used in the treatment of HIV-1 infection. abacavir hypersensitivity syndrome (AHS) is a CD8$^+$ T cell-mediated drug hypersensitivity reaction strongly associated with human leukocyte antigen (HLA)-B*57:01. AHS clinically manifests with systemic symptoms such as fever, gastrointestinal disturbances and rash, which becomes more severe with continued use. The molecular mechanism of AHS involves the occupation of the peptide-binding cleft of HLA-B*57:01 by abacavir, which alters the repertoire of self-peptides presented to T cells. The broad range of novel epitopes induced by the drug results in a polyclonal response by CD8$^+$ T cells. However, abacavir/peptide/HLA complexes that stimulate specific CD8$^+$ T cell receptors (TCRs) have yet to be identified.

Here, we describe the identification epitopes of two abacavir-specific TCRs, which differ in their requirement for the transporter associated with antigen processing (TAP). HLA-B*57:01 peptide ligands were identified from drug-treated HLA-B*57:01$^+$ cell lines (TAP-sufficient C1R.B*57:01 and 1106 KERTr, and TAP-deficient T2.B*57:01). The immunopeptidome workflow involved immunoaffinity purification of HLA molecules, fractionation of peptides and HLA molecules by RP-HPLC and peptide identification by tandem mass spectrometry. Selected peptides were screened with SKW3 reporter cells transduced with TCRs derived from abacavir-stimulated T cells of HLA-B*57:01$^+$ donors, using a CD69 upregulation assay. The screening was performed with 39 peptide ligands of HLA-B*57:01 (identified in abacavir-treated C1R.B*57:01) and 10 signal peptides (identified in the overlap of abacavir-treated C1R.B*57:01, 1106 KERTr and T2.B*57:01 cell lines). The stimulation via the first TCR was TAP-dependent, and the epitope mapped to a widely expressed protein, Ribophorin II. The stimulation of the second TCR was TAP-independent and the epitope mapped to the signal peptide of HLA-DPB1.

These findings highlight different pathways and sources of epitopes for abacavir-specific TCRs, which likely contribute to the complexity of AHS immunopathology. Additionally, this study demonstrates that the discovery of drug-induced peptides by mass spectrometry can be combined with functional assays to identify drug-specific T cell epitopes. Current work strives to characterise the molecular interactions within the abacavir/peptide/HLA/TCR complexes that could be targeted to abrogate these life-threatening immune responses.

5

Identification of novel proteins encoded by the human genome

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

1. Cancer Precision Medicine Group, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

2. Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia

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

The estimate of protein-coding genes is largely unchanged in two decades since the completion of the human genome project. According to current estimates, there are ~20,500 protein-coding genes. This catalog serves as the basis for most biomedical research. Therefore, any missing gene in this catalog is less likely to draw the attention of most researchers. Recently, various ribosome profiling and mass-spectrometry studies have reported several novel proteins encoded by lncRNAs and UTR regions of protein-coding genes. Some of these novel proteins 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 may have potentially missed some of the protein-coding regions and might have annotated them as non-coding. Most of these transcripts lack obvious open reading frames (ORFs>300nt) and display poor evolutionary conservation across vertebrate lineage. We developed a workflow to identify potential novel proteins encoded by lncRNAs and UTR regions of protein-coding genes and generated ORFome database by computationally translating unique lncRNAs from GENCODE, NONCODE and LNCipedia and UTRs of known protein-coding genes. ORFs that do not qualify the cross-species conservation and NMD (Nonsense Mediated Decay) criteria were filtered. Our ORFome database contains 44,844 candidates ORFs. This database was searched against the proteomic data available for 30 human tissues. We identified hundreds of lncRNA and UTR ORFs with protein evidence. Moreover, a subset of them showed tissue-specific expression pattern. Novel proteins encoded by lncRNAs and UTRs of protein-coding genes should enable researchers to elucidate their roles in various diseases.

SCIEX
The Power of Precision

6

Multi-*omics* of B-cell precursor acute lymphoblastic leukemia cells with *MLL* rearrangement (MLL-r) revealed a deep modification of their glycosylation machinery

Tiago Oliveira[1], Eun J Joo[2], Hisham Abdel-Azim[3], Andreia Almeida[1], Kathirvel Alagesan[1], Mingfeng Zhang[2], Francis Jacob[4], Nicolle H Packer[1, 5, 6], Mark von Itzstein[1], Nora Heisterkamp[2], Daniel Kolarich[1, 6]

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

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

3. Division of Hematology/Oncology and Bone Marrow Transplant, Children's Hospital Los Angeles, Los Angeles, CA, USA

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

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

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

Mixed-lineage Acute Leukemia (MLL) is one of the most high-risk forms of pediatric cancer. Although the long-term survival rates of pediatric acute lymphoblastic leukemia have increased over the past 40 years, the current chemotherapeutic treatment schemes often fail in MLL. The biological mechanisms resulting in drug resistance, however, are still not fully understood and there is an urgent need to identify novel diagnostic and therapeutical targets. We have established the first integrated multi-*omics* investigation of primary patient MLL samples and control precursor B bone-marrow (BM) cells from healthy donors, mapping their proteome, transcriptome and glycome.

4-6 million cells from 3 normal BM and 2 MLL samples were used for analysis on our multi-*omics* platform. Porous-Graphitised Carbon (PGC) nanoLC-ESI-MS/MS was used for Glycomics analyses after the enzymatic and chemical release of N- and O-glycans, respectively. The proteome was explored using RP-LC-ESI-MS/MS analyses on an Orbitrap Fusion, performed after offline high-pH fractionation, in addition to RNA-seq analyses.

Overall, 4225 proteins were identified across the patient MLL and control BM cells, of which 216 were overexpressed in MLL ($p<0.01$, log2dif>2). Analyses of RNA-seq and proteomics data revealed significant correlations between gene and protein expression levels, with overexpression in MLL of important glycoprotein signalling receptors and extracellular matrix proteins as well as various transcription factors. Offline fractionation was the key to identify and quantify numerous important glycosyltransferases and other protein expression changes.

The O-glycosylation initiating enzyme GALNT7 was specifically overexpressed in MLL cells, correlating with a significant increase in Core 2 type O-glycans. Core 2 O-glycans account for (in average) 50% of the *O*-glycans found in MLL cells, whilst in Normal BM this percentage decreases to half. Core 1 *O*-glycans, however, experienced a significant decrease in MLL cells (63% in Normal BM to 37% in MLL), accompanied by an increase in *GCNT1* and a decrease of *ST6GALNAC1* transcript levels in MLL cells. Changes on N-glycans were less pronounced, with just a slight increase in complex type glycan levels in MLL cells, largely due to an increase in sialylated N-glycans.

Our data shows that MLL cells undergo and extensive remodelling of the cell surface glycocalyx, in particular on the type of protein O-glycosylation. Next to confirming previous reports describing increased levels of a number of MLL glycoprotein markers such as FLT3, our integrated multi-omics workflow identified a number of hitherto not-reported diagnostic/therapeutic protein candidates that provide novel clues to understand MLL-r pathogenesis.

7

Clinical Proteomics - improving outcomes for children

Vera Ignjatovic[1, 2]

1. Murdoch Childrens Research Institute, Parkville, VIC, Australia

2. Department of Paediatrics, The University of Melbourne, Parkville, VIC, Australia

Whilst the advantages of proteomics, as a functional phenotypic assessment of an individual are well known, the use of proteomics as an assessment of clinical phenotypes has been underutilised, especially when it comes to children. This presentation will outline examples where proteomics and specifically SWATH-MS is being utilised to improve the outcomes for children who are unwell and will focus on the results of one such investigation.

8

Predicting And Monitoring Chemoresponse Using Patient Derived Samples Of Ovarian Cancer

Manuela Klingler-Hoffmann[1]

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

Chemoresistance remains the major barrier to effective treatment of ovarian cancer and predicting chemoresponse is an unmet clinical need. We developed a carboplatin resistant ovarian cancer cell line model and observed enhanced migratory and invasive capabilities in resistant cells. Proteomics analysis was able to separate these populations based on their molecular features. However, the comparable analysis of patients' samples did not show a clear distinction between samples from sensitive and resistant tumours, highlighting the interpatient variability. To be able to monitor direct response of patient derived samples to chemotherapy, we used patient derived ovarian cancer spheroids (3-dimensional multicellular cell clusters), which can be easily isolated from malignant ascites. We performed MALDI-mass

SCIEX
The Power of Precision

spectrometry imaging on spheroids and were not only able to visualise three distinct layers (proliferating, quiescent, and necrotic) within the spheroids, but were also able to monitor drug response. For potential translation into a clinical setting, we developed an assay ready tissue culture plate, which needs minimal sample handling of malignant ascites and can be preloaded with chemotherapy combinations of choice. The chemotherapy response of the patient derived cells/spheroids can be easily monitored using standard fluorochromic stains.

1. Acland M, Mittal P, Lokman NA, Klingler-Hoffmann M, Oehler MK, Hoffmann P. Mass Spectrometry Analyses of Multicellular Tumor Spheroids. Proteomics Clin Appl. 2018 May;12(3):e1700124. doi: 10.1002/prca.201700124. Epub 2018 Feb 9. PMID: 29227035.
2. Mittal P, Price ZK, Lokman NA, Ricciardelli C, Oehler MK, Klingler-Hoffmann M, Hoffmann P. Matrix Assisted Laser Desorption/Ionization Mass Spectrometry Imaging (MALDI MSI) for Monitoring of Drug Response in Primary Cancer Spheroids. Proteomics. 2019 Nov;19(21-22):e1900146. doi: 10.1002/pmic.201900146. Epub 2019 Oct 15. PMID: 31474002.

9

Moving diagnostic quantitative proteomics for rare disease from the bench to the clinic

Daniella H Hock[1], Alison G Compton[2,3], Guy Helman[2,3], Sumudu S Amarasekera[2,3], Ann E Frazier[2,3], Mike T Ryan[4], Zornitza Stark[2,3,5], Cas Simons[2,3], John Christodoulou[2,3,5], David R Thorburn[2,3,5], David A Stroud[1]

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

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

3. Department of Pediatrics, University of Melbourne, Melbourne, Victoria, Australia

4. Monash University, Melbourne, VICTORIA, Australia

5. Victorian Clinical Genetics Services, Royal Children's Hospital, Melbourne, Victoria, Australia

Mitochondrial diseases are a group of highly heterogeneous disorders caused by mutations in over 300 known genes, most of which directly or indirectly impair mitochondrial energy generation. Around 50% of patients with suspected mitochondrial disease remain undiagnosed after massively parallel genome sequencing approaches, often due to the identification of variants of uncertain significance (VUS) in multiple genes and a lack of appropriate functional tests to demonstrate their pathogenicity in a timely manner. The gold-standard and National Association of Testing Authorities (NATA) accredited functional test measures the enzyme activities of mitochondrial respiratory chain complexes I-IV, however this test is not specific as defects in hundreds of genes can lead to an abnormal result. Diagnosis of most patients requires the development of bespoke tests in non-clinical labs, which means patients undergo a diagnostic odyssey of months to years. We believe quantitative proteomics can fill this gap by linking a functional phenotype in the proteome to a VUS, prioritising the variant for sequencing follow-up and providing strong evidence for pathogenicity. Here, I will present a number of recently published[1,2,3] and unpublished case studies demonstrating our proteomics-based approaches to VUS prioritisation and functionalization with and without suggestive genome sequencing leads. As we consider moving our approaches from the research bench to the clinic, I will also discuss the utility of our approach for diagnosis of other rare diseases[4], our investigations into the use of less invasive sample types and faster modes of acquisition with a view toward acute paediatric care, and the next steps toward accreditation of our technique for routine clinical use.

[1] Lake, N. J. *et al.* **Am J Hum Genet** 101, 239-254 (2017).

[2] Frazier, A. E. *et al*. **Med** 1, 1-25 (2020).

[3] Helman, G. et al. **Hum Mutat** *in press* (2020).

[4] Van Bergen, N. J. *et al*. **J Exp Med** 217, e20192040 (2020).

10

A multi-omic approach to study interactions between the gut virome, microbiome and metaproteome during development of islet autoimmunity

Emma E Hamilton-Williams[1], Patrick G gavin[1], Ki Wook Kim[2], Maria E Craig[2], Michelle Hill[3]

1. University of Queensland Diamantina Institute, Brisbane, QLD

2. Faculty of Medicine, University of New South Wales , Sydney, NSW, Australia

3. QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

The gastrointestinal ecosystem is a highly complex environment that can have a profound influence on human health. Inflammation in the gut, linked to increased intestinal permeability and an altered gut microbiome has been linked to the development of multiple human diseases including type 1 diabetes (T1D). Viruses infecting the gastrointestinal tract are also thought to play a role in T1D pathogenesis possibly via similar mechanisms. Yet the inter-relationship between the gut virome, microbiome and host-function have not been studied. Here, we have applied an integrative approach to combine fecal virome, microbiome and metaproteomic datasets from samples collected before and after the onset of islet autoimmunity in a cohort of children at risk of T1D. We show strong age-related effects across all three 'omic datasets. Mastadenovirus infection was linked to profound functional changes in the metaproteome. Multiomic factor analysis modelling revealed proteins derived from the genus Faecalibacterium were linked to case-control status. The methods we have developed form a framework for future large-scale studies investigating disease progression in the context of an altered host-microbe relationship.

SCIEX
The Power of Precision

11

Multi-omics approach to identify cancer neo-antigens for immunotherapy

Harsha Gowda[1, 2, 3]

1. QIMR Berghofer, Brisbane, 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

Cancer is one of the leading causes of death in the world. Recent advances in immunotherapy strategies have revolutionized cancer treatment. This is exemplified by therapeutic efficacy of immune checkpoint inhibitors in treating melanoma and lung cancers. However, most other cancer types do not show similar response to immune checkpoint inhibitors. This has prompted researchers to pursue other therapeutic avenues that can activate immune system to target cancer cells. Some of these strategies rely on T cell recognition of cancer cells based on neo-antigens presented on cell surface. This strategy requires identification of specific neo-antigens that are presented by MHC complex on the surface of cancer cells. Most cancer genome sequencing studies predict cancer neo-antigens based on somatic mutations identified in tumours. However, these approaches can result in several false positives as many coding mutations may not even be expressed in these cancers cells. We carried out whole genome, exome and transcriptome sequencing of gall bladder cancers from South Korea, India and Chile. By combining genomic and transcriptomic data, we identified coding mutations that are expressed in these tumours and predicted cancer neo-antigens from frequently mutated genes. Mutant peptides from ELF3, ERBB2 and TP53 were found to activate T-cells suggesting these peptides are potential cancer vaccine candidates. In order to characterize sequence determinants and other parameters that determine which neo-antigens are presented by MHC complex, we carried out exome, transcriptome, proteome and immunopeptidome analysis on melanoma, lung and breast cancer cell lines. We identified thousands of MHC bound peptides including several mutant peptides expressed in these cancer cell lines. Our data shows that most coding mutations observed at genomic level are not presented by MHC complex. By integrating multi-omics dataset from these cell lines, we investigated features that determine peptides that are most likely to be presented by MHC complex.

12

SWATH glycoproteomics of sparkling wine

Cassandra Pegg[1], Toan K Phung[1], Christopher H Caboche[1], Suchada Niamsuphap[2], Marshall Bern[3], Kate Howell[4], Benjamin L 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, Queensland, Australia

3. Protein Metrics Inc, Cupertino, California, United States of America

4. Faculty of Veterinary and Agricultural, The University of Melbourne, Parkville, Victoria, Australia

Sparkling wine is enjoyed around the world in celebrations and festivities. The sensory properties of sparkling wine depends on both consumer expectations and on a complex interplay between the chemical and biochemical components in the final product. Glycoproteins have been linked to positive and negative qualities in sparkling wine, but glycosylation profiles of sparkling wine have not been previously investigated in detail. Here, we used SWATH-MS to study the glycoproteme based on glycopeptides identified by Byonic. We applied the optimised workflow to three pairs of experimental sparkling wines highlighting several differences between conditions and the grape used. These included differences in protein abundance, glycans at specific sites and the overall abundance of yeast or grape glycoproteins. The workflow developed for SWATH/DIA glycoproteomics is applicable to the study of other glycoproteomes and could also be applied to still or base wine and complements other tools that measure the sensory properties of sparkling wine.

13

Protein extraction protocols for optimal proteome measurement and arginine kinase quantitation from cricket Acheta domesticus

Utpal Bose[1], James Broadbent[1], Angela Juhasz[2], Shaymaviswanathan Karnaneedi[3], Elecia Johnston[3], Sally Stockwell[1], Keren Byrne[1], Vachiranee Limviphuvadh[4], Sebastian Maurer-Stroh[4, 5, 6], Andreas Lopata[3], Michelle Colgrave[1, 2]

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

2. School of Science, Edith Cowan University, Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, Joondalup, Western Australia, Australia

3. Molecular Allergy Research Laboratory, Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, Queensland, Australia

4. Biomolecular Function Discovery Division, Bioinformatics Institute, Agency for Science, Technology and Research, Singapore, Singapore

5. IFCS Programme, Singapore Institute for Food and Biotechnology Innovation, Agency for Science, Technology and Research, Singapore

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

Insects have been consumed by people for millennia and have recently been proposed as a complementary, sustainable source of protein to feed the world's growing population. The development of insects for food requires technology to explore their allergenic potential. Insects and crustaceans both belong to the arthropod family. Crustacean (shellfish) allergies are both relatively common and potentially severe;

SCIEX
The Power of Precision

hence, the cross-reactivity of the immune system with insect proteins (e.g. tropomyosin and arginine kinase) is a potential health concern. In this study, LC-MS/MS was used to explore the proteome of whole, roasted whole and roasted powdered cricket products. Eight protein extraction protocols that encompass different buffer compositions, including one with a pre-extraction defatting step, were compared using the total number of protein and distinct peptide identifications as a qualitative measure. Within this data, 20 putative cross-reactive shellfish allergens were identified, of which three were arginine kinase (AK) proteoforms. Subsequently, a multiple reaction monitoring MS assay was developed for the AK proteoforms and applied to the five extraction methods that yielded the highest number of protein identifications. The urea/thiourea buffer was shown to maximally extract AKs from cricket samples, while the abundance of AKs was noted to vary between different cricket extracts. This targeted assay demonstrated that allergen abundance/ detectability varies according to the different food processing approaches.

14

Phosphoproteome rearrangement of rice leaves in prolonged water stress and varying nitrogen supplementation

Sara Hamzelou[1], Vanessa J Melino[2], Darren C Plett[3], Karthik Shantharam Kamath[4], Arkadiusz Nawrocki[5], Martin R. R Larsen[5], Brian J Atwell[6], Paul A Haynes[1]

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

2. King Abdullah University for Science and Technology, 2955-6990, The Kingdom of Saudi Arabia

3. The Plant Accelerator, Australian Plant Phenomics Facility, The University of Adelaide, PMB 1, Glen Osmond, SA, 5064,, Australia

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

5. Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK 5230 Odense M, Odense, Denmark

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

The rice crop is considered as a strategic commodity in most Asian countries due to its importance in providing food security and introducing employment opportunities. Water and nitrogen accessibility are two of the most important environmental factors in determining the productivity of rice. Of all required nutrients, plants need nitrogen in the largest quantities, as nitrogen is a crucial component of proteins, enzymes, nucleotides, and plant hormones. Even though it is a prominent component of food security, rice production is projected to deteriorate due to climate change and natural resource depletion. Achieving food security along with the efficient use of resources are the main challenges in rice production.

In this study, *Oryza sativa* cv. Nipponbare was exposed to a prolonged water stress (25 days) in the presence of two levels of nitrogen supplementation. Leaves from water-stressed (40% field capacity) and well-watered (100% field capacity) plants grown in 500 and 1500 mg nitrogen supplementation per kilogram of soil were harvested. Proteins were extracted and peptides and phosphopeptides enriched using TiO_2 chromatography were analysed by nanoLC-MS/MS analysis coupled with label-free quantitation.

We hypothesized that phosphorylation, as one of the important post-translational modifications (PTMs), would play significant roles in addition to the proteome rearrangements observed in response to prolonged environmental changes. Extensive changes in phosphorylation were seen in transporter proteins, such as nitrogen-dependent phosphorylation of aquaporin PIP2-6, suggesting the regulatory role of PTM in leaf cell water maintenance. Dephosphorylation of channel proteins was detected in the droughted rice plants, which is mostly in correlation with inhibition of their activity for those well-characterized phosphoproteins such as plasma membrane ATPase and aquaporin PIP2-6.

Our study also revealed that phosphorylation in RNA processing proteins and those involved in carbohydrate metabolism may regulate the signalling cascades required for plant response to drought and nitrogen resources. In the comparison of well-watered and droughted plants grown in 500mg nitrogen, two isomers of sucrose-phosphate synthase, and two of GAPC (Glyceraldehyde-3-phosphate dehydrogenase C), were found to be differentially phosphorylated. Phosphorylation changes have been proposed as the main initial factor for stress-dependent translocation of GAPC to the nucleus. The results of this study show that phosphorylation plays an important signalling role in rice acclimation to drought.

15

Targeted quantitative analysis of conglutin seed storage proteins in lupin seeds

Arineh Tahmasian[1], Angela Juhasz[1], James A Broadbent[2], Mitchell Nye-Wood[1], Thao T Le[3], Michelle L Colgrave[1, 2]

1. Edith Cowan University, Joondalup, WA, Australia

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

3. Auckland university of technology, Auckland, AK, New Zealand

Plant-based sources are under increasing demand to provide new protein ingredients. Lupin seeds with an average protein content of 35-44% are an emerging source of alternative protein. Lupin fortified foods have many health-promoting benefits such as lowering cholesterol and blood pressure, managing glucose levels and helping to fight obesity by lowering energy intake while increasing satiety. The sweet lupin varieties such as narrow-leafed lupin (Lupinus angustifolius, blue lupin, or NLL) are increasingly popular pulse crops, because of their nutritional and nutraceutical properties in addition to their lower content of anti-nutritional alkaloids.

In this study, the major lupin seed proteins, referred to as conglutins, were identified in the mature seeds of NLL commercial cultivars Tanjil and Unicrop and the wild accession P27255. The influence of protein extraction buffer composition on the number of identified conglutin proteins and the abundance of measured marker peptides attributed to these proteins were evaluated. In addition, the reproducibility of extraction protocols was assessed, to develop and optimise an efficient extraction protocol for identification and quantitation of conglutin seed storage proteins.

SCIEX
The Power of Precision

Discovery proteomics analysis was performed on a TripleTOF 6600 (SCIEX) mass spectrometer and the conglutin expression profiles across these three NLL cultivars were studied through quantitative multiple reaction monitoring workflow undertaken on a QTRAP 6500+ (SCIEX) system. The Tris buffer resulted in the lowest number of proteins identified across all the cultivars and yielded lower levels of gamma conglutins. The comparative study of conglutins within the selected cultivars, revealed different expression levels in some beta conglutins which are known as putative allergens.

16

Quantitative proteomics to investigate rabbit haemorrhagic disease virus infection

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

1. CSIRO, Canberra, ACT, Australia

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

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

Rabbit haemorrhagic disease virus (RHDV) is used in Australia to control feral European rabbits that cause excessive damage to the agriculture and Australian ecosystems. Infected rabbits experience extremely rapid onset of liver failure, haemorrhages and death within 72 hours after infection. The virus belongs to the *Caliciviridae* family and contains a single-stranded RNA genome that encodes a capsid protein and several non-structural proteins. The functions of some of the non-structural proteins remain unknown, mostly due to the lack of a robust cell culture system. In order to elucidate functions of the uncharacterised proteins and better understand the infection process, we performed stable isotope labelling of amino acids in cell culture (SILAC) coupled with co-immunoprecipitations of the tagged proteins. This approach aimed at identification of cellular interaction partners of non-structural RHDV proteins in transiently transfected rabbit kidney cells. We found that one of the non-structural proteins, p23, binds chaperone proteins (heat shock proteins 70 and 110). This protein is also predicted to form a membrane-spanning channel which may disrupt ion balances and induce cellular stress. Our results allow to suggest that the heat shock proteins are involved in the stress response induced by p23. Another non-structural protein, the RNA-dependent RNA polymerase, which replicates viral genome, interacts with BRO1 domain-containing proteins. These proteins are involved in cellular membrane remodeling and cytoskeletal dynamics. This finding is in line with previous observations that the polymerase disrupts the Golgi architecture in a wide range of transfected cells, affecting normal cellular trafficking. Interactions between the BRO1 domain-containing proteins and the polymerase may facilitate the formation of the virus replication factories.

17

Proteomics of developing cotton pollen reveals temporal patterns under heat stress

Farhad Masoomi-Aladizgeh[1], **Matthew McKay**[2], **Mehdi Mirzaei**[2], **Dana Pascovici**[2], **Paul Haynes**[3], **Brian Atwell**[1]

1. Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia

2. Australian Proteome Analysis Facility, Macquarie University, North Ryde, NSW 2109, Australia

3. Department of Molecular Sciences, Macquarie University, North Ryde, NSW 2109, Australia

Heatwaves resulting from global warming are a leading threat to cotton crops because they specifically affect reproductive (gametophytic) development. We investigated pollen development using proteomics after exposure to heat in order to identify how gene products respond in tetrads, uninucleate and binucleate microspores, and mature pollen. A library consisting of 5257 *G. hirsutum* proteins was constructed using SWATH-MS, which led to quantification of 4501 proteins at the four distinct stages. Data analysis revealed that heat stress resulted in differential expression of 170 – 880 proteins at different stages of development. Differentially expressed proteins were associated with structural molecule activity, transporter activity, multicellular organismal process, amino acid activation and cell cycle. We conclude that protein abundances responded differentially to heat according to when it was applied, for example by supporting cell division and growth in the haploid germ line. Physiological and proteomic evidence supported our observation that the sensitivity of pollen to heat declined as the gametophyte matured. Identifying the stage-specific proteins will reveal key heat-responsive genes and new genetic tools for improved resilience of crops as climates keep warming.

18

A novel approach for biomarker discovery: spatial mapping of *N*-glycans on human knee osteoarthritis cartilage-bone tissue using matrix-assisted laser desorption/ionisation mass spectrometry imaging

Yea Rin Olivia Lee[1,2,3], **Matthew Briggs**[3], **Julia Kuliwaba**[1], **Paul Anderson**[2], **Peter Hoffmann**[3]

1. School of Medicine, The University of Adelaide, Adelaide, SA, Australia

2. Clinical and Health Sciences, Health and Biomedical Innovation, University of South Australia, Adelaide, South Australia, Australia

3. Future Industries Institute, University of South Australia, Mawson Lakes, South Australia, Australia

Objective:

Knee osteoarthritis (KOA) is the most common form of arthritis, but the biomolecular involvement of its onset and progression is controversial. Several studies have shown that the alterations of *N*-glycans on proteins contribute to the pathophysiology and progression of various diseases. However, the biomolecular distribution of *N*-glycans on KOA cartilage-bone tissue is poorly understood. Thus, the aim of this study was to spatially compare *N*-glycans from formalin-fixed paraffin-embedded (FFPE) cartilage-bone tissue of KOA patients and cadaveric controls (CTL).

Methods:

SCIEX
The Power of Precision

Human FFPE cartilage-bone tissue from end-stage KOA patients (2-Female; aged 58 and 79 years) and CTL individuals (2-Female; 44 and 54 years) was analysed by matrix-assisted laser desorption/ionisation mass spectrometry imaging (MALDI-MSI). In order to do so, we have developed and applied a novel and cost-effective sample preparation workflow in which commercial conductive ITO slides are pre-coated with gelatin and chromium potassium sulfate dodecahydrate to improve the adherence of tissue sections. Based on the theoretical masses, *N*-glycan peaks were manually selected, and ion intensity maps were generated using FlexImaging and SCiLS Lab software. Putative *N*-glycan structures were annotated using the following tools: GlycoMod, which calculates the theoretical monosaccharide composition, and Glycoworkbench to create individual *N*-glycan structures.

Results:

MALDI-MSI revealed differential *N*-glycan profiles between KOA patients and CTL individuals within the cartilage region only. Overall, 13 *N*-glycans were identified in KOA cartilage compared to 9 *N*-glycans in CTL cartilage, with approximately a 3-fold increase in the signal intensity. Interestingly, ion intensity maps of KOA cartilage-specific hybrid/complex-type *N*-glycans, *m/z* 1501.7 ± 0.5 Da, *m/z* 1647.2 ± 0.5 Da, and *m/z* 1663.4 ± 0.5 Da, showed higher intensity localisation to the superficial fibrillated area of degraded cartilage (cartilage histological grade 2-2.5) with underlying bone sclerosis, compared to the adjacent region with less damaged cartilage tissue (cartilage histological grade 0-1), associated with non-sclerotic bone.

Conclusion:

Our preliminary results demonstrate the novel application of MALDI-MSI to identify and localise KOA cartilage-specific *N*-glycans. The alterations of these hybrid/complex-type *N*-glycans could evolve into a potential cartilage degradation marker and may play an important role in the development of underlying bone sclerosis. This could also mean that *N*-glycans are a possible new target for treatment of cartilage degradation in patients with KOA.

19

Comprehensive proteomic profiling of blood extracellular vesicles via SWATH mass spectrometry: a new avenue for glioma tumour monitoring

Susannah Hallal[1, 2], Heng Wei[2, 3], Maggie Lee[2, 3], Hao-Wen Sim[4, 5, 6], Brindha Shivalingam[1, 2], Michael Buckland[7, 8, 2], Kimberley Alexander-Kaufman[7, 1, 2]

1. Neurosurgery department, Chris O'Brien Lifehouse, Sydney, NSW, Australia

2. BrainStorm Brain Cancer Research Group, Brain and Mind Centre, University of Sydney, Sydney, NSW, Australia

3. Department of Neuropathology, Royal Prince Alfred Hospital, Sydney, NSW, Australia

4. Department of Medical Oncology and NHMRC Clinical Trials Centre, Chris O'Brien Lifehouse, Sydney, NSW, Australia

5. Central Clinical School, University of Sydney, Sydney, NSW, Australia

6. The Kinghorn Cancer Centre, St Vincent's Hospital, Sydney, NSW, Australia

7. Neuropathology, Royal Prince Alfred Hospital, Sydney, NSW, Australia

8. Pathology, Sydney Medical School, The University of Sydney, Sydney, NSW, Australia

Diffuse gliomas (grades II-IV) are the most frequent and devastating primary brain tumours of adults. Currently, the clinical management for glioma patients involves a tissue biopsy for diagnostics and routine neuro-radiographic assessments, both of which have major limitations for accurate clinical assessment. With efforts to improve the clinical management of glioma, comes a growing trend to design minimally-invasive liquid biopsies (i.e. blood tests) that can routinely measure glioma-derived molecules in body fluids and allow for tumour evolution to be assessed in real-time. In this regard, extracellular vesicles (EVs; 30-1000 nm membranous particles) hold major promise as biomarker reservoirs. EVs encapsulate molecules that reflect the identity and molecular state of their cell-of-origin and their release is upregulated in neoplasia. EVs also cross the blood-brain-barrier into the circulation where they are stable and readily accessible. Despite their suitability as biomarkers, in-depth proteomic characterisation of circulating-EVs by traditional *shot-gun* proteomics has been hindered by the complexity of the blood and the co-isolation of highly abundant blood proteins. In this study, a data-independent acquisition (DIA) proteomics platform, sequential window acquisition of all theoretical fragment ion spectra (SWATH), was used in conjunction with a targeted data extraction strategy to achieve in-depth protein profiles of circulating-EVs from glioma patients. EVs were isolated by size exclusion chromatography from the plasma of pre-operative glioma II-IV patients and controls. Nanoparticle tracking and transmission electron microscopy confirmed the isolation of small-EV subtypes (< 200 nm). The plasma-EV peptides were sequenced by SWATH-MS, and the identities and quantities of the proteins were extracted using a custom spectral library comprised of 8662-protein species, developed using peptide samples from a range of glioma specimens, including cell lysates, tumour tissues and EVs. A total of 4054 proteins were identified in the plasma-EVs of all sample groups. Of these proteins, 463 changed significantly across the glioma and non-glioma cohorts (adj. $p < 0.05$), and included proteins previously reported to have significance in glioma-EVs. Principal component analysis showed excellent discrimination between the patient groups, with samples observed to cluster with their respective glioma subtype/grade. Using SWATH mass spectrometry we describe the most comprehensive proteomic plasma-EV profiles for glioma reported to-date, for which future studies using larger longitudinal cohorts could define a set of circulating-EV biomarkers, capable of stratifying glioma patients and detecting recurrence, progression and treatment resistance.

20

Proteomic analysis of adipose depots after intermittent fasting reveals visceral fat preservation mechanisms

Dylan J Harney[1], Michelle Cielesh[1], Renee Chu[1], Kristen C Cooke[1], David E James[1], Jacqueline Stöckli[1], Mark Larance[1]

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

Intermittent fasting (IF) is a beneficial diet that improves metabolic health independent of weight loss. Adipose tissue plays a key role in the fasting response, mobilising fatty acid reserves for use by other tissues such as the liver. There are many different adipose tissue depots

SCIEX
The Power of Precision

around the body and each is known to have a distinct physiological response to fasting. But the proteomic response of each distinct adipose depot is currently poorly defined. Here, we explore the response of three key adipose depots, visceral white adipose tissue (vWAT), subcutaneous white adipose tissue (scWAT) and brown adipose tissue (BAT), to every-other-day fasting (EODF) in mice using proteomics. A key change in both vWAT and scWAT was increased mitochondrial protein content associated with browning after EODF. These included metabolic enzymes crucial to pyruvate metabolism and the citric acid cycle as well as upstream cytosolic enzymes involved in glycolysis. These effects were correlated with increased fatty acid synthesis enzymes, such as the ACC1 protein complex, in both WAT depots, but not brown adipose tissue. UCP1 was differently regulated by EODF where it was increased in scWAT but not vWAT even although both had significant mitochondrial accumulation after EODF. Strikingly, EODF treatment downregulated lipolysis specifically in vWAT, mediated by a large decrease in the protein abundance of the catecholamine receptor (ADRB3). This decrease resulted in reduced intracellular lipolytic signalling, which lead to reduced free fatty acid release when challenged with an overnight fast and reduced sensitivity to lipolysis-inducing drugs. Together these changes would be important for the preservation of the visceral lipid store during EODF. Enrichment analysis also revealed downregulation of inflammatory collagen IV specifically in vWAT, which may contribute to the improved insulin sensitivity phenotype in these animals. Further analysis demonstrated decreased abundance of macrophage-specific and eosinophil-specific protein markers that could lead to a further reduction in adipose tissue inflammation. This resource is provided as a web-based interactive visualisation (larancelab.com/fat-eodf) for adipose depot-specific fasting adaptations in mice.

21

DIA based plasma proteomics for detection of diabetes related pathologies: readiness to support the FIELD clinical trial

Matthew B O'Rourke[1], Andrzrj januszewski[2], Alicia Jenkins[2], Mark P Molloy[1]

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

2. NHMRC Clinical Trials Centre, The University of Sydney , Camperdown, NSW, Australia

Diabetic related complications have long been reported as one of the key co-morbidity factors associated with both Type 1 and Type 2 diabetes. The FIELD study (Fenofibrate Intervention and Event Lowering in Diabetes) was developed to examine the effects of the lipid-lowering drug fenofibrate and its association with prevention of vascular pathologies such as diabetic retinopathy, limb amputation and cardiovascular incidents. The study was a double-blind placebo, controlled trial conducted with a cohort of 9,975 patients with Type 2 diabetes tracked over a period of 12 years. Blood samples were taken at baseline and during follow-up period, with initial collection after a 6 week run-in where all patients received the fenofibrate, followed by annual collections and 12 years post study close-out. Plasma samples were stored at -80C for subsequent analyses. A custom high throughput, DIA based plasma proteomics pipeline was developed and tested on an initial set of 15 patients samples that had blood collected at trial baseline and at the end of active run-in. We identified several protein markers including C4B binding protein, serotransferrin and kallistatin that show a significant change in regulation pre and post drug treatment. We are currently employing this approach to the wider FIELD study consisting of 1720 patient samples.

22

GlypNirO: An automated workflow for quantitative N- and O-linked glycoproteomic data analysis

Toan K. Phung[1], Cassandra L. Pegg[1], Ben L. Schulz[1, 2]

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

2. ARC Training Centre for Biopharmaceutical Innovation, The University of Queensland, Brisbane, Queensland, Australia

Mass spectrometry glycoproteomics is rapidly maturing, allowing unprecedented insights into the diversity and functions of protein glycosylation. However, quantitative glycoproteomics remains challenging. We developed GlypNirO, an automated software pipeline which integrates the complementary outputs of Byonic and Proteome Discoverer to allow high-throughput automated quantitative glycoproteomic data analysis. The output of GlypNirO is clearly structured, allowing manual interrogation, and is also appropriate for input into diverse statistical workflows. We used GlypNirO to analyse a published plasma glycoproteome dataset and identified changes in site-specific N- and O-glycosylation occupancy and structure associated with hepatocellular carcinoma as putative biomarkers of disease.

23

The post translational modification landscape of commercial beers using quantitative proteomics

Edward D Kerr[1], Cassandra L Pegg[1], Chris H Caboche[1], Toan K Phung[1], Mark T Howes[2], Kate Howell[3], Ben L Schulz[1]

1. University of Queensland, Brisbane, QLD, Australia

2. Newstead Brewing Company, Brisbane, QLD, Australia

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

Beer is one of the most popular beverages, with ~1.95 billion hectolitres produced annually around the globe. As a product from complex agricultural ingredients and processes, the final beer is highly molecularly complex. We used SWATH-MS to investigate the proteomic complexity and diversity of a wide range of commercial Australian beers. While the overall complexity of the proteome was modest, with contributions from barley and yeast proteins, we uncovered a very high diversity of post-translational modifications, especially proteolysis, glycation, and glycosylation. We used newly developed data analysis pipelines to efficiently extract and quantify site-specific PTMs from SWATH-MS data, and showed incorporating these features extended analytical precision. We found that the key differentiator of the beer

SCIEX
The Power of Precision

glyco/proteome was the brewery, followed by the beer style. Targeting our analyses on beers from a single brewery, Newstead Brewing Ltd, allowed us to identify beer style-specific features of the glyco/proteome, and show that abundant surface-active proteins from barley and yeast correlate with foam formation and stability.

24

Illuminating the dark lipidome of cancer using isomer-resolved lipidomics

Reuben SE Young[1], Andrew P Bowman[2], Kaylyn D Tousignant[3], David L Marshall[4], Shane R Ellis[2, 5], Berwyck LJ Poad[4], Martin C Sadowski[3, 6], Stephen J Blanksby[1, 4]

1. School of Chemistry and Physics, Queensland University of Technology, Brisbane, QLD, Australia

2. The Maastricht MultiModal Molecular Imaging Institute (M4I), Maastricht University, Maastricht, Netherlands

3. Australian Prostate Cancer Research Centre - Queensland, Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Faculty of Health, Queensland University of Technology/Translational Research Institute, Brisbane, QLD, Australia

4. Central Analytical Research Facility - Institute for Future Environments, Queensland University of Technology, Brisbane, Not US or Canada, Australia

5. Molecular Horizons and School of Chemistry and Molecular Bioscience, University of Wollongong, Woolongong, NSW, Australia

6. Institute of Pathology, University of Bern, Bern, Switzerland

One hallmark of cancer metabolism is the increased production of membrane lipids during cellular proliferation. This characteristic necessitates a transition from the predominant supply of fatty acids from extracellular sources to their generation through *de novo* synthesis mechanisms. It follows therefore, that the cancer lipidome should carry signatures of this metabolic shift. Recent research has highlighted that significant variation within the cellular lipidome of cancer can remain 'dark' to conventional lipidomics. In contrast, isomer-resolved lipidomics can illuminate the once 'dark lipidome' and reveal otherwise hidden changes in cancer cells lipid metabolism. Here we deploy these next-generation technologies to trace the metabolic fate of exogenously supplied fatty acids in prostate cancer cell lines.

Using a combination of stable isotope labelled fatty acids with high-resolution mass spectrometry that combines collision- and ozone-induced dissociation modes, we trace metabolic end-products of extracellular fatty acids and contrast these with isotopologues derived from *de novo* synthesis. These isomer-resolved strategies elucidate full molecular structure(s) of labelled glycerophospholipids, including assignment of fatty acyl chain position on the glycerol-backbone (*sn*-position) and carbon-carbon double bond location(s). This analysis reveals that glycerophospholipids carrying labelled-extracellular fatty acids (or their metabolites) have distinctive regiochemical profiles. Notably, extracellular palmitic acid and its desaturated progeny are enriched at the *sn*-1 position on the glycerol-backbone compared to *de novo*-derived isotopologues. This trait is also shared by extracellular stearic acid, however metabolites of stearic acid display enrichment at the *sn*-2 position. Probing these ratiometric isomer changes can therefore provide signatures of *de novo* and extracellular sourcing of fatty acids.

Unlike *in vitro* cell models, the reduced blood flow (and hence restricted nutrient supply) that is common to *in vivo* tumours will limit extracellular fatty acid supply for uptake. Based on the regioisomeric changes observed within the *in vitro* LNCaP cell experiments, we hypothesised that exploring the dark lipidome of tumour tissues would reveal otherwise hidden fatty acid metabolic behaviours. To test this, we obtained resected LNCaP xenograft tumours and combined isomer-resolved lipidomics with mass spectrometry imaging to monitor the regioisomeric lipid distribution across the tissues. This analysis revealed distinct domains across the tumour, indicating changes to fatty acid supply. Regions marred with the isomeric-signature of extracellular uptake were also found to contain high levels of arachidonic acid and DHA containing glycerophospholipids – fatty acids that must be exogenously sourced within mammals. Together, these findings implicate the lipase enzyme, PLA_2, may have a unique influence on tumour lipid remodelling.

25

In vivo phosphoproteomics of age-associated insulin-signalling in bone tissue

Benjamin Parker[1]

1. The University of Melbourne, Melbourne, Australia

There is growing evidence to suggest insulin signalling in bone tissue plays a critical role in the regulation of whole-body glucose and energy metabolism. However, a systems biology analysis to map *in vivo* signalling architecture has yet to be performed. Furthermore, whether this network of signalling is rewired during ageing and insulin-resistance is unknown. To this effect, we present the first mouse bone phosphoproteomics landscape from 8-week and 73-week old C57BL/6J mice following insulin or vehicle treatment. Tryptic bone lysates were labelled with Tandem Mass Tags and phosphopeptides enriched using titanium dioxide and analysed by multi-dimensional ultra-high-pressure liquid chromatography coupled to tandem mass spectrometry (2DLC-MS/MS), to compare *in vivo* age-related insulin signalling and to identify key regulators of the osteo-insulin pathway. The mouse bone phosphoproteome revealed 16,502 phosphosites that were mapped to 4528 bone phosphoproteins. Of these, we identified 4,661 novel phosphosites (~28% of the total phosphosites) that are associated with proteins involved in several regulatory mechanisms including the insulin signalling pathway. Comparison of insulin activated 8-week old bone to the control revealed 2,953 insulin-regulated phosphosites ($q<0.05$; Limma moderated t-test with Benjamini Hochberg FDR), including several substrates of insulin-responsive kinases such as AKT and mTOR. Furthermore, an age associated differential expression of 2,175 insulin-regulated phosphosites was observed in 8-week versus the 73-week old bone revealing dramatic rewiring and defects in insulin signalling ($q<0.05$; Limma ANOVA with Benjamini Hochberg FDR). We next utilised machine-learning based on kinase motif and expression profiles to predict substrates of Akt/S6K, mTOR and PKA kinase. Integration of these data with systems genetic analysis of human GWAS against bone mineral density enabled us to prioritise high confidence targets for functional validation. We further integrated these data with phosphosite evolutionary conservation analysis to prioritise functional phosphorylation sites with an emphasis on zebrafish allowing us to

SCIEX
The Power of Precision

establish a high-throughput functional screen of zebrafish bone. We hope our functional analysis of the bone phosphoproteome will further enhance our understanding of the signalling mechanisms controlling bone biology and whole-body energy metabolism.

26

Temporal quantitative proteomics and phosphoproteomics reveals ubiquitin proteasome system is essential to mount adaptive response to hypoxia in melanoma cells

Keshava Datta[1], **Parthiban Periasamy**[1,2], **Sonali Mohan**[1,2], **Rebekah Ziegman**[1], **Harsha Gowda**[1,2,3]

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

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

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

Hypoxia is a common feature in various solid tumors including melanoma. 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 cancer cell survival in hypoxic environments can reveal potential vulnerabilities that can be exploited to improve efficacy of anti-cancer therapies. We carried out temporal proteomic and phosphoproteomic profiling in melanoma cell lines to identify hypoxia induced protein expression and phosphorylation changes. By employing TMT based quantitative proteomics strategy, we identified and quantified >7,000 proteins and >10,000 phosphosites in melanoma cell lines grown in hypoxia. Proteomic data showed metabolic reprogramming as one of the prominent adaptive responses in hypoxia. We identified several novel hypoxia mediated phosphorylation changes that have not been reported before. They reveal kinase-signaling pathways that are potentially involved in modulating cellular response to hypoxia. In addition to known protein expression changes, we identified several novel proteomic alterations associated with adaptive response to hypoxia. We show that cancer cells are dependent on ubiquitin-proteasome system to survive in hypoxia. Inhibition of proteasome activity affects cell survival and may provide a novel therapeutic avenue to target cancer cells in hypoxia. Our study can serve as a valuable resource to pursue several novel candidates to target hypoxia in cancers and improve efficacy of anticancer therapies.

27

Controlling the controllers: the post-translational regulation of histone methylation enzymes in yeast

Ryan J Separovich[1], **Mandy WM Wong**[1], **Tyler R Chapman**[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 key epigenetic modification that is central to the regulation of eukaryotic transcription. In the lower eukaryote, *Saccharomyces cerevisiae* (Baker's yeast), it is controlled by a system of four methyltransferases (Set1p, Set2p, Set5p, and Dot1p) and four demethylases (Jhd1p, Jhd2p, Rph1p, and Gis1p). While the histone targets for these enzymes are well characterised, the connection of the enzymes with the intracellular signalling network and thus their regulation is poorly understood, in yeast and in all other eukaryotes. Here we report the detailed mass spectrometric characterisation of the eight *S. cerevisiae* enzymes, and show that they carry a total of 75 phosphorylation sites, 93 acetylation sites, and two ubiquitination sites. All enzymes are subject to phosphorylation, although demethylases Jhd1p and Jhd2p contained one and five sites respectively whereas other enzymes carried 14 to 36 sites. Enzymes with large numbers of sites are potential integrators of signalling data[1]. Phosphorylation was absent or under-represented on catalytic and other domains but strongly enriched for regions of disorder on methyltransferases, suggesting a role in the modulation of protein-protein interactions. We show that a phosphorylation cluster within an acidic and intrinsically disordered N-terminal region of methyltransferase Set2p regulates H3K36 methylation levels *in vivo,* thus supporting the functional relevance of disordered phosphosites. Strikingly, some phosphorylation events were evolutionarily conserved between yeast Set1p, Set2p, Rph1p, Gis1p, and their respective mammalian orthologs, and further sequences that contain phosphosites in yeast enzymes show relevant amino acid conservation in human. While most kinases upstream of the yeast histone methylation enzymes remain unknown, we model the possible connections between the signalling network and the histone-based gene regulatory system and propose an integrated regulatory structure. Our results provide a foundation for future, detailed exploration of the role of specific kinases and phosphosites in the regulation of histone methylation.

1. Separovich, R.J., Pang, C.N.I., Wilkins, M.R. (2020) Controlling the controllers: regulation of histone methylation by phosphosignalling. Trends in Biochemical Sciences 45 (12), 1035-1048.

28

Exploiting *pglB* negative and positive *Campylobacter jejuni* to identify occupied and unoccupied sites of *N*-linked protein glycosylation

Joel A Cain[2,1], **Ashleigh L Dale**[2,1], **Stuart J 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 a Gram negative human pathogen that contains a unique *N*-linked protein glycosylation system that is encoded by the *pgl* gene cluster. The heptasaccharide *N*-glycan is attached to membrane-associated proteins in the periplasm by the PglB

SCIEX
The Power of Precision

oligosaccharyltransferase at the consensus sequon Asp/Glu-X-Asn-X-Ser/Thr. Over 500 sequons are contained in the *C. jejuni* proteome and >130 sites have been experimentally verified. We performed quantitative proteomics of *C. jejuni* NCTC11168 compared with a *pglB* deletion mutant (Δ*pglB*) and a *pglB* restored strain (Δ*pglB*::*pglB*) with the aim of identifying differentially regulated sequon-associated peptides. We quantified 164 sequon-associated peptides corresponding to 120 unique sequons, of which 73 contained previously experimentally verified glycosites. Analysis of differential abundances for these peptides allowed us to make inferences regarding site-level occupancy by the N-glycan *in vivo.* By referencing the subcellular topology of modified proteins we were able to show that these predictions were in agreement with the presentation of sequons relative to the N-glycosylation machinery. Finally we utilised a series of proteases (trypsin, Glu-*C*, Asp-*N*, thermolysin, pepsin and chymotrypsin) to define the glycoproteome of wild-type and *pglB* restored *C. jejuni.* We identified 1915 glycopeptides corresponding to 140 glycosylation sites, including the identification of 31 sites not previously observed and 17 highlighted by the Δ*pglB* proteomic analysis.

29

Uncovering glycoproteome signatures associated with prostate and colorectal cancer progression using integrated 'omics

Rebeca Kawahara[1], Liisa Kautto[1], Seong Beom Ahn[2], Giuseppe Palmisano[3], Morten Thaysen-Andersen[1, 4]

1. Macquarie University, Sydney, NSW, Australia

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

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

4. Biomolecular Discovery Research Centre (BDRC), Macquarie University, Sydney, NSW, Australia

Glycosylation is a prevalent and complex type of protein modification central to many biological processes underpinning cancer development and progression. Recent advances in separation science, mass spectrometry and informatics have enabled glycoscientists to study molecular features of the heterogenous cancer glycoproteome including the system-wide mapping of glycosylation sites and glycan compositions to discrete sites. However, significant bottlenecks still exist in determining the glycan fine structures, their site-specific micro- and macro-heterogeneity, their dynamics and cellular origins within in the complex tumour environment. To this end, we have developed an integrated LC-MS/MS-based glycomics, glycoproteomics and proteomics approach that makes use of the Tissue Atlas and Blood Atlas to identify glycoproteome signatures and progression-related changes within the prostate cancer (PCa) and colorectal cancer (CRC) tumour environment. This strategy was first applied to surgically-removed PCa tissues spanning five histological grades (n = 10/grade) and benign prostatic hyperplasia control tissues (n = 5). Grade-specific dynamics of pauci- and oligomannosylation of known bone-marrow and prostate-derived glycoproteins, respectively, and increased *N*-glycan branching and core 2-type *O*-glycosylation of known extracellular matrix glycoproteins were found to be key changes associated with PCa progression [1]. In another study, peripheral blood mononuclear cells (PBMNCs) and fresh snap-frozen CRC tissue samples collected from 28 CRC patients spanning four different disease stages (n = 7/stage) and 8 tissue samples from clinically-relevant control individuals were investigated using the same integrated omics strategy. Paucimannosylation was shown to be a significant signature associated with CRC progression in both the tumour tissues and PBMNCs. Correlation analysis between the glycome and proteome datasets showed that the paucimannosidic glycan levels correlated closely with the expression of granulocytic protein markers in the tumour tissues and monocytic protein markers in the PBMNC samples. Collectively, these findings have advanced our knowledge of the molecular and cellular makeup of the PCa and CRC tumour microenvironment by documenting, amongst other key findings, that paucimannosidic glycan-containing proteins of innate immune cell origin are strong yet still largely underexplored signatures of PCa and CRC that may play important roles associated with tumour progression.

1. Kawahara, R., Recuero, S., Srougi, M., Leite, K. R. M., Sr., Thaysen-Andersen, M., & Palmisano, G. (2020). The complexity and dynamics of the tissue glycoproteome associated with prostate cancer progression. Mol Cell Proteomics. doi:10.1074/mcp.RA120.002320

30

Microbial glycoproteomics; how do you find glycopeptides when you're not sure what you're looking for?

Nichollas E Scott[1]

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

Microbial glycoproteomics is a rapidly growing field which seeks to characterise glycosylation at a proteome scale in Bacteria and Protozoa. Although interest in microbial glycoproteins as potential biomarkers or vaccine candidates has increased the atypical nature of these events has limited the study of microbial glycosylation to a hand full of specialised laboratories. To improve the accessibility of microbial glycosylation we have sought to develop agnostic bioinformatic and enrichment approaches to streamline the analysis of microbial glycoproteomes using members of the Burkholderia genus as model systems. Using open searching we have now shown that microbial glycosylation events can be identified in an unsupervised manner without prior knowledge of the glycans used for glycosylation. Using open searching determined glycan profiles we have also shown that glycan utilisation can be rapidly compared allowing genus wide comparisons of glycosylation diversity. To streamline the identification of microbial glycopeptides we have also explored the use of high-field asymmetric waveform ion mobility spectrometry (FAIMS) fractionation demonstrating that at high compensation voltages (CVs) short aliphatic glycopeptides can be readily isolated. This FAIMS based glycopeptide enrichment enables glycopeptide selection from complex samples providing an alternative means to identify glycopeptides recalcitrant to hydrophilic based enrichment approaches. Combining these approaches, we have now shown the glycoproteome of Burkholderia species are nearly 10 times larger than initially thought (~200 glycoproteins per species) and that O-linked glycosylation only occurs on serine residues. Furthermore, by combining multiple FAIMS CVs within a single analytical run this enables both proteome and glycoproteome analysis of even limited samples. Taken together, these results demonstrate that open searching and FAIMS enrichment approaches are valuable tools for glycoproteomic analysis.

SCIEX
The Power of Precision

31

Characterisation of the O-glycoproteome of *Porphyromonas gingivalis*

Paul D Veith[1], **Nichollas E Scott**[2], **Eric C Reynolds**[1]

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

2. Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria, Australia

Porphyromonas gingivalis is a Gram-negative oral pathogen belonging to the Bacteroidetes phylum. A general O-glycosylation system targeting exported proteins has been reported in this phylum and the major O-glycans have been characterized in a few species. In this study we identify the major O-glycans in *P. gingivalis* and identify the glycosylation sites. Proteins from membrane fractions were either fractionated by SDS-PAGE or subjected to glycoprotein enrichment using ZIC-HILIC. Proteins were digested with trypsin and analysed by LC-MS/MS with an Orbitrap equipped with a FAIMS source. Two different FAIMS settings were used and peptides containing the HexNAc oxonium ion were selected for ETD and CID fragmentation to aid elucidation of the glycosylation site and glycan sequence respectively. In a complementary approach, protein fractions were also partially deglycosylated with trifluoromethanesulfonic acid (TFMS) prior to MS analyses. So far, ~320 unique peptide sequences encompassing ~200 sites from ~120 proteins have been identified. The proteins included surface proteins, outer membrane proteins, lipoproteins and periplasmic proteins. The proteins exhibited between one and seven glycosylation sites which conformed to the known motif of D[T/S]X, where X was known to include A,V,L,I, M and T. In *P. gingivalis* we now show that X can also be S, C, F and G. The two major glycans have delta masses of 1436 and 1394 respectively and appear to be the same apart from an additional acetyl group. The preliminary glycan sequence of the 1436 Da form is Hex-(HexA)-Hex-(dHex)-dHex-HexNAc-X-HexNAc where X=238 Da. With the use of deglycosylation, the advantage of being able to select for HexNAc-containing glycopeptides was lost, however it was beneficial for identifying additional glycopeptides and was very helpful for glycan sequencing since it truncated the glycans to form a ladder. In addition, the truncated glycoforms yielded ETD fragmentation patterns more amenable to identification of glycosylation sites.

32

In-house packed porous graphitised carbon columns for nano-liquid chromatography-mass spectrometry analysis of *N*-glycans

Matthew T Briggs[1], **Mark R Condina**[1], **Clifford Young**[1], **Gurjeet Kaur**[2], **Martin K Oehler**[3], **Peter Hoffmann**[1]

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

2. Institute for Research in Molecular Medicine, Universiti Sains , Pulau Pinang, Malaysia

3. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

Glycosylation is one of the most important post-translational modifications (PTMs) as it is responsible for the homeostasis of cellular immunity and adhesion, and protein translation and degradation. There are two main types of glycosylation; *N*-linked glycans which are complex sugars attached to asparagine residues and *O*-linked glycans which are complex sugars attached to serine or threonine residues. Of these two types, it is estimated that 60% of proteins are glycosylated with 90% of glycoproteins being *N*-glycosylated. One of the most common analytical techniques to analyse *N*-glycans is liquid chromatography coupled to mass spectrometry (LC-MS), which results in the separation of *N*-glycans followed by MS acquisition and quantitative analysis. Hydrophilic interaction chromatography (HILIC) is a well-established separation technique whereby polar compounds are separated based on hydrogen bonding, ionic interactions and dipole-dipole interactions. Although HILIC is advantageous for separating *N*-glycans, an alternative separation technique, such as porous graphitised carbon (PGC), possesses improved resolving capability that cannot be achieved by HILIC. For multiple groups, PGC capillary columns have become the gold standard due to this resolving capability. However, there currently are limited commercial options for such columns. Therefore, our group have successfully developed an in-house packed PGC nano column using Hypercarb material (3um) and a 15cm glass capillary (75µm ID, 360µm OD). To assess these columns, we firstly released *N*-glycans from an in-house glycoprotein standard and formalin-fixed paraffin-embedded (FFPE) egg white as a quality control. These *N*-glycans were then separated on an Agilent 1290 Infinity II UHPLC with an Agilent Infinity UHPLC nanodaptor coupled to an Agilent 6550 iFunnel Q-TOF Mass Spectrometer. This optimsied workflow was then applied to *N*-glycans released from FFPE ovarian and endometrial cancer tissue to (1) further validate this novel PGC nano column and (2) investigate *N*-glycan alterations between these gynaecological cancers which may lead to the discovery of diagnostic markers or therapeutic targets. It is envisioned that this novel PGC nano column could be packed in-house by other groups, thereby overcoming the caveat of commercial options while maintaining sensitivity and reproducibility.

33

Serum *N*-Glycomics Stratifies Bacteremic Patients Infected with Different Pathogens

Sayantani Chatterjee[1,2], **Harry C. Tjondro**[1,2], **Rebeca Kawahara**[1,2], **Marni Nenke**[3,4,5], **David J. Torpy**[3,4], **Morten Thaysen-Andersen**[1,2]

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

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

3. Endocrine and Metabolic Unit, Royal Adelaide Hospital, Adelaide, SA, Australia

4. School of Medicine, University of Adelaide, Adelaide, SA, Australia

5. Department of Endocrinology and Diabetes, The Queen Elizabeth Hospital, Woodville South, SA, Australia

Bacteraemia, the presence of pathogens in the bloodstream, is associated with long-term morbidity and is a potential precursor condition to life-threatening sepsis. Timely detection of bacteraemia is therefore critical to reduce patient mortality, but existing methods lack the

SCIEX
The Power of Precision

precision, speed, and sensitivity to effectively stratify bacteremic patients from uninfected individuals and identify the infecting pathogen. Herein, we test the potential of quantitative serum *N*-glycomics performed using an established porous graphitised carbon LC-MS/MS platform to stratify bacteremic patients infected with different pathogens i.e. *Escherichia coli*, *Staphylococcus aureus* (both n = 11), *Pseudomonas aeruginosa*, and *Streptococcus viridans* (both n = 5) from healthy donors (n = 39). In total, 62 *N*-glycan isomers spanning 41 glycan compositions primarily comprising complex-type core fucosylated and α2,3- and α2,6-sialylated structures were profiled across all samples using label-free relative quantitation. Excitingly, unsupervised hierarchical clustering of the serum *N*-glycome data accurately separated the patient groups. *P. aeruginosa*-infected patients displayed the most aberrant serum *N*-glycome involving elevated core fucosylation and reduced α2,6-sialylation relative to healthy donor sera as further supported by quantitative lectin blotting using *Aleuria aurantia* lectin (AAL) and *Sambucus nigra* lectin (SNA), respectively. Importantly, receiver operating characteristic analyses demonstrated that a single *N*-glycan isomer can effectively stratify each of the four patient groups from the healthy donor cohort (ROC 0.93-1.00). We conclude that the serum *N*-glycome represents a new hitherto unexplored class of potential diagnostic markers for bloodstream infections.

34

Multi-omic profiling of Australian Paralysis tick, *Ixodes holocyclus* in understanding the symbiosis with *Midichloria mitochondrii.*

Amrita Vijay[1], Thomas Karbanowicz[2], Quentin Gouil[1], Balu Balan[1, 3], Samantha J Emery-Corbin[1], Louise Baker[1], Stefano Gaiarsa[4], Jan Riemer[5], Fabrizia Stavru[6], Davide Sassera[7], Peter Czabotar[1], Tony Papenfuss[1], Aaron R Jex[1, 3]

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

2. The University of Queensland, Brisbane, Queensland, Australia

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

4. Microbiology and Virology unit at Policlinico San Matteo, Fondazione IRCCS, Pavia, Province of Pavia, Italy

5. Department for Chemistry, Institute for Biochemistry , University of Cologne, Cologne, Germany

6. Unité de Biologie Evolutive de la Cellule Microbienne, Institut Pasteur, Paris, France

7. Department of Biology and Biotechnology, University of Pavia, Pavia, Italy

Ancient endosymbiotic interactions between alphaproteobacteria and archaea/proto eukaryotes gave rise to the mitochondrion, on which all eukaryotic evolution is built. Intriguingly, a novel Alphaproteobacterium, *Candidatus Midichloria mitochondrii*, is the only known example of an organism that has evolved to colonise these organelles. This bacterium lives in the cellular cytoplasm and the intermembranal space (IMS) of the mitochondrion of its host, *Ixodes ricinus* (the European castor bean tick). and this is known as Intramitochondrial tropism, (IMT) The same species is also found in the Australian paralysis tick, *Ixodes holocyclus*. However, in the latter host, it is confined to the cytoplasm only. The functional consequences of IMT for the host, whether beneficial or pathogenic, are not known. It is not clear if or how the bacterium invades the mitochondria, nor whether or how it is able to repair the mitochondrial outer membrane upon invasion. In *I. ricinus, M. mitochondrii* appears to replicate within the IMS and degrade the mitochondrial matrix, yet this does not appear to result in cell death. Understanding this fascinating endobiotic interaction has intriguing implications for evolutionary and cellular biology. Beyond its relevance to tick biology, understanding IMT has the potential to greatly improve our understanding of mitochondrial targeting, membrane repair and degradation. These have broad implications in mitochondrial disorders, cancers and aging.

However, study of this system is inhibited by a number of substantial obstacles. In addition to the requirement for methods to cultivate these bacteria *in vitro*, or to functionally perturb them or their host cells, there are major, fundamental knowledge gaps in the basic biology of ticks. There is a need for greater understanding of their cell death systems, and for resources to study them, including reference genomes, transcriptomes or proteomes. Here, we described the development of a reference standard genome and transcriptome for *I. holocyclus*. We curated this resource to undertake proteomic analyses of adult male and female *I. holocyclus*, in both the presence and absence of blood-feeding. We then undertook detailed *in silico* analyses of *I. holocyclus* proteins to identify conserved orthologs of mitophagy and apoptotic proteins. This work underpins our efforts to explore the host cell changes associated with *M. mitochondrii* endobiosis. This will be a critical resource for comparison with *I. ricinus* laboratory colonies with or without *M. mitochondrii* endobiosis and IMT.

35

Comprehensive analysis of histone post-translational modifications in mouse CD8$^+$ T cells.

Swati Varshney[1], Annabell Bachem[2], Sammy Bedoui[2], Michael Leeming[1], Ching-Seng Ang[1], Shuai Nie[1], Nicholas Williamson[1]

1. Bio21 Molecular Science & Biotechnology Institute, University of Melbourne, Parkville, VIC, Australia

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

Post-translational modifications (PTM's) on histones plays a crucial role in determining the immune cell's fate in terms of differentiation, development, and disease. Mass spectrometry has emerged as a powerful tool for global histone PTM studies. Analysis of histone PTM's is very complex owing to the enormous number and types of isobaric histone peptides. Histone has four forms (H2A, H2B, H3, and H4) attached with a linker H1 and multiple isoforms.

We used a "bottom-up" MS-based proteomics strategy to identify histone PTMs and to determine their abundance in CD8$^+$ T cells of the mouse. Briefly, histones proteins were subjected to chemical derivatization by Propionic anhydride to generate peptides suitable for MS analysis. The MS was operated in DDA mode and data analysis was carried out with Epiprofile 2.1 basic and Proteome discoverer 2.3 (PD) against the UniProt database. Epiprofile is a MATLAB-based software specially developed for Histone PTM's analysis using Hela cells. EpiProfile was designed to extract ions for a defined set of histone peptides (5 Histones, 24 isoforms, 70 peptides, and 213 modified/unmodified forms) by determining the RT for non-isobaric peptides by the MS1 and discriminate isobaric peptides using unique fragment ions in the MS2 spectra. We found that the Epiprofile worked well however comparison of the Epiprofile results with our PD

SCIEX
The Power of Precision

results revealed that only 31 peptides were found in common between the two analysis tools. The limited library of peptides in Epiprofiler limiting its utility. Using PD we found 5 Histones, 28 isoforms, and 700 modified/unmodified forms of 160 histone peptides. The limited overlap between two tools is a direct example of the complexity of histone modifications and the differences between the Hela cells used for the development of Epiprofile and the $CD8^+$ T cells used in our own study. This work will now form the basis of our future work examining histone modifications and epigenetic regulation of $CD8^+$ T cells.

36

Gene transfection: An alternative approach to discover epitopes of highly pathogenic pandemic pathogens

AKM Mr Muraduzzaman[1], Patricia Dr Illing[1], Anthony Prof Purcell[1], Nicole Dr Mifsud[1]

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

Background: In last two decades, the world has experienced several outbreaks of potential pandemic viruses including highly pathogenic avian influenza (H5N1), Middle East Respiratory Syndrome Corona Virus (MERS-CoV), influenza A (H1N1)pdm09, Ebola, Zika and, most recently, Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2). As emerging pathogens, there is no pre-existing immunity at the population level, generating a need for effective vaccines. In virally infected cells, antigens are processed by proteasome complex and transported to Endoplasmic Reticulum where it further trims to short peptides (8-11 aa) and presented through Major Histocompatibility complex (MHC) on to the cell surface for the T cells scanning and response. Peptide-based T cell vaccines are promising candidates to induce broadly protective immunity against multiple circulating strains of the pathogen (e.g. influenza), by targeting conserved proteins of the pathogen. This study provides an alternative approach to examine the immunopeptidome of avian influenza H5N1 by transfecting antigen-presenting cells (APCs) with targeted influenza proteins and comparing these data with natural influenza infection using mass spectrometry. Given the low avian influenza H5N1 susceptibility incidences in South-Asia region, we are interested in exploring the immunopeptidome presented by Human Leukocyte Antigen (HLA)-A*33:03 allele (highly expressed in South Asian populations) to determine whether there are protective effects in play.

Methods: Class I reduced (C1R) APCs expressing HLA-A*33:03 allele (C1R.A*33:03) were used as an APCs for transfection of targeted genes (hemagglutinin (HA), polymerase basic 2 (PB2), matrix & nucleoprotein (NP)) of avian influenza H5N1 and influenza A/X31. Transfected or infected C1R.A*33:03 was grown to 10^9 cells to generate cell pellets, underwent immunoaffinity purification, RP-HPLC and LC-MS/MS to identify influenza-specific peptides. Raw data were analysed using PEAKS software version X+.

Results: We have interrogated the HLA-A*33:03-restricted immunopeptidome of selected proteins derived from both avian influenza H5N1 (NP; 18 peptides identified) and influenza A/X31 (total of 40 peptides identified; HA=25, matrix=7, NP=18) transfectants. We have identified 41 peptides (HA=8, matrix=7, NP=22) from the influenza A/X31 naturally infected cells. HLA-binding motifs and length distribution were similar between transfectants and natural infection. Investigation of avian influenza H5N1 for HA, PB2 and matrix proteins is currently underway.

Conclusion: Using this workflow we have successfully identified large numbers of H5N1 derived HLA-A*33:03 restricted peptides. Thus, utilisation of transfectants offers a novel approach for identification of peptide targets derived from highly infectious viruses for development as immunotherapeutics, circumventing the need to handle such pathogens in high containment facilities.

37

Functional phosphoproteomic analysis of insulin signalling in aging bone

Mriga Dutt[1], Ronnie Blazev[1], Audrey Chan[1], Natalie K.Y. Wee[2], Jeffrey Molendijk[1], Luoping Liao[1], Vanessa R. Haynes[1], Yaan Kit-Ng[1], Veronica U. Sokolov[1], Kelly A. Smith[1], Garron T. Dodd[1], Gordon G. Lynch[1], Natalie A. Sims[2], Benjamin L. Parker[1]

1. Department of Physiology, Faculty of Medicine, School of Biomedical Sciences, University of Melbourne, Parkville, VIC, Australia

2. St. Vincent's Institute of Medical Research and Department of Medicine at St. Vincent's Hospital, University of Melbourne, Parkville, VIC, 3010

There is growing evidence to suggest insulin signalling in bone tissue plays a critical role in the regulation of whole-body glucose and energy metabolism. However, a systems biology analysis to map *in vivo* signalling architecture has yet to be performed. Furthermore, whether this network of signalling is rewired during ageing and insulin-resistance is unknown. To this effect, we present the first mouse bone phosphoproteomics landscape from 8-week and 73-week old C57BL/6J mice following insulin or vehicle treatment. Tryptic bone lysates were labelled with Tandem Mass Tags and phosphopeptides enriched using titanium dioxide and analysed by multi-dimensional ultra-high-pressure liquid chromatography coupled to tandem mass spectrometry (2DLC-MS/MS), to compare *in vivo* age-related insulin signalling and to identify key regulators of the osteo-insulin pathway. The mouse bone phosphoproteome revealed 16,502 phosphosites that were mapped to 4528 bone phosphoproteins. Of these, we identified 4,661 novel phosphosites (~28% of the total phosphosites) that are associated with proteins involved in several regulatory mechanisms including the insulin signalling pathway. Comparison of insulin activated 8-week old bone to the control revealed 2,953 insulin-regulated phosphosites ($q<0.05$; Limma moderated t-test with Benjamini Hochberg FDR), including several substrates of insulin-responsive kinases such as AKT and mTOR. Furthermore, an age associated differential expression of 2,175 insulin-regulated phosphosites was observed in 8-week versus the 73-week old bone revealing dramatic rewiring and defects in insulin signalling ($q<0.05$; Limma ANOVA with Benjamini Hochberg FDR). We next utilised machine-learning based on kinase motif and expression profiles to predict substrates of Akt/S6K, mTOR and PKA kinase. Integration of these data with systems genetic analysis of human GWAS against bone mineral density enabled us to prioritise high confidence targets for functional validation. We further integrated these data with phosphosite evolutionary conservation analysis to prioritise functional phosphorylation sites with an emphasis on zebrafish. This enabled us to establish a CRISPR/Cas9 loss-of-function screen in zebrafish to identify novel phosphoproteins regulating bone formation and glucose metabolism. We present preliminary and ongoing advances in establishing high-throughput functional screening of zebrafish bone integrated with proteomics. We hope our functional analysis of the bone phosphoproteome will further enhance our understanding of the signalling mechanisms controlling bone biology and whole-body energy metabolism.

SCIEX
The Power of Precision

38

Mapping Lipidomic Complexity Using Advanced Mass Spectrometry Imaging Technologies

Shane Ellis[1, 2]

1. University of Wollongong, Wollongong, NSW, Australia

2. Maastricht MultiModal Molecular Imaging Institute (M4I), Maastricht University, Maastricht, Netherlands

Mass spectrometry has evolved into an indispensable tool in the (bio)chemical sciences. However, analyses are typically performed on extracted molecules from homogenised samples meaning the spatial context in which molecules are present is sacrificed. Understanding not only what molecules are present, but *where* there are located within a tissue or a cell and how they are altered during the progression of disease is of utmost importance for understanding complex biochemical processes occurring in spatially heterogeneous tissues. In this presentation, I will demonstrate how recent advances in mass spectrometry imaging (MSI) now provide a suite of potent approaches to study chemical diversity direct from complex surfaces like biological tissues. Unlike other imaging approaches, MSI allows the parallel imaging of thousands of molecules without the need for labels and with cellular-level resolution.

I will give an introduction to matrix-assisted laser desorption/ionisation-MSI (MALDI-MSI) and its applications for visualising biochemical processes within tissues. Particular emphasis will be placed on innovative technology developments made by my group that enable enhanced chemical resolution and detection of previously undetectable molecular classes. These include the development of MALDI-MSI ion sources for Orbitrap mass spectrometers [1], highly parallelised MS/MS imaging acquisitions sequences [2], laser-induced post-ionisation methods [3-4] and the imaging of isomeric lipids using selective gas-phase ion/molecule reactions [5]. The coupling of MSI with high resolving power analysers also enables localised molecular dynamics to be visualised in tissues, and we have applied this in mouse models of respiratory disease. These advances pave the way of dynamic MSI where lipid turnover rates are provided concurrently with imaging data, thus providing a new window into tissue heterogeneity.

References

(1) Belov, M. E., Ellis, S. R., Dilillo, M., Paine, M. R. L., Danielson, W. F., Anderson, G. A., de Graaf, E. L., Eijkel, G. B., Heeren, R. M. A., McDonnell, L. A., Design and Performance of a Novel Interface for Combined Matrix-Assisted Laser Desorption Ionization at Elevated Pressure and Electrospray Ionization with Orbitrap Mass Spectrometry. *Anal. Chem.* **2017**, *89,* 7493-7501.

(2) Ellis, S. R., Paine, M. R. L., Eijkel, G. B., Pauling, J. K., Husen, P., Jervelund, M. W., Hermansson, M., Ejsing, C. S., Heeren, R. M. A., Automated, Parallel Mass Spectrometry Imaging and Structural Identification of Lipids. *Nat. Methods* **2018**, *15,* 515-518.

(3) Ellis, S., Soltwisch, J., Paine, M. R. L., Dreisewerd, K., Heeren, R., Laser Post-Ionisation Combined with a High Resolving Orbitrap Mass Spectrometer for Enhanced MALDI-MS Imaging of Lipids. *Chem. Commun.* **2017**, *53,* 7246-7249

(4) Bowman, A. P., Bogie, J. F. J., Hendriks, J. J. A., Haidar, M., Belov, M., Heeren, R. M. A., Ellis, S. R., Evaluation of Lipid Coverage and High Spatial Resolution MALDI-Imaging Capabilities of Oversampling Combined with Laser Post-Ionisation. *Anal. Bioanal. Chem.* **2020**, *412,* 2277-2289.

(5) Paine, M. R. L., Poad, B. L. J., Eijkel, G. B., Marshall, D. L., Blanksby, S. J., Heeren, R. M. A., Ellis, S. R., Mass Spectrometry Imaging with Isomeric Resolution Enabled by Ozone-Induced Dissociation. *Angew. Chem. Int. Ed.* **2018**, *57,* 10530-10534.

39

Multiplexed bioaffinity mass spectrometry (MX-BaMS): Highly-parallelised and rapid screening of complex mixtures for ligands that bind to druggable protein targets

Sherrie Liu[1], Giang T. H. Nguyen[1], Jack L. Bennett[1], Daniel L. Winter[1], Dominic J. Glover[1], William A. Donald[1]

1. University of New South Wales, UNSW SYDNEY, NSW, Australia

In drug discovery, natural product libraries have the distinct advantages of high molecular diversity, complex stereochemistry, and biochemical specificity. Although far less than 10% of the world's biodiversity has been evaluated for potential biological activity, the exploration of natural product libraries has declined since the 1970s in favour of synthetic libraries that can be more readily screened. One potential approach to substantially improve the screening of natural products is to use native mass spectrometry (MS), owing to its high sensitivity, low sample consumption, and rapid analysis time. However, conventional native MS is limited by its intolerance to non-volatile salts and impurities. In practice, current native MS methodology is usually limited to screening one isolated candidate compound at a time, which substantially limits throughput.

Recent research has demonstrated that up to 11 protein-ligand interactions can be measured directly from a single mass spectrum by using nanoscale ion emitters in native MS.[1-2] The use of nanoscale emitters can significantly reduce the adduction non-volatile salts common to complex mixtures such as natural extracts.[1-3] Here, we report the application of nanoscale emitters in native MS to achieve highly parallelised screening of complex mixtures for ligands of druggable protein targets. The proposed workflow, entitled multiplexed bioaffinity mass spectrometry (MX-BaMS), involves: (i) incubation of the target protein with a crude natural extract and (ii) native MS detection of intact protein-ligand complexes. As a proof-of-concept, three carbonic anhydrases (CA) were selected as protein targets, and five natural products were screened. A total of 15 hits were detected, corresponding to 11 unique compounds putatively identified using metabolomics data from liquid chromatography tandem MS, and multiple hits were confirmed using authentic standards. Two new CA ligands were identified, and an entirely new CA-binding scaffold was discovered. In addition, recent related research in the application of MX-BaMS to the Papain-like protease from SARS-CoV-2, a promising therapeutic target for COVID-19, will also be discussed. Overall, MX-BaMS is a highly promising approach for rapid and efficient screening of hundreds of compounds simultaneously for ligands that bind to protein targets. We envision that nanoscale ion emitters should also be useful for improving native proteomics and live-single cell MS measurements.

References

SCIEX
The Power of Precision

1. Nguyen, G. T. H.; Nocentini, A.; Angeli, A.; Gratteri, P.; Supuran, C. T.; Donald, W. A. *Analytical Chemistry* **2020,** *92* (6), 4614-4622.
2. Nguyen, G. T. H.; Tran, T. N.; Podgorski, M. N.; Bell, S. G.; Supuran, C. T.; Donald, W. A. *ACS Central Science* **2019,** *5* (2), 308-318.
3. Nguyen, G. T. H.; Leung, W. Y.; Tran, T. N.; Wang, H.; Murray, V.; Donald, W. A. *Analytical Chemistry* **2020,** *92* (1), 1130-1137.

40

Blink of an Eye

Sarah Alexander[1], Tom Covey[1]

1. SCIEX, Mount Waverly, VIC, Australia

Echo MS is a mass spectrometry (MS) analytical platform resulting from the novel integration of acoustic droplet ejection (ADE) technology, an open-port interface (OPI), and electrospray ionization (ESI) MS that creates a transformative system enabling high-speed sampling and label-free analysis. The technology delivers nanoliter droplets in a touchless manner with high speed, precision and accuracy; subsequent sample dilution within the OPI, in concert with the capabilities of modern ESI-MS, eliminates the laborious sample preparation and method development required in current approaches. This novel new platform has already been applied to a variety of experiments, including high-throughput (HT) pharmacology screening, label-free in situ enzyme kinetics, in vitro and in vivo adsorption, distribution, metabolism, elimination, pharmacokinetic (PK) and biomarker analysis, and HT parallel medicinal chemistry. I will present some of the latest applications and developments for this game changing Technology.

Why a Blink of an Eye?

A Blink of an eye is about 400-500 milliseconds depending on your source of information. A baker's dozen is 13 and this system can analyse 6-12 samples in the time it takes for one or two eye blinks!!

41

High Field Asymmetric Waveform Ion Mobility (FAIMS) increases the depth and coverage of cross-linked peptide identification

Ashleigh L Dale[1, 2], Dylan J Harney[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

High Field Asymmetric Waveform Ion Mobility (FAIMS) is an ion fractionation device attached front-end to an Orbitrap tribrid mass spectrometer that acts to separate and fractionate ions in the gas phase according to their physicochemical properties including charge, size, mass, and dipole moment. The use of FAIMS for cross-linking mass spectrometry (XL-MS), compared to traditional off-line fractionation techniques such as size-exclusion chromatography (SEC), has gained significant traction in recent years given its ability to enrich for low abundance cross-linked species. An advantage of FAIMS is the freedom it gives the user to modify the compensation voltage or combination of voltages applied to peptide ions allowing them to leave the device and enter the mass spectrometer. Therefore, the same sample can yield significantly different results based on the voltage scheme applied. Replicate LC-MS/MS analyses of *Campylobacter jejuni* whole cell protein digests and membrane proteins, fractionated with SEC and/or FAIMS and a combination of gradient lengths, were compared using a tribrid fragmentation scheme (CID-MS2-HCD-MS3-EThcD-MS2) and a combined compensation voltage approach with internal-stepping (CV= -50, -60, -75). For complex whole cell tryptic digests, the coupling of FAIMS to LC-MS/MS enabled a 3.2-fold increase in cross-linked spectral matches (CSMs) compared to without FAIMS, and a 2.3-fold increase in identifications compared to SEC fractionation. Single 180-minute gradient injections with FAIMS yielded greater than two-fold more CSMs and cross-linked peptides than any single SEC fraction or injection without differential ion mobility. The combination of FAIMS with SEC resulted in minimal differences in cross-link identification, highlighting a loss of orthogonality. Therefore, the usefulness of FAIMS lies in decreasing the sample amount, MS-time and sample preparation time needed for successful, rapid and in-depth analysis of complex cross-linked samples.

42

Targeted and exploratory studies of protein-protein interactions

Xuanzhao Jiang[1], Daniela-Lee Smith[2], Chi Nam Ignatius Pang[3], Marc R Wilkins[3], David J Tremethick[1], Gene Hart-Smith[4]

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

2. ProCan Laboratory, Children's Medical Research Institute, Sydney, NSW, Australia

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

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

Interactomics can be defined as the large-scale study of biomolecular relationships. It is often simply characterised as the concurrent investigation of many protein-protein interactions; one of life's most important biomolecular relationships. Not all approaches for studying protein-protein interactions are equal, and this presentation will tease apart some of the benefits and drawbacks of these different approaches when assessed within the broader aims of interactomics.

Some approaches for studying protein-protein interactions are targeted towards specific interactions of interest. These approaches are capable of providing deep mechanistic insights. Our recent investigation of H2A.B, a histone variant which plays a role in oncogenesis in

SCIEX
The Power of Precision

Hodgkin lymphoma, provides a case in point. After uncovering a series of novel post-translational modifications on the N-terminus of this protein, we employed biotinylated peptide pull-downs in conjunction with SILAC to identify protein-protein interactions mediated by these modifications. This uncovered a remarkable new function for H2A.B: the genomic targeting of the chromatin remodelling complex SWI/SNF is regulated by asymmetric dimethylation of H2A.BR8.

Targeted approaches for characterising protein-protein interactions do, however, feature limitations. It has been shown that when interactions uncovered in such studies are collated on an interactome scale, systemic biases towards interactions involving highly studied and/or highly expressed proteins become apparent [1]. Our findings involving H2A.B illustrate how such biases can emerge.

To avoid such biases, approaches for studying protein-protein interactions that are less targeted and more exploratory can be employed. These approaches include proteome scale affinity purification followed by mass spectrometry (AP-MS), yeast two-hybrid (Y2H) systems, co-fractionation mass spectrometry (CF-MS), and cross-linking mass spectrometry (XL-MS) of complex samples. The bird's eye views of the interactome provided by these approaches do, however, come with a caveat: there is strikingly little overlap between the interactions identified across these different approaches [2,3].

To gain insight into this phenomenon, we assessed the congruence of protein-protein interactions identified from concurrent XL-MS and CF-MS analyses of *Saccharomyces cerevisiae* protein complexes. This helped test the assay complementarity hypothesis, which suggests that different interactome mapping techniques characterise different subsets of protein-protein interactions, thus explaining the small overlaps between interactome maps [2,4,5]. Our experiments reveal that protein-protein interactions identified from identical samples using XL-MS and CF-MS do indeed produce strikingly little overlap, supporting the idea of assay complementarity.

Together these examples speak to the relative merits of targeted versus exploratory studies of protein-protein interactions, and identify possible avenues towards fulfilling the promise of systems-level interactomics.

1. 1. Rolland, T. et al. (2014) A proteome-scale map of the human interactome network. Cell 159, 1212–1226.
2. 2. Luck, K. et al. (2017) Proteome-scale human interactomics. Trends in Biochemical Sciences 42, 342–354.
3. 3. Bartolec, T.K. et al. (2019) Cross-linking mass spectrometry analysis of the yeast nucleus reveals extensive protein−protein interactions not detected by systematic two-hybrid or affinity purification-mass spectrometry. Analytical Chemistry 92, 1874–1882.
4. 4. Braun, P. et al. (2009) An experimentally derived confidence score for binary protein–protein interactions. Nature Methods 6, 91–97.
5. 5. Chen, Y.C. et al. (2010) Exhaustive benchmarking of the yeast two-hybrid system. Nature Methods 7, 667–668.

43

Investigating the role of histidine methylation on ribosome assembly using quantitative large-scale cross-linking mass spectrometry

Tara K Bartolec[1], Joshua J Hamey[1], Andrew Keller[2], Juan D Chavez[2], James E Bruce[2], Marc R Wilkins[1]

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

2. Department of Genome Sciences, University of Washington, Seattle, Washington, USA

Cross-linking mass spectrometry (XL-MS) involves the use of chemicals which covalently link amino acids within spatially constrained reactions. Recent technological advances have enabled XL-MS of complex samples and our group generated the first large-scale XL-MS dataset for yeast [1]. This revealed novel insights into the protein interactome and unveiled the conformations of hundreds of unresolved protein structures. Studies of protein translation have revealed mounting evidence that ribosomes are both highly regulated and heterogeneous in structure, composition, interactions and post-translational modifications, making them an ideal system to explore using XL-MS. To investigate the function of one form of heterogeneity, protein methylation, we applied quantitative large-scale XL-MS to study methyl-histidine-243 of ribosomal protein uL3, which is deposited by the Hpm1 methyltransferase.

Wild-type and Hpm1 knockout yeast were cultured using forward/reverse SILAC. Intact spheroplasts were crosslinked with the biotin-containing and MS-cleavable crosslinker PIR. Peptides were fractionated by SCX, enriched with avidin and analysed with stepped-HCD. Crosslinks were identified using the Mango/Comet/XlinkProphet pipeline, and crosslink abundance ratios extracted using MethylQuant/MassChroQ. This resulted in 8481 crosslink spectral matches at a 1% FDR, which represent 1268 unique lysine-lysine pairs. Over half of these were successfully quantified, with a significant proportion quantified in both biological replicates. Many of the detected changes were within or between ribosomal proteins/members of co-translational processes, indicating an altered ribosome structure/interactome. Interesting, we also detected large-scale yet reproducible changes in seemingly unrelated domains including the mitochondria and histones which may reveal unexplored aspects of Hpm1 biology.

This study is one of the few to date to perform differential interactomics analysis using *in vivo* quantitative large-scale XL-MS, and the first in yeast. It illustrates how this technique can be used to profile the diverse impacts of cellular perturbation, even as small as modulating a single methylation site.

1. Cross-linking Mass Spectrometry Analysis of the Yeast Nucleus Reveals Extensive Protein–Protein Interactions Not Detected by Systematic Two-Hybrid or Affinity Purification-Mass Spectrometry. Tara K. Bartolec, Daniela-Lee Smith, Chi Nam Ignatius Pang, You Dan Xu, Joshua J. Hamey, and Marc R. Wilkins. Analytical Chemistry 2020 92 (2), 1874-1882. DOI: 10.1021/acs.analchem.9b03975

SCIEX
The Power of Precision

44

The Benzylpenicillin haptenome and its impact on the HLA-A*02:01 immunopeptidome

Shawn Goh[1], **Katherine Scull**[1], **Kirti Pandey**[1], **Robert Puy**[2], **Robyn O'Hehir**[2], **Anthony Purcell**[1], **Nicole Mifsud**[1], **Patricia Illing**[1]

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

2. Department of Allergy, Immunology and Respiratory Medicine, Monash University, Central Clinical School, The Alfred Hospital, Melbourne, VIC, Australia

Penicillins are a class of β-lactam antibiotics used as a first-line treatment against a broad array of bacteria. However, following administration around 10% of individuals experience immune-mediated hypersensitivity reactions, which restricts treatment options in these individuals. Clinical investigations suggest that penicillin hypersensitivity is associated with HLA-A*02:01, a common HLA allotype found in 40% of the Caucasian population that is involved in T cell-mediated immune responses. We have previously demonstrated the presence of Benzylpenicillin-specific CD8+ T cells in a HLA-A*02:01+ hypersensitive patient and hypothesise that the presentation of drug-modified peptides by HLA-A*02:01 is driving the immunological drug response.

Study Aims:

1) Identify intracellular targets of Benzylpenicillin and predict if peptides encompassing these haptenation sites can furnish ligands for HLA allotypes associated with increased risk of penicillin-hypersensitivity

2) Identify naturally presented HLA ligands with the Benzylpenicillin adduct

To map the Benzylpenicillin haptenome of treated antigen presenting cells (APCs), we employed an MS-based bottom-up approach. Benzylpenicillin-modified proteins and sites were identified based on a mass addition at Lys or Arg residues, and further validated by the presence of diagnostic Benzylpenicillin ions in spectra used for peptide assignment. Utilising a predictive tool (NetMHCpan, v4.1b), we have shown that peptides encompassing haptenation sites are predicted ligands of HLA allotypes associated with increased risk of penicillin-hypersensitivity. This suggests that haptenated intracellular proteins undergo antigen processing to be presented by HLA presentation for T cell recognition. Furthermore, we have isolated drug modified peptides directly from immunoaffinity purified HLA-A*02:01 of drug treated cells.

This study provides essential knowledge of drug modified antigens generated by Benzylpenicillin and their capacity to enter the antigen presentation pathway for display at the cell surface by HLA molecules. Together, the data will help inform rational intervention strategies based on the immunogenicity of identified drug-modified peptides and their interactions with T cells.

45

OzWheat – A genetic diversity panel for classification and prediction of wheat traits

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

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

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

Proteome measurements from genetic diversity panels offer the opportunity to robustly classify and predict traits in complex populations. At the same time, the analysis of the proteomes arising from these populations present specific challenges. One challenge lies in establishing a pan-proteome database and/or search strategy such that sequence variation can be robustly identified across genetic diversity. Another challenge lies in establishing a unified set of protein coordinates with which to map all samples' measurements such that quantitative data are directly comparable between samples. While inroads have been made that seek to meet these challenges, a complete solution has not yet become mainstream. In this regard, approaches to this problem that uncouple the requirement for peptide sequence assignment / protein attribution from classification and prediction offer an alternative solution.

Wheat varieties from the OzWheat genetic diversity panel were grown in glasshouses under simulated long and short daylight conditions using a randomised planting scheme. Samples were taken at the two-leaf stage while sibling plants were grown to full term to capture trait information, such as grain yield, flowering time and spike length. Proteomes were measured for a subset of the cultivars using SWATH acquisition and peak area information extracted using PeakView software. The resulting data were used to classify samples by day length and predict flowering time using a selection of machine learning methods. Recent work has focussed on the conversion of raw SWATH data to images with the goal of applying image classification methods, such as convolution neural networks, for determining day length and predicting yield traits without prior knowledge of peptide/protein coordinates. This image classification work is ongoing at the time of abstract submission with progress to be presented at the conference.

46

Interactive statistical and functional analysis of phosphoproteomics data with *Phosphomatics*

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

1. Bio21 Mass Spectrometry and Proteomics Facility, University of Melbourne, Parkville, VIC, Australia

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 stimuli. While these experiments can routinely detect and quantify thousands of

SCIEX
The Power of Precision

phosphorylated peptides, interpreting this data, and extracting biological meaning, remains challenging. In part, this is because comprehensive analysis can frequently involve many different software packages, web sites and databases resulting in a complicated and iterative process that is both prohibitive for non-experts and cumbersome and time-consuming for experienced researchers. Here, we present a substantial expansion to our phosphoproteomics data analysis website - *'Phosphomatics'* – that incorporates a suite of new tools and resources for statistical and functional analysis that aim to simplify the process of extracting meaningful insights from experimental results.

Phosphomatics can natively import search and quantitation results from major search engines including MaxQuant and Proteome Discoverer and employs intuitive 'wizards' to guide user through data preprocessing routines such as filtering, normalization and transformation. A graphical platform of interactive univariate and multivariate analysis features is provided that allow subgroups of the uploaded data containing phosphosites of statistical interest to be created and interrogated through further functional analysis. For example, features are incorporated to assess hyperactivation of putative upstream kinases, consensus phosphorylation motifs for peptides groups, pathway/gene ontology enrichment and interaction networks. A range of databases have been integrated that, for example, provide ligand and inhibitor information for key proteins or highlight key modification sites known to be involved in functional state regulation. At each step, published literature is natively incorporated along with a 'bibliography builder' that allows references of interest to be assembled and exported in various formats. Taken together, these expanded features aim to provide a 'one-stop-shop' for phosphoproteomics data analysis.

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

47

CoffeeProt: An online tool for correlation and functional enrichment of proteome-wide systems genetics

Jeffrey Molendijk[1], **Marcus M. Seldin**[2], **Benjamin L. Parker**[1]

1. Department of Physiology, University of Melbourne, Melbourne, Victoria, Australia

2. Department of Biological Chemistry and Center for Epigenetics and Metabolism, University of California, Irvine, California, USA

The integration of genomics, proteomics and phenotypic traits across genetically diverse populations is a powerful approach to discover novel biological regulators. The increasing volume of complex data require new and easy-to-use tools accessible to a variety of scientists for the discovery and visualization of functionally relevant associations. To meet this requirement, we developed *CoffeeProt*, an open-source tool that analyzes genetic variants associated to protein networks and phenotypic traits. *CoffeeProt* uses proteomics data to perform correlation network analysis and annotates protein-protein interactions, subcellular localizations and drug associations. It then integrates genetic and phenotypic associations along with variant effect predictions. We demonstrate its utility with a step-by-step tutorial of the workflow in the analysis of mouse data, highlighting the rapid identification of genetic variants associated with druggable proteins and clinical traits. We expect that *CoffeeProt* will serve the proteomics and systems genetics communities, leading to the discovery of novel biologically relevant associations. *CoffeeProt* is available at www.coffeeprot.com.

48

SRL generation using neural network analysis in lung carcinoma SWATH-MS acquisition

Daniel Bucio Noble[1], **Erin Humphries**[1], **Jennifer Koh**[1], **Steve Williams**[1], **Daniela Smith**[1], **Erin Sykes**[1], **Clare Loudon**[1], **Dylan Xavier**[1], **Natasha Lucas**[1], **Peter Hains**[1], **Phil Robinson**[1]

1. Children's Medical Research Institution, Westmead, NSW, Australia

Data Independent Acquisition (DIA) strategies have been widely adopted in Proteomics studies involving clinical samples in large cohorts. Given its robustness and reproducibility, SWATH-MS data can be routinely analysed using spectral reference libraries (SRLs) derived from Information Dependent Acquisition (IDA). The emergence of novel tools for SRL generation employing neural network analysis, such as DIA-NN, may offer alternatives to conventional methods. Despite the importance of SRLs in peptide identification and quantification, there is a lack of information in regards to SRL comparison derived from IDA and DIA-NN in clinical settings. Here we present such comparison analysis using formalin-fixed parafilm embedded (FFPE) and fresh frozen (FF) preserved Adenocarcinoma and Squamous Stage 1 Lung cancer tissue. Our data shows that a neural network trained library built on high pH fractions acquired in SWATH mode improves the number of identified proteins in 8% in contrast to a conventional high pH-fractionated SRL acquired with IDA mode. In addition, our DIA-NN SWATH library improves the number of quantified proteins in around 22% compared to libraries built on different methods. DIA-NN allows us to build subset specific SRLs using exclusively SWATH files. For instance, FF-Adenocarcinoma and Squamous libraries differentiate malignant from their corresponding normal matching tissue as shown in principal component analysis. We also determined that an Adenocarcinoma SRL made from FFPE and FF samples, containing 4779 proteins, and an SRL composed of a mix of Adenocarcinoma, Squamous, Lung Neuroendocrine Carcinoma (LNEC) and Large-cell lung carcinoma (LCC)-FFPE samples, containing 5447 proteins, when both applied to the same Adenocarcinoma-FFPE cohort, effectively segregate malignant from normal tissue. The employment of SWATH data for library generation is of relevance in studies where protein yield is limited for IDA acquisition such is often the case for clinical samples. These also suggest the suitability of DIA-NN in producing comprehensive and complementary SWATH-derived subset specific libraries regardless of their preservation method. Our results show DIA-NN as a valuable tool in clinical settings for SRL production.

SCIEX
The Power of Precision

49

SARS-CoV-2 structural coverage map reveals state changes that disrupt host immunity

Sean I. O'Donoghue[1], Andrea Schafferhans[2], Neblina Sikta[1], Christian Stolte[1], Sandeep Kaur[3], Bosco K. Ho[1], Stuart Anderson[4], James Procter[5], Christian Dallago[6], Nicola Bordin[7], Matt Adcock[4], Burkhard Rost[6]

1. Garvan Institute of Medical Research, Darlinghurst, NSW, Australia

2. Weihenstephan-Tr. University of Applied Sciences, Freising, Germany

3. University of New South Wales, Kensington, NSW, Australia

4. Data61, CSIRO, Canberra, ACT, Australia

5. University of Dundee, Dundee, United Kingdom

6. Technical University of Munich, Munich, Germany

7. University College London, London, United Kingdom

In response to the COVID-19 pandemic, many life scientists are focused on SARS-CoV-2. To help them use available structural data, we systematically modeled all viral proteins using all related 3D structures, generating 872 models that provide detail not available elsewhere. To organise these models, we created a structural coverage map: a novel, one-stop visualization summarizing what is — and is not — known about the 3D structure of the viral proteome. The map highlights structural evidence for viral protein interactions, mimicry, and hijacking; it also helps researchers find 3D models of interest, which can then be mapped with UniProt, PredictProtein, or CATH features. The resulting Aquaria-COVID resource (https://aquaria.ws/covid) helps scientists understand molecular mechanisms underlying coronavirus infection. Based on insights gained using our resource, we propose mechanisms by which the virus may enter immune cells, sense the cell type, then switch focus from viral reproduction to disrupting host immune responses.

1. https://doi.org/10.1101/2020.07.16.207308

50

Immunopeptidomics for the identification of SARS-CoV-2 T cell epitopes.

Patricia Illing[1], Sri Ramarathinam[1], Sinead Williams[2], Leonard Izzard[3], Pouya Faridi[1], Rochelle Ayala[1], Asolina Braun[1], Shan Zou Chung[1], Nathan Croft[1], Ziyi Huang[1], Chen Li[1], Nicole Mifsud[1], Kate Scull[1], Michelle Baker[2], Anthony Purcell[1]

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

2. CSIRO, Australian Centre for Disease Preparedness, Health and Biosecurity Business Unit, Geelong, Victoria, Australia

3. CSIRO, Australian Centre for Disease Preparedness, Australian Animal Health Laboratory Business Unit, Geelong, Victoria, Australia

The SARS-CoV-2 pandemic has had unprecedented global impact. To date, there have been over 67 million confirmed infections and 1.5 million deaths worldwide. Diverse outcomes to infection are observed, with asymptomatic presentation and mild illness at one end of the spectrum, and pneumonia-like disease requiring intensive care at the other. Despite recent promising developments in vaccines, our understanding of the immune response to this novel pathogen, including why responses vary so greatly, remains immature.

The cellular immune response relies on cell surface presentation of pathogen-derived peptidic antigens by the Human Leukocyte Antigen (HLA) class I and II molecules (HLA-I and HLA-II) for surveillance by $CD8^+$ and $CD4^+$ T cells, respectively. We have employed an immunopeptidomics workflow to identify naturally processed and presented peptides derived from SARS-CoV-2 in infected airway epithelial cell lines. In brief, post-infection, cells were lysed, and the HLA-I and HLA-II immunoaffinity purified using a pan HLA-I (W6/32) monoclonal antibody (mAb) and a pan HLA-II mAb cocktail (combines anti HLA-DR, anti HLA-DQ and anti HLA-DP). Peptides were acid eluted and analysed by LC-MS/MS.

Our analyses have revealed HLA-I mediated presentation of Spike glycoprotein derived peptides incorporating deamidated asparagine (N) residues, a residual post-translational modification indicative of removal of prior N-linked glycosylation. In contrast, the majority of HLA-II restricted peptides mapped to the nucleoprotein, ORF3a, membrane protein and ORF9b, revealing differences in sampling of the SARS-CoV-2 proteome by the HLA-I and -II antigen processing pathways. Understanding which regions of the virus are responsible for eliciting protective immune responses will have applications in future vaccine design, work to define the correlation between specific T cell responses and infection outcome, and our ultimate understanding of this newly emerged pathogen.

51

Exploring the origins, evolution and function of methylation in disease-causing protists

Samantha J Emery-Corbin[1, 2], Joshua J Hamey[3], Brendan R.E Ansell[1, 2], Balu Balan[4, 1, 2], Swapnil Tichkule[1, 2], Andreas J Stroehlein[4], Crystal Cooper[5], Bernie V McInerney[6], Soroor Hediyeh-zadeh[1, 2], Daniel Vuong[7], Andrew Crombie[7], Ernest Lacey[7], Melissa J Davis[1, 2], Marc R Wilkins[3], Melanie Bahlo[1, 2], Staffan G Svärd[8], Robin B Gasser[4], Aaron Jex[4, 1, 2]

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

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

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

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

5. Central Analytical Research Facility (CARF), Queensland University of Technology, Brisbane, QLD, Australia

SCIEX
The Power of Precision

6. Australian Proteome Analysis Facility, Macquarie University, North Ryde, NSW, Australia

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

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

Methylation of arginine (Arg/R/R-Me) and lysine (Lys/K/K-Me) are essential protein post-translational modifications (PTM) of eukaryotic proteins. While detailed studies of eukaryotic methylproteomes exist for mammals (e.g., rodents, humans) and experimentally tractable model organisms (e.g., yeast), we lack species from basal ("ancient") eukaryotic taxa. Without these species we cannot contexualise the numerous evolutionary expansions in methylproteomes of higher eukaryotes, or distinguish evolved ("new") interactions and components from original methylation networks. The protists *Trichomonas vaginalis*, *Entamoeba histolytica* and *Giardia duodenalis* are among the earliest extant representatives of Eukaryota, and are major human and veterinary parasites. These protists are valuable as evolutionary models from basal eukaryotic taxa, and in representing unique phenotypic diversity (i.e. disease-causing species).

We have demonstrated functional methylproteome networks in protist lineages which pre-date current eukaryotic model species by ~1 billion years. We used sequence-based domain-mapping and 3D-structural modelling to bioinformatically curate Class I and Class V methyltransferase (MTases) enzymes in seven protist species. This identified putative lysine-MTases, including novel annotations of MTases of eukaryote Elongation Factor 1 alpha (eEF1a) in the earliest diverging extant lineages of the Eukaryota. It also demonstrated that diplomonad species including *Giardia* lacked methylarginine enzymes, and had converted R-Me sites and RGG motifs in substrate orthologs. To explore these losses, we performed *in vitro* analyses of methylation in *Giardia* using immunoblotting, Amino Acid Analysis, and immunoaffinity purification of methylation-modified peptides combined with LC-MS/MS. While we readily identified methyllysine in *Giardia*, no methylarginine was detected. This confirmed *Giardia* as the first eukaryote lacking conserved methylarginine networks. In contrast, we bioinformatically inferred the enzymes and preferred substrates of methylarginine are present in other protists including *Entamoeba* and *Trichomonas*, highlighting intriguing contrasts in methyl-regulation between basal species. In order to facilitate large-scale methyl-site identifications in these protists, we have validated methyl-site confidence-filtering pipelines in R software as an alternative to incompatible orthogonal methyl-peptide identification approaches (i.e., heavy-methyl SILAC). To date, we have identified >200 methyllysine sites in *Giardia* using these pipelines, including sites in the histone H3 variant and eEF1a which demonstrate basal origins of these eukaryotic regulatory mechanisms. However, the majority of *Giardia* methyllysine proteins identified were species- or lineage-specific as compared to model eukaryotes, including gene families enriched in coiled-coil features involved in cytoskeletal regulation. Together, these protists' methylproteomes represent both specialized adaptations for parasitic lifestyles and eukaryote-conserved mechanisms, and provide a more complete understanding of the natural history of methylation through the Eukaryota.

52

The Human Proteome Project High-Stringecy Blueprint

Mark S. Baker[1], Subash Adhikari[1]

1. Macquarie University, NSW, Australia

HUPO launched the Human Proteome Project (HPP) in 2010, creating an international framework for global collaboration, reanalysis of community shared data, quality assurance and accurate annotation of the genome-encoded proteome. During the past decade, the HPP established collaborations, developed guidelines/metrics, and reanalyzed previously-deposited community uploaded data, resulting in continuous increases in coverage of the human proteome. Celebrating its recent 10th anniversary, the HPP reported a 90.4% complete high-stringency human proteome blueprint. This presentation addresses strategies for making more of the currently invisible human proteome credibly visible. The HPP blueprint highlights how human proteome knowledge is essential for discerning molecular processes, and with other "omics" data playing a key role in our understanding, diagnoses and treatment of human disease.

60

Understanding how azithromycin reduces asthma exacerbations and the underlying mechanisms of macrolides

Lisa M Jurak[1], John W Upham[1, 2], Michelle M Hill[1, 3], Ian A Yang[4, 5], Jodie L Simpson[6]

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

2. Princess Alexandra Hospital, Woolloongabba, QLD, Australia

3. 4QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

4. 5UQ Thoracic Research Centre, The University of Queensland, Brisbane, QLD, Australia

5. The Prince Charles Hospital, Brisbane, QLD, Australia

6. University of Newcastle, Hunter Medical Research Institute, Newcastle, NSW, Australia

Abstract:

Introduction/Aim:

New interventions are needed for non-eosinophilic or non-T2 asthma phenotypes. Although the AMAZES study (Lancet 2017) showed that azithromycin (AZM) reduces asthma exacerbations, there is limited understanding of involved mechanisms.

This study's aim was to identify differentially expressed proteins in sputum from AMAZES participants, comparing sputum obtained before and after 48 weeks of AZM or placebo treatment in order to generate an AZM-sensitive protein signature.

Methods:

SCIEX
The Power of Precision

Using a robust high throughput method (mass spectroscopy) for the global analysis of the sputum proteome, 52 samples were selected with clinical characteristics balanced to ensure these were representative of the entire AMAZES cohort. Protein intensities were extracted for external statistical analyses using *R studio*. Differentially enriched proteins were identified by a log-fold change of 1.5 and a p-value of 0.05.

Results:

1446 identified proteins were common across all samples. A further 240 proteins were uniquely expressed at week 48 in AZM-treated patients and 214 in placebo-treated patients. Using a univariate model, 32 proteins were differentially expressed in AZM-treated patients, comparing week 48 to baseline. In addition, 90 proteins were differentially expressed at week 48 between AZM and placebo-treated samples. Finally, placebo-treated patients showed no significant changes in protein expression between baseline and week 48.

Multivariate modelling identified a unique signature of 60 proteins that distinguish AZM from placebo in comparison to the baseline proteome. Finally, protein interactions and pathway analysis identified overrepresented pathways including apoptosis, phagocytosis, IL-5 pathway, endogenous TLR/chemokine signalling and bacterial invasion, independent of asthma phenotype.

Conclusion:

We have identified known and unique proteins that change with add-on AZM therapy. Analysis of the sputum proteome provides a unique insight into AZM's mechanisms of action. This will facilitate development of novel treatment options for severe persistent asthma.

61

SWATH Quantification of blister fluid proteome from paediatric burn injuries using an integrated ion library

Tuo Zang[1], Daniel Broszczak[1], Tony Parker[1], Leila Cuttle[1]

1. School of biomedical sciences, Queensland University of Technology, Brisbane, QLD, Australia

Burn injury is a highly traumatic event for children and their families. The degree of burn severity (superficial-, deep-, or full-thickness injury) dictates the clinical management of the injury and the extent of scar formation. It can often take several days for the true depth of a burn injury to become apparent. Additionally, the biological processes by which burn injuries continue to develop and worsen over days or weeks (burn wound conversion) are not well understood. Burn blister fluid (BF), which develops after injury, contains proteins that reflect both the systemic and local microenvironment response to the injury and have the potential to diagnose burn wound severity. BF is collected non-invasively and therefore, it is ideal for studying the burn wound proteome in paediatric patients. BF samples were collected from 65 children in Brisbane (Queensland, Australia) and quantitatively analysed using SWATH mass spectrometry. An integrated peptide ion library, consisting of data-dependent analysis (DDA) data from 1) pooled BF samples, 2) previous DDA data from fractionated BF samples, and 3) published human DDA data, was used to extract abundance values for more than 2000 proteins from the SWATH data derived from each individual sample. An optimized normalization method was used prior to data processing. Using the integrated ion library, we found significant differences between burn severities, based on relative protein abundance. In addition, the biological processes relating to different burn depths were profiled through gene ontology enrichment analysis. Through this study, the quantitative profiling of burn blister fluid enabled identification of proteins related to burn severity and extent of injury. These will be further investigated as potential biomarkers to assist with clinical diagnosis of bun wound severity.

62

Investigation of recombinant human factor IX posttranslational modification differences in fed-batch and perfusion processes of Chinese Hamster Ovary cells against plasma-derived factor IX

Dinora D Roche Recinos[2, 3, 1], Cassandra C Pegg[4], Toan K. TK Phung[4], Ellen E Otte[3], Mark M Napoli[3], Campbell C Aitken[3], Yih Yean YY Lee[3], Ben B Schulz[4], Christopher B. C Howard[2, 1]

1. Australian Institute of Bioengineering and Nanotechnology (AIBN) at The University of Queensland (UQ), Brisbane, QLD, Australia

2. Centre for Biopharmaceutical Innovation (CBI), Australian Institute of Bioengineering and Nanotechnology (AIBN) at The University of Queensland (UQ), Brisbane, QLD, Australia

3. CSL Limited, Melbourne, Victoria, Australia

4. School of Chemistry and Molecular Biology, University of Queensland, Brisbane, QLD, Australia

Background and novelty

Human coagulation factor IX (FIX) is a protein that relies on an extensive spectrum of posttranslational modifications that enable it to function correctly and efficiently in the coagulation pathway [1, 2]. These consist of seven disulfide bridges, two N-glycans and six O-linked glycans, one sulfation site, one phosphorylation site, 12 γ-carboxylation (GLA) sites as well as one β-hydroxylation sites [3-9]. This study aims to investigate the differences in the posttranslational modifications of human recombinant factor IX (rFIX) produced in CHO fed-batch and perfusion cultures, compared with plasma-derived factor IX (PD-FIX).

Experimental approach

The cell line used was a CHO-K1SV expressing rFIX. Two fed-batch bioreactors were conducted using commercial CD-CHO media and EfficientFeed A and B respectively. Perfusion cultures were conducted in the same base medium using an Applisens Biosep acoustic perfusion unit at a dilution rate of one reactor volume a day. The bioreactors were sampled daily for off-line measurements to track cell growth,

SCIEX
The Power of Precision

metabolism and productivity. These samples were also used for Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH-MS) analysis [10]. In parallel, purified rFIX from these cultures was analyzed after an in-gel digestion using Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC/ESI-MS/MS) to characterize the PTMs.

Results and discussion

The fed-batch cultures responded differently to each of the feeds despite achieving similar peak cell densities of ~15x106 cells/mL. Almost all the PTMs of PD-FIX were observed in rFIX of both fed-batch cultures, although they showed partial occupancy and higher heterogeneity. Preliminary qualitative analysis also suggested that gamma-carboxylation in the rFIX GLA domain is more complete in one fed-batch compared to the other. As a comparison, pseudo steady-states were established in the perfusion cultures at 15x106 cells/mL via bleeding of the cultures under the control of an online turbidity probe. Samples were collected from these steady-states and purified for PTM analysis to establish comparison across the different modes of cultures and the native PD-FIX.

Acknowledgements & Funding

CSL Limited, Melbourne, Australia, supported this research. We would also like to thank our colleagues from both CSL and the Australian Research Council, Centre for Biopharmaceutical Innovation (CBI) who provided insight and expertise that greatly assisted the research.

References

1.Zogg, T. and H. Brandstetter, Activation mechanisms of coagulation factor IX. Biol Chem, 2009. 390(5-6): p. 391-400.

2.Zögg, T. and H. Brandstetter, Structural Basis of the Cofactor- and Substrate-Assisted Activation of Human Coagulation Factor IXa. Structure, 2009. 17(12): p. 1669-1678.

3.Kumar, S.R., L.L.C. Adimab, and N.H. Lebanon, Industrial production of clotting factors: Challenges of expression, and choice of host cells. Biotechnology Journal, 2015. 10: p. 995 - 1004.

4.Monroe, D.M., et al., Characterization of IXINITY(R) (Trenonacog Alfa), a Recombinant Factor IX with Primary Sequence Corresponding to the Threonine-148 Polymorph. Adv Hematol, 2016. 2016: p. 7678901.

5.Kovnir, S.V., et al., A Highly Productive CHO Cell Line Secreting Human Blood Clotting Factor IX. Acta Naturae, 2018. 10(1): p. 51-65.

6.Monahan, P.E., W.H. Velander, and S.P. Bajaj, Coagulation factor IXa, in Handbook of proteolytic enzymes, N.D. Rawlings and G. Salvasen, Editors. 2013, Elseviert, Ltd.

7.Kurachi, K. and E.W. Davie, Isolation and characterization of a cDNA coding for human factor IX. Proc Natl Acad Sci USA, 1982. 79: p. 6461 - 6464.

8.Lee, M.H., et al., Recombinant human factor IX produced from transgenic porcine milk. Biomed Res Int, 2014. 2014: p. 315375.

63

Proteogenomic analysis to identify cancer neo-antigens

Sonali V Mohan[1, 2], Keshava K Datta[2], Corey Smith[3, 4, 5], Harsha Gowda[1, 2, 3]

1. Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

2. Cancer Precision Medicine Group, , QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

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

4. QIMR Centre for Immunotherapy and Vaccine Development, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

5. Translational and Human Immunology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

Advent of immunotherapies has revolutionized cancer treatment. Recent success with immunotherapy is predominantly due to checkpoint inhibitors that block inhibitory signals and enable T cell activation that target cancer cells. Other strategies including adoptive cell transfer and cancer vaccines are being investigated in parallel to increase available arsenal for immune therapy. Cancer genome sequencing studies have identified several genomic alterations including single nucleotide variations, insertions/deletions and structural variations across various cancers. It is known that some proteins encoded by mutated genes are processed and presented on the cell surface. These MHC presented mutant peptides serve as neo-antigens that are recognized by T cells. Identification of such neo-antigens can strengthen cancer immunotherapy efforts and reveal neo-antigens that can be potentially targeted. However, it is unclear what fraction of mutant alleles in cancer genomes are expressed at the protein level and what fraction of these are presented on cell surface by MHC complex. Mass spectrometry based immunopeptidome datasets can provide large-scale datasets that can be used to gain insights into sequence features and other principles that potentially determine peptides that are presented by MHC complex. We have combined whole-exome sequencing and transcriptome analysis with proteomics and MHC peptidome mass spectrometry to identify potential neo-antigens from melanoma, lung and breast cancer cell lines. We identified thousands of MHC bound peptides from cancer cell lines including mutant peptides that can potentially serve as neo-antigens. Our proteogenomics analysis revealed that the proportion of genomic variants that are presented by MHC class I complex is significantly small. These observations can prove useful for developing better experimental strategies and prediction tools to identify potential cancer neo-antigens. Reliable identification of cancer neo-antigens can accelerate development of novel therapeutic approaches that can exploit host immune system to treat cancers.

SCIEX
The Power of Precision

64

Glycoproteomics to determine site-specific autopolysialylation in human ST8 sialyltransferases

Ruby Pelingon[1], Cassandra Pegg[1], Toan Phung[1], Matthew Hardy[2], Lucia Zacchi[1], Christopher Howard[1], Catherine Owczarek[2], Ben Schulz[1]

1. Univeristy of Queensland, Brisbane, QUEENSLAND, Australia

2. CSL Limited, Melbourne, Victoria, Australia

Human ST8 sialyltransferases are enzymes that transfer oligo- or polysialic acid to the non-reducing terminal glycans of proteins and lipids. This polysialic acid (PSA) is a chain of α2-8-linked sialic acid which has recently had increased attention from biomedical researchers because of PSA's role in neurological diseases and tumour metastases. PSA is also extensively researched in biotechnology as a potential natural polymer to conjugate to therapeutic proteins to improve pharmacokinetics. To understand the mechanisms and constraints of polysialylation, we studied the six human ST8 sialyltransferases, and in particular ST8SIA2 and ST8SIA4, which are known to autopolysialylate. We expressed and purified these enzymes from HEK293 cell culture, and used mass spectrometry glycoproteomics to measure the site-specific glycosylation profiles and polysialylation sites of these autopolysialylated sialyltransferases.

65

Discovery of organelle membrane remodelling associated with CAV1/CAVIN1 in prostate cancer using integrative Protein and Lipid Organelle Profiling (iPLOP)

Harley Robinson[1, 2], Michelle Hill[3, 2]

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

2. Precision and Systems Biomedicine, QIMR Berghofer Medical Research Institute , Brisbane, QLD, Australia

3. UQ Diamantina Institute , Brisbane, QLD, Australia

Introduction:

Mammalian cells are compartmentalised into membrane bound organelles, where dysregulated membrane composition leads to loss of organelle homeostasis and may underlie chronic diseases. The cholesterol-binding protein caveolin-1 (CAV1) normally forms specialised caveolae structures at the plasma membrane with the co-factor protein cavin-1 (CAVIN1). However, CAV1 is expressed in advanced prostate cancer without CAVIN1, exemplified by PC3 cells. Ectopic expression of CAVIN1 in PC3 cells lead to caveola formation at the plasma membrane and attenuated prostate cancer progression *in vivo* (1). In addition to plasma membrane structure changes, PC3-CAVIN1 cells showed redistribution of cholesterol to intracellular compartments (2), therefore, we hypothesize that membrane remodelling modulated by CAVIN1 regulates oncogenic organelle function in prostate cancer.

Methods:

To enable comprehensive evaluation of organelle composition and proteo-lipid interactions, we developed an integrative Protein and Lipid Organelle Profiling (iPLOP) workflow which incorporates density gradient organelle fractionation with lipidomic (Dynamic MRM) and proteomic (DDA MS/MS) profiling. Here, we used iPLOP to discover organelle lipidomic changes associated with PC3-CAVIN1 cells compared to PC3-CONT cells, and then conducted validation using cell-based assays.

Results:

iPLOP workflow yielded two major findings which may underpin the reduced aggressiveness in PC3-CAVIN1 cells: reduced long unsaturated lipids in the mitochondria and endoplasmic reticulum (ER), and elevated lipid droplets. Lipid droplets, mitochondrial and ER membrane composition and function have been associated with cancer hallmarks (3). These results were functionally validated using cell and spectrometry-based assays. Combined with further functional assessment revealed possible links between the membrane remodelling and oncogenic phenotypes.

Conclusions:

iPLOP allowed discovery of CAVIN1-induced organelle remodelling in prostate cancer. This unbiased organelle profiling method can facilitate discovery in other cell systems and disease models associated with membrane remodelling.

1. 1. Moon et al. 2014 Oncogene 33:3561-70.
2. 2. Inder et al. 2012 Mol Cell Proteomics. 11:M111.012245
3. 3. Molendijk et al. 2020 Molecular Omics. 17:6-18.

66

The Role of RBBP7 in Maintaining Genomic Stability

Alejandra Fernandez[1], Janelle Hancock[2], Dorothy Loo Oey[2], Amila Suraweera[1], Joshua Burgess[1], Ken O'Byrne[1], Derek J Richard[1]

1. School of Biomedical Sciences, Queenland University of Technology, Woolloongabba, Queensland, Australia

2. Proteomics, Translational Research Institute, Woolloongabba, Queensland, Australia

Introduction: This project aims to investigate the role of retinoblastoma binding protein 7 (RBBP7) in maintaining genomic stability (a universal cancer hallmark) through an interaction with the components of a previously undescribed chromatin-remodelling complex named as the Histone Ubiquitin Remodelling Complex (HURC). Additionally, the project will determine the role of the HURC complex in DNA double strand break (DNA DSB) repair and investigate the role RBBP7 plays in lung and breast cancer development, progression and resistance.

We hypothesise that RBBP7 functions as part of a novel Histone Ubiquitin Remodelling Complex (HURC) that is critical for genomic stability.

SCIEX
The Power of Precision

Methods: Liquid chromatography mass spectrometry (LC-MS) using a method of Q exactive HF was used to determine which components of the HURC complex interact with RBBP7. Confirmation of protein interactions were performed by co-immunoprecipitations.

Results: LC-MS results have shown that RBBP7 interacts with several proteins that are part of the copper metabolism gene MURR1 (COMMD) family. Co-immunoprecipitations have confirmed that RBBP7 interacts with some of the HURC components.

Conclusion: LC-MS has opened a range of unexplored possible interactions that can be further confirmed through co-immunoprecipitations and that might be involved in genomic stability and other relevant molecular pathways. RBBP7 interacts with some of the HURC components and other DNA repair proteins suggesting that it may play a role in DNA repair.

67

LC-DAD-MS/MS methodology for anthocyanin analysis in pigmented corn kernels

Hung T. Hong[1]

1. The University of Queensland, Coopers Plains, QLD, Australia

Pigmented corn is a rich source of anthocyanins, natural blue, purple, and red pigments that have demonstrated various potential health benefits. Several studies have attempted to extract and quantify the anthocyanins of pigmented corn. However, low stability of anthocyanins following extraction, and a strong tendency for anthocyanins to remain bound to the corn matrix are still major issues to be addressed. In this study, an optimal extraction procedure and an ultra-high performance liquid chromatography–diode array detector-mass spectrometry (UHPLC-DAD-MS) method was developed for characterisation and quantification of anthocyanin components in mature corn kernels. A total of eighteen anthocyanins, mainly cyanidin-, peonidin-, and pelargonidin-based glucosides, were identified and quantified. Cyanidin-based glucosides were the major pigments of purple-pericarp sweetcorn (75.5%) and blue-aleurone maize (91.6%), while pelargonidin-base glucosides constituted the main anthocyanins of reddish-purple-pericarp sweetcorn (61.1%) and cherry-aleurone maize (74.6%). Importantly, it could be clearly demonstrated that previously reported acetylated and succinylated anthocyanins in corn kernels are generated during the extraction process and are not genuine corn pigments. Acidified solutions stabilise and support the release of anthocyanins during the extraction. However, the solutions also hydrolyse and esterify the ester linkages between malonic acid and Cy3G, Pg3G and Pn3G in the malonated anthocyanins and a carboxyl group of malonic acid. Consequently, nine 'artefact' anthocyanins were identified by LC-DAD-ESI-MS. The creation of artefact anthocyanins can be reduced by the use of low temperature during extraction, instrumental analytical operation, and storage of extracts. These findings are crucial for correct anthocyanin profiling of pigmented corns, and emphasizes the importance of using acidification during the extraction process for corn-based anthocyanins, and potentially other anthocyanin-containing commodities.

1. Hong, H. T.; Netzel, M. E.; O'Hare, T. J., Optimisation of extraction procedure and development of LC–DAD–MS methodology for anthocyanin analysis in anthocyanin-pigmented corn kernels. Food Chem 2020.

68

The use of tetracycline-off system to study yeast with deficiency in both Ost3p and Ost6p

Chun Zhou[1], **Benjamin Schulz**[2, 1, 3]

1. School of Chemistry and Molecular Biosciences, 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

3. Australian Infectious Disease Research Centre, The University of Queensland, St Lucia, QLD, Australia

N-glycosylation plays an essential role in protein folding and function in eukaryotic cells. Transfer of glycan to selected asparagine residues in polypeptides is catalysed by oligosaccharyltransferase (OTase), which is a multimeric complex consisting of eight subunits. Ost3p and Ost6p are mutually exclusive subunits in the yeast OTase, defining two OTase isoforms with distinct substrate-specificities. The oxidoreductase activity of Ost3p/6p mediated by the N-terminal thioredoxin domain is important for efficient site-specific glycosylation. Absence of both Ost3p and Ost6p in yeast causes underglycosylation at many glycosylation sites and therefore results in a severe growth defect. To study the functions of Ost3p and Ost6p in and out of glycosylation, a tetracycline-off system was recruited for conditional knock-down of *OST3*, while *OST6* was genomically deleted by replacing the whole gene with a HisMx cassette. This double-deficient strain has a normal growth rate when tetracycline is absent, as the expression of *OST3* is stimulated by the attachment of the tetracycline-controlled transactivator to the tetracycline responsive element in the Tet promoter. Interestingly, we found that the double-deficient yeast had a more severe growth defect than the double-knockout yeast, suggesting that additional mutations had accumulated in the double-knockout yeast. To identify possible suppressor mutations, we isolated fast-growing colonies of the double-deficient yeast and sequenced their genomes. This sequencing data showed that in multiple independent colonies mutations accumulated in the tetracycline-off system rather than in native yeast genes, suggesting that single mutational events that can suppress the important role of *OST3* and *OST6* in yeast are very rare.

SCIEX
The Power of Precision

69

A new, algorithm-based analysis workflow to explore non-traditional cereals for the presence of gluten-like proteins

Sophia Escobar-Correas[2, 1], Angela Juhasz[1], James A Broadbent[2], Michelle L Colgrave[2, 1]

1. School of Science, Edith Cowan University, Perth, WA, Australia

2. Agriculture and Food, CSIRO, Brisbane, QLD, Australia

Gluten proteins are the main storage proteins within cereal grains such as wheat, rye and barley. Consumption of these cereals can lead to serious digestive problems in those with Coeliac disease, an autoimmune disorder, that affects 1-2% of the global population. The only solution for these patients is adherence to a strict gluten-free diet; however, symptoms persist in about 30% of patients despite committing to this strict regimen.

These persistent symptoms may arise from the presence of gluten-containing grains from agricultural co-mingling, that is cross-contamination occurring during cereal harvesting due to the presence of weeds. In this study, the most common weed infesting grain fields in Australia: ryegrass, family *Lolium*, was studied This grass has small dense seeds that are difficult to eliminate during automated grain cleaning. Ryegrass, like wheat and other cultivated crop species is a member of the grass species (Poaceae) wherein the storage proteins comprise gluten-like proteins.

In this study we developed a novel proteogenomic approach to identify 30 gluten-like proteins with potential immunogenic properties from cereal species underrepresented in public protein sequence repositories. Herein, 32 ryegrass genotypes were analysed by ELISA and LC-MS/MS. Data were processed and analysed using in-house developed workflows, in which a novel database was constructed from transcriptomic and genomic data. Amino acid substitution searches were performed supported by a peptide validation algorithm. Primary sequence alignments were used to filter gluten proteins and identify sequence regions with epitope-like features and thereby potential immunogenic properties. The proteins aligned to five protein types with known immune-reactivity: omega-gliadin, gamma-gliadin, high-molecular-weight glutenin, avenin and avenin-like proteins. Protein motif searching uncovered potential immunotoxic peptides similar to wheat, barley and oats. The next stage will be to explore whether these non-traditional gluten sources enter the food supply chain and to confirm their potential to trigger Coeliac disease.

70

Integrating SWATH-MS and Multi-omics Data Analysis to Study Impact of Gene Methylation on Protein Abundance

Manika Singh[1], Annette McGrath [2], Michelle Colgrave[3], Selvam Paramasivan[4], Tony Parker[5], Kevin Dudley[1], Pawel Sadowski[1]

1. Central Analytical Research Facility, Queensland University of Technology, Brisbane, QLD, Australia

2. Data61, CSIRO, Brisbane, QLD, Australia

3. Agriculture and Food, CSIRO, Brisbane, QLD, Australia

4. School of Veterinary Science, University of Queensland, Gatton, QLD, Australia

5. School of Biomedical Science, Faculty of Health, Queensland University of Technology, Kelvin Grove, Brisbane, QLD, Australia

DNA methylation is one of several epigenetic mechanisms that organisms use to control gene expression. Aberrant DNA methylation disrupts normal gene expression and has been associated with various disorders including cancer. The exact role of region-specific DNA methylation, however, is not yet fully understood.

The experimental design entails the analysis of a cerebral brain cortex region of 10-week-old male mice from two strains (C57BL/6J and DBA/2J) in a steady-state model. The study interrogates DNA methylation using whole-genome bisulfite sequencing (WGBS), transcriptome analysis using RNA-seq, and quantitative proteomics using SWATH-MS. R packages and in-house scripts are then used to explore correlations of various omics datasets.

Our preliminary analysis identified 8,907 significant DMRs (differentially methylated regions), 1,426 significant DEGs (differentially expressed genes), and 116 significant DAPs (differentially abundant proteins) between the two mice strains under study. Interestingly, 241 DEGs genes had DMRs, and 8 genes encoding significant DAPs also harbored DMRs. Our observations are in line with known phenotype differences between these two extensively studied mouse strains and are suggestive of a critical role for DNA methylation in this process.

71

A sweet way of understanding and predicting viral zoonoses by comprehensive study of vertebrate host glycome using state-of-the-art glycomics

Abarna Vidya Mohana Murugan[1], Tiago Oliveira[1], Samantha Richardson[2], Yasin Mojtahedinyazdi[3], Kathirvel Alagesan[1], Kimberly Finlayson[3], Jason Van De Merwe[3], Daniel Kolarich[1]

1. Institute for Glycomics, Griffith University, Gold Coast, QLD, Australia

2. RMIT University, Melbourne, Victoria, Australia

3. Institute for Glycomics, Griffith University, Gold Coast Campus, Southport, QLD, Australia

Zoonoses are diseases that are transmissible across difference species. Glycosylation is fundamental for this inter-species transmission and host-pathogen co-evolution. It is well established that pathogen strategies for infection require recognition of host species glyco-epitopes, adhesion and invasion into the host tissue. In consequence, there has been a constant evolutionary pressure controlling the interplay between evolution of pathogen glycan binding proteins such as hemagglutinins and its corresponding host glycosylation and vice versa.

SCIEX
The Power of Precision

However, to date the glyco-epitope niche across Class Vertebrata is still largely undefined thereby limiting our ability to understand and systematically investigate the extant of cross-species transmission potential of zoonotic diseases.

Serum/plasma contains the largest dynamic range of easily accessible, species-specific glycoproteins including important immunity associated circulatory proteins such as Immunoglobulins (Igs). Hence, it provides an archive of species-specific pathophysiological glyco-information. The major Igs existing across the members of the Class vertebrata (such as IgG in mammals, IgY in birds and reptiles, the primordial IgNARs in sharks) is intriguingly diverse signifying host-pathogen co-evolution.

We have laid the foundation for a novel database to systematically curate glycan diversity existing in vertebrate serum/plasma and their corresponding Igs. We have enriched the major Immunoglobulin component of all our vertebrate candidates. We have been systematically mapping the glycome in vertebrate serum/plasma and their corresponding Igs by applying an orthogonal set of different N- and O-glycomics approaches (Porous Graphitised Carbon nano-Liquid Chromatography coupled to ESI MS/MS [PGC nano-LC ESI MS/MS], MALDI TOF MS) including separate arrays for capturing sialic acid diversity (HPLC of DMB-labelled sialic acids). We have hitherto accomplished mapping the serum/plasma N- and O- glycome of 35 different vertebrate species spanning 21 mammals, 6 birds, 5 reptiles and 3 fish with at least one corresponding individuals per species. With a well-established PGC based chromatographic system in our lab, we could differentiate the intensity variation of sialic acid linkage based on the retention time of glycans. The same was verified on the MALDI TOF MS by derivatising the glycans by ethyl esterification which gave a distinct m/z peak for α2-6 and α2-3 linked sialic acids.

Higher incidence of α-Gal epitopes was witnessed in Green sea turtles and predominantly Neu5Ac associated mono, disialylated glycans in birds. The same glycomic pattern was observed in Chicken IgY enriched from the corresponding serum. The results are organised as a species-based heat map and the profound difference and similarities in vertebrate serum/plasma glycan expression is registered.

72

Identification and characterisation of sPEPs in *Cryptococcus neoformans*

Paige Erpf[1, 2], James Fraser[1, 2], Ben Schulz[2], Joseph Rothnagel[2], Edward Kerr[2], Sheena Chua[1, 2], Toan Phung[2]

1. Australian Infectious Disease Research Centre, St Lucia, Queensland, Australia

2. SCMB, School of Chemsitry and Molecular Biosciences, St Lucia, Queensland, Australia

Short open reading frame (sORF)-encoded peptides (sPEPs) have been found across a wide range of genomic locations in a variety of species. To date, their identification, validation, and characterisation in the human fungal pathogen *Cryptococcus neoformans* has been limited due to a lack of standardised protocols. We have developed an enrichment process that enables sPEP detection within a protein sample from this polysaccharide-encapsulated yeast, and implemented proteogenomics to provide insights into the validity of predicted and hypothetical sORFs annotated in the *C. neoformans* genome. Novel sORFs were discovered within the 5' and 3' UTRs of known transcripts and non-coding RNAs. One novel candidate, dubbed *NPB1*, that resided in an RNA annotated as "non-coding" was chosen for characterisation. Through the creation of both specific point mutations and a full deletion allele with the aid of the newly developed *amdS2* Blaster for *C. neoformans*, the function of the new sPEP, Npb1, was shown to resemble that of the bacterial *trans*-translation protein SmpB.

73

Rapidly Advance Quantitative Proteomics with a High-Throughput SWATH® Acquisition Solution

Christie Hunter[1], Amy Heffernan[1]

1. SCIEX, Toronto, ON, Canada

The combination of microflow LC with SWATH® Acquisition for large scale quantitative proteomics studies is becoming increasingly more widespread, due to the improved robustness and throughput obtained relative to the traditional nanoflow LC approach. High quality quantitative datasets have been generated using a standard 1 hour gradient, demonstrating large numbers of proteins quantified routinely.1

Here, an exploration into the impact of gradient length on proteins identified quantified using data dependent (DDA) and data independent acquisition (DIA) was undertaken to provide researchers expanded workflow options with microflow SWATH acquisition. Using microflow liquid chromatography gradients as short as five minutes (total run time <15mins), acquisition parameters were first optimized then used on multiple TripleTOF® 6600 systems to study impact of shortened separations on number of proteins identified and quantified.

Next the fast methods were applied to a small set of cell lines to further characterize the impact of fast gradients on quantitative accuracy.

MATERIALS AND METHODS

Sample Preparation: Cell lysates (HEK, K562 and yeast) were digested with trypsin using standard protocols. Sample loading of 1 µg of total protein were used for each injection.

Chromatography: A NanoLC™ 425 System plumbed for microflow chromatography was used (5 µµL/min) and operated in trap/elute mode. Column temperature was controlled at 30°C. Gradients of 5, 10, 20, or 45 minutes were tested. More information on LC configuration can be found in the SWATH Performance Kit SOP.2

Mass Spectrometry: All data was acquired using a TripleTOF 6600 System with the Turbo V™ Source equipped with the 25 µm hybrid electrodes for microflow LC or the OptiFlow® TurboV source. SWATH acquisition data were collected using a variety of acquisition strategies. The TOF MS scan was 150 msec and the # of Q1 windows/cycle and accumulation time was varied.

Data Processing: Data independent acquisition (DIA) data was processed using SWATH Acquisition microapp in PeakView® Software 2.2 and the Pan Human Library.3 . Results were evaluated using the SWATH Acquisition Replicates Template.

REFERENCES

1. Microflow SWATH Acquisition for Industrialized Quantitative Proteomics. SCIEX Technical note RUO-MKT-02-3637-B
2. SWATH Performance Kit SOP Protocol.pdf

SCIEX
The Power of Precision

3. Rosenberger G et al. (2014) Scientific Data, 1, 140031.

74

Cryptosporidiosis modulates gut microbiome metabolism and the immune response in an infected host.

Avinash Karpe[1]

1. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Dutton Park, QUEENSLAND, Australia

Background and aims: Cryptosporidiosis is a major global human health concern. Our understanding of the infection mechanism remains limited, compounding the difficulty of clinical diagnosis. This study investigated the underlying biochemistry of host-microbiome-parasite relationships during infection.

Methods: C57BL/6J mice were infected with 1 x 10^5*Cryptosporidium parvum* oocysts via oral gavage. Faecal samples were collected daily, while blood, liver tissues and luminal contents of the small and large intestines were collected 10 days post infection. High-resolution liquid chromatography and low-resolution gas chromatography coupled with mass spectrometry were used to analyse the proteomes and metabolomes of faeces, serum, liver, and luminal contents. Faecal samples were additionally subjected to 16S rRNA gene sequencing. Univariate and multivariate statistical analysis were applied to all datasets.

Results: Host and microbial energy pathways altered during infection. Glycolysis/citrate cycle metabolites, such as malate and lactate, were elevated in the large intestine. Short-chain fatty acids, formate and acetate, increased in the small intestine, while butanoate increased in the caecum-colon. This correlated with an increased abundance of bacteria associated with a stressed host environment, including *Lactobacillus* (small intestine) and *Coriobacteriaceae* (throughout the intestine, but more prominently in colon). The expression of host electron transfer flavoprotein, phosphoglycerate kinase and acetyl CoA binding proteins, yeast glyceraldehyde-3-phosphate dehydrogenase, and *Lactobacillus* glyceraldehyde-3-phosphate dehydrogenase significantly increased in the infected gut. Liver oxalate increase was also seen during infection.

Conclusions: The microbiome-parasite relationship is more influential than the previously thought host-parasite relationship, in mediating major biochemical changes in the mouse gut during cryptosporidiosis. Defining this parasite-microbiome interaction is the first step towards building a comprehensive cryptosporidiosis model.

75

Immunopeptidogenomics: harnessing RNA-seq to illuminate the dark immunopeptidome

Katherine E Scull[1], Kirti Pandey[1], Sri H Ramarathinam[1], Anthony W Purcell[1]

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

Background: Human leukocyte antigen (HLA) molecules are cell-surface glycoproteins that present peptide antigens for surveillance by T lymphocytes seeking signs of disease. Mass spectrometric analysis allows us to identify large numbers of these peptides (the immunopeptidome) following affinity purification of HLA-peptide complexes from cell lysates. However, in recent years there has been a growing awareness of the 'dark side' of the immunopeptidome: unconventional peptide epitopes, including neoepitopes in cancer, which elude detection by conventional search methods because their sequences are not present in reference protein databases.

Methodologies: Here we establish a bioinformatic workflow to aid identification of peptides generated by non-canonical translation of mRNA or genome variants. The workflow incorporates both standard transcriptomics software and novel computer programs to produce cell line-specific protein databases based on 3-frame translation of the transcriptome. The final protein database also includes sequences resulting from variants determined by variant calling on the same RNA-seq data. We then search our experimental data against both transcriptome-based and standard databases using PEAKS Studio. Finally, further novel software helps to compare the various result sets arising for each sample, pinpoint putative genomic origins for the identified unconventional sequences, and highlight potential neoepitopes.

Results: We have trialled the workflow to study the immunopeptidome of the acute myeloid leukaemia cell line THP-1, using RNA-seq and mass spectrometric immunopeptidome data. We confidently identified over 14000 peptides from 3 replicates of purified THP-1 HLA peptides using UniProt. Using the transcriptome-based database, we recapitulated >75% of these, and also identified over 927 unconventional peptides, including 14 sequences caused by non-synonymous variants.

Conclusions: Our workflow, which we term 'immunopeptidogenomics', can provide databases which include pertinent unconventional sequences, allowing neoepitope discovery in cancer studies, without becoming unsearchably large. Immunopeptidogenomics is a step towards the unbiased search approaches needed to illuminate the dark side of the immunopeptidome.

76

The role of non-proteasomal proteases in antigen presentation in melanoma

Ritchlynn Mr Aranha[1], Anthony Prof Purcell[1], Nathan Dr Croft[1], Pouya Dr Faridi[1], Andreas Dr Behren[2]

1. Monash University, Clayton, VIC, Australia

2. Cancer Immunobiology, Olivia Newton-John Cancer Research Institute, Austin Hospital, Heidelberg, Victoria, Australia

Background: The proteasome is a multi-subunit protease that degrades intracellular proteins into short peptides, which are then presented on the cell surface in complex with Human Leukocyte Antigen (HLA) molecules to T-cells. This HLA-bound peptide repertoire is termed the immunopeptidome. These peptides conventionally derive from contiguous sequences of proteins, yet recent studies have identified that a proportion can result from splicing of non-contiguous regions of the same or different proteins. Such spliced peptides have also been identified as being immunogenic in human cancers. Whilst most studies suggest a predominant role of the proteasome in catalysing peptide

SCIEX
The Power of Precision

splicing, other proteases crucial to antigen presentation may also contribute to the spliced immunopeptidome. Endoplasmic reticulum aminopeptidase (ERAP) is a protease that trims HLA-bound peptides at their N-terminus. Studies have shown that its inhibition leads to an increase in the presentation of immunogenic peptides, but the extent to which this inhibition affects peptide splicing remains unexplored. This study has utilised a novel ERAP inhibitor on patient-derived melanoma cell lines to investigate its impact on the immunopeptidome and identify potential novel peptide targets for T-cell immunotherapy.

***Methods**:* This study combines peptide sequencing performed by data-independent acquisition (DIA) mass spectrometry with PEAKS Studio software and novel bioinformatics algorithms to identify and quantify linear and spliced peptides presented by patient-derived melanoma cells under normal and ERAP1 inhibitor-treated conditions (500ng/ml and 2µg/ml; novel ERAP1 inhibitor from Grey Wolf Therapeutics).

***Results**:* We generated a spectral library of ~120,000 peptides from replicates of immunopeptidomics data derived from patient-derived melanoma cell lines. This library was then used to interrogate DIA data, which revealed ~5000-11000 unique peptides quantified per cell line. Spliced peptides made up between 23-34% of each immunopeptidome. ERAP1 inhibition led to a remodelling of the immunopeptidome with only ~20% overlap between peptides presented under untreated and inhibited conditions, an increase in 10-12mer peptides, and an overall decrease in the proportion of spliced peptides by ~7%. Importantly, inhibition of ERAP led to a significant increase in the presentation of peptides derived from melanoma-associated antigens (MAA).

***Conclusion**:* In this study, we observed that ERAP1 inhibitor treatment led to significant changes in the HLA-bound immunopeptidome. The repertoire of ERAP inhibited peptides contains significantly more MAA derived peptides as compared to the conventional peptidome. ERAP1 inhibition also led to a decrease in the proportion of spliced peptides, which shows the potential role of ERAP in splicing. These results suggest that inhibition of non-proteasomal proteases can change the immunopeptidome creating potentially novel targets for T-cell immunotherapy.

77

Benchmarking targeted analysis pipelines in data-independent acquisition mass spectrometry-based immunopeptidomics

Mohammad Shahbazy[1], Pouya Faridi[1], Sri H. Ramarathinam[1], Emma C. Jappe[1, 2], Patricia T. Illing[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

2. Evaxion Biotech, Bredgade 34E, DK-1260 Copenhagen, Denmark

Background Immunopeptidomes are the peptide antigen repertoires bound by the molecules encoded by the major histocompatibility complex (MHC) (human leukocyte antigen (HLA) in humans). These HLA-peptide complexes are presented on the surface of cells for recognition by T cells of the immune system. Immunopeptidomics denotes the utilization of tandem mass spectrometry (MS/MS) to identify and quantify peptides bound to HLA molecules. Data-independent acquisition (DIA) has emerged as a powerful strategy for deep proteome-wide profiling, but DIA's application to immunopeptidomic analyses has so far seen limited use. Further, of the many DIA data processing tools currently available, there is no consensus in the immunopeptidomics community on the most appropriate pipeline for in-depth and accurate HLA peptide identification. Herein we benchmarked four conventional and recently developed spectral library-based pipelines for processing and targeted analysis of DIA data for label-free immunopeptidome quantification.

Methodology We immunoprecipitated HLA molecules from replicates of 5×10^7 cells of two cell lines, C1R-B*57:01 and C1R-A*02:01, and eluted HLA-bound peptides for DIA analysis on an Orbitrap Fusion™ Tribrid™ mass spectrometer. Data analysis was evaluated across four peptide-centric DIA software tools (Skyline, Spectronaut, DIA-NN, and PEAKS X+) using an extensive DDA library previously acquired from both cell lines.

Results DIA analyses allowed for comparisons of immunopeptidome coverage, reproducibility between replicates, and estimates of external false-discovery rates for each data set. In general, DIA-NN was found to achieve the greatest number of peptide identifications on average at FDR 1% for both cell lines (~3000 peptides for C1R-B*57:01 and 2000 peptides for C1R-A*02:01), with the three other tools reporting fewer peptides (~1200-3000 peptides for C1R-B*57:01 and ~1200-1500 peptides for C1R-A*02:01). Through a hybrid spectral library containing HLA-B*57:01 and HLA-A*02:01 peptides, we examined external false-positive rates achieved by each software. In this context, Skyline and DIA-NN achieved lower false-positive rates for C1R-B*57:01 data, whereas Spectronaut and DIA-NN showed better performance for C1R-A*02:01. In examining the HLA peptide identification robustness of pipelines across replicates, Spectronaut, and DIA-NN provided higher reproducibility. We also reported linearity metrics and other analytical figures of merit in pairwise comparisons to validate the results.

Conclusions Through this extensive analysis, we propose a pathway for users to choose an appropriate tool based on research aims. The current data suggest a combined strategy of applying at least two complementary DIA software tools to achieve the greatest degree of confidence and in-depth coverage of immunopeptidome data.

79

Explorations of major eukaryotic "cell fate regulator" at the base of eukaryogenesis: the Pumilio homology superfamily in 'basal' eukaryote Giardia duodenalis

Balu Balan[1, 2], Jarrod J Sandow[1], Ahmad Wardak[1], Waruni Abeysekera[1], Swapnil Tichkule[1], Brendan Robert E. Ansell[1], Amrita Vijay[1], Laura Rojas[3], Olivia S Rissland[4], Samantha J Emery-Corbin[1], Peter Czabotar[1], Andrew I Webb[1], Staffan G Svärd[3], Aaron R 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

4. Molecular Biology, University of Colorado, Denver, Colorado, USA

SCIEX
The Power of Precision

Pre-programming ('cell-fating') has been explored extensively in eukaryotic cells, including developmental biology and stem-cell differentiation. It is heavily regulated through post-transcriptional regulation (PTR), and particularly translational repression. RNA-binding proteins (RBPs) are essential in PTR, including in cell-fating and translation repression. Despite significant conservation of RBPs across the tree of life, there are notable periods of transition, particularly with eukaryogenesis. As eukaryotic lineages evolved, they expanded on bacterial and archaeal RNA Binding domains (RBDs) and formulated "novel" domains, notably including the emergence of pumilio homology proteins (PUF). The emergence of PUFs was a major innovation underpinning "cell fate" decisions that guide eukaryotic stage-transition, cellular development and embryogenesis. Intriguingly, RBP biology and functionality appears almost entirely conserved from yeast through to humans. This leaves open questions as to when eukaryotic RBP regulation evolved, what functionality these early RBP systems had, and what the early eukaryotic RBPome looked like. Such systems are essential to study basic elements of RBP biology and regulation in, presumably, their simplest form.

The phylum Metamonada (Protista), which includes human gastrointestinal parasite *Giardia duodenalis,* is one of the earliest diverging extant lineages of the Eukaryota and appears to be the oldest lineage in which many eukaryotic RBP systems, include PUF proteins, first appear. As part of a broader exploration of the earliest eukaroytic RBPome, we explored these PUFs for complex eukaryotic functions. We undertook bioinformatic mining and 3D protein structural homology modelling of *Giardia* PUF proteins followed by molecular simulations to map homologous RNA interacting residues within PUFs. We used transcriptomic and proteomic data to explore the kinetic behaviour of PUFs through *Giardia*'s stage transitions, combined with interactome capture of poly-A mRNA-bound proteins to confirm that the giardial PUFs interact with RNA. We further employed PAR-CLIP strategies to identify the diverse target RNAs bound to *Giardia* PUFs and assessed the functional impact of CRISPR-based PUF-silencing on *Giardia* stage transition. Lastly, we used *in vitro* phase separation assays to test if *Giardia* PUF can mediate RNA granule formation, which is an essential step in RBP-based regulation of cell-fate. This first, comprehensive study of the earliest eukaryotic PUFs show that the RBP systems are as complex at the base of the eukaryotic tree as they are at the tips. We propose that the evolution of eukaryotic RBPs likely occurred at, and may have played a crucial role in, the emergence of eukaryotes.

80

Proteomic profiling of human trophectoderm small extracellular vesicles reveal insights into implantation and remodeling through endometrial cell reprogramming

Qi Hui Poh[2, 1], **Alin Rai**[2], **Iska Carmichael**[3], **Lois A Salamonsen**[4], **David W Greening**[2, 1]

1. Biochemistry and Genetics, La Trobe University, Melbourne, Victoria, Australia

2. Baker Heart and Diabetes Institute, Melbourne, VIC, Australia

3. Monash Micro Imaging, Monash University, Melbourne, VIC, Australia

4. Hudson Institute of Medical Research, Melbourne, VIC, Australia

Embryo implantation into a receptive endometrium is a critical step in establishing pregnancy, with implantation failure accounting for up to two-thirds of unsuccessful pregnancies. This tightly regulated process requires reciprocal signalling between the blastocyst and endometrium, though factors regulating this crosstalk remain poorly understood. Maternally-derived factors, including endometrial extracellular vesicles (EVs), are known to signal to the embryo during implantation, however, the role of embryo-derived EVs remains largely unknown. Here, we provide a comprehensive proteomic characterisation of a major class of EVs, termed small EVs (sEVs), released by cells from the outer layer of human embryos (trophectoderm). Highly purified sEVs obtained by density-gradient separation were characterised based on morphology, shape, size, and expression of classical sEV markers. A striking finding through proteomic dissection of sEVs was the enrichment of proteins involved in embryo development, immune regulation, and antioxidant activity; processes critical for successful implantation. We demonstrate that the sEVs are readily internalised by human endometrial cells, where they potentially upregulate expression of adhesion molecules in endometrial cells. Our findings provide critical molecular insights into trophectoderm sEV-mediated signalling to the endometrium during implantation.

81

Straight from the heart: Characterisation of cardiac extracellular vesicles

Bethany Claridge[2, 1], **Haoyun Fang**[2], **Alin Rai**[2], **David W Greening**[2, 1]

1. Biochemistry and Genetics, La Trobe University, Melbourne, VIC, Australia

2. Baker Heart and Diabetes Institute, Melbourne, VIC, Australia

Intra- and inter-organ signalling are critical for normal function, contribute to disease progression, and are processes in which extracellular vesicles (EVs) are emerging as fundamental mediating entities. However, our current understanding of cardiac EVs (cEVs) has been acquired from *in vitro* culture systems, a limitation resultant of challenges with tissue-EV extraction. Here, we present a novel strategy to isolate cEVs using gentle enzymatic perfusion of an intact heart. cEVs are 100-300 nm in diameter, attainable at levels of 2-3µg protein/heart, and carry many classical EV markers associated with both surface and endosomal origin. Proteomic dissection of cEVs highlights enrichment of components associated with cardiac regulatory functions, including signalling, peptide-hormone processing, and extracellular matrix maintenance, as well proteins enriched in organs (i.e., liver, kidney) other than the heart. This unique insight into cEVs directly from heart provides the basis for further investigation into cEVs and how cardiovascular diseases dysregulate their form and function in intra-cardiac and inter-organ communication.

SCIEX
The Power of Precision

82

Multi-Omic exploration of human skeletal muscle after endurance training identifies extensive mitochondrial biogenesis and remodelling.

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

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

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

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

4. Metabolomics Laboratory, Baker Heart & Diabetes Institute, Melbourne, Victoria, Australia

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

In addition to generating the bulk of cellular energy, mitochondria direct a vast array of biological functions essential for cellular homeostasis. Mutations affecting mitochondrial function and biogenesis can lead to a variety of pathological conditions that affect tissues of high energy demand, such as the skeletal muscle. A long-standing question in biology concerns the biogenesis of mitochondria in these tissues and its regulation in response to stress and the metabolic needs of the cellular environment. Increased need for mitochondrial energy, for example during times of increased muscle contraction, represents a major challenge to both these pathways, with exercise being arguably one of the most 'natural' perturbations experienced by our tissues. Despite this, there has been little study into exactly how mitochondria adapt to changing demands of the host tissue. In order to further explore the effects exercise has on the mitochondria within skeletal muscle, ten participants performed three different training volumes over 12 weeks. Training phases included a combination of normal-, high- and reduced-volume training regimens, with muscle biopsies taken following each phase. A combination of RNA-seq, quantitative proteomics, and lipidomics was performed on tissue biopsies and mitochondrial isolates, allowing for a holistic visualisation of the effects that exercise has on the mitochondria and associated pathways. Surprisingly, the significant changes in tissue respiratory capacity with increasing exercise volume could be attributed to an increase in mitochondrial content, rather than an increase in the efficiency of mitochondrial respiration within mitochondria. To accommodate extensive biogenesis of new mitochondria, we found that mitochondria initially prioritise tricarboxylic acid cycle linked fatty acid oxidation, with biogenesis of oxidative phosphorylation (OXPHOS) complexes peaking at high volumes of exercise. Related transcripts as well as proteins involved in the biogenesis of OXPHOS complexes preceded this phase, peaking following normal volume training, suggesting the delay is related to biogenesis and assembly of the complexes. Although we observed an increase in abundance of cardiolipins with exercise volume, a lipid solely found within the membranes of mitochondria, we found that at high volumes the dominant tetra-linoleoyl cardiolipin was supplanted by other species. Cessation of high-volume exercise rapidly reversed most, but not all of these changes. Our findings therefore provide important insights into how tissues accommodate the acute proliferation of mitochondria while maintaining the need for uninterrupted energy supply.

83

Proteomics at scale: Experimental design and automated data management for large clinical cohorts

Ahmed Mohamed[1, 2], Julian Kelabora[1], Laura Dagley[1], Melissa Davis[2], Andrew Webb[1]

1. Colonial Foundation Healthy Ageing Centre, WEHI, Parkville, Vic, Australia

2. Bioinformatics division, WEHI, Parkville, Vic, Australia

Scaling proteomics profiling to large clinical cohorts presents challenges in mitigating experimental variability and streamlining data management. The prevalence and effects of unanticipated confounding variables such as batching effect or instrument fluctuations over time, can directly hampers the discovery of robust biomarkers for such studies. Additionally, traditional software tools for proteomics data analysis are not yet tailored to processing thousands of samples. Here, we present a scalable experimental design allows minimization and estimation of confounding variables, through combination of batch design, global QC samples and spiked-in internal standards. We also established a cloud-based fully automated data processing and archiving workflow that is petabyte scalable. The workflow is hooked to a user-friendly interactive interface allowing rapid identification of instrument- and sample-related issues. The workflow will be utilised to acquire and process proteomics data for the 22,000 samples of the ASPREE cohort.

84

Evaluation of prmPASEF for multiplexed targeted proteomics

Antoine Lesur[1], Jens Decker[2], Sven Brehmer[2], Elisabeth Letellier[3], Francois Bernardin[1], Pierre-Olivier Schmit[4], Gunnar Dittmar[1]

1. Quantitative Biology Unit, Luxembourg Institute of Health, Strassen Luxembourg

2. Bruker Daltonik GmbH, Bremen, Germany

3. Department of Life Sciences and Medicine, University of Luxembourg, Luxembourg

4. Bruker France, Wissembourg, France

Introduction

prm-PASEF is a targeted acquisition method that fully exploits the multiplexing capability and the high resolution of TIMS-Q-TOF mass spectrometers. Multiple peptides can be sequentially measured from a single ion mobility scan without compromising the sensitivity. We evaluated the reproducibility, sensitivity, accuracy and dynamic range of the method by using aqua peptides spiked both in a Hela cell line digest and a depleted serum samples digest. . Finally, we applied the method to quantify the mutations and isoforms of the Ras oncoproteins family in cancer cell lines.

SCIEX
The Power of Precision

Methods

The quantitative performance of prm-PASEF was evaluated with a tryptic digest of HeLa cells and depleted human serum samples spiked with 201 AQUA peptides and 15 light peptides. Serial dilutions ranging from 5.5 to 50,000 amole/µl were generated with 15 heavy/light peptides pairs. The other 186 AQUA peptides were spiked at the constant concentration of 2 fmole/µl. Ten cancer cell lines were analyzed for screening the Ras mutations. All acquisition were performed on a timsTOF Pro instrument.

Results

The Aqua peptides calibration curves showed a signal response fitted by a linear regression over a concentration factor of 2900 (from 17.2 to 50,000 amole injected column), and an averaged RSD of 3% for the heavy/light peptide signal ratios. For label free quantification, the 2 fmole 186 AQUA peptides measured over 30 LC-MS runs showed a median RSD of 10.5% and a median of 25 data points across the chromatographic peaks. Finally, the G12 mutations of Ras protein family and the three isoforms (NRas, KRas and HRas) were identified and quantified in 10 colon and lung cancer cells lines using a 10 min chromatography separation.

Conclusions

We developed a new targeted acquisition method that takes advantages of trapped ion mobility and fast high-resolution Q-TOF. The method is well suited for clinical applications that require to measure high density of targets with fast chromatography separations.

85

High throughput proteomics - Application of dia-PASEF for short gradients

Adam Rainczuk[1]

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

Stephanie Kaspar-Schoenefeld[1], Markus Lubeck[1], Adam Rainczuk[2], Thomas Kosinski[1], Scarlet Koch[1], Oliver Raether[1], Gary Kruppa[1]

Institutes

[1]Bruker Daltonik GmbH, 28359 Bremen, Germany

[2]Bruker Australia, Melbourne, VIC

Introduction

Data-independent acquisition (DIA) promises reproducible and accurate protein identification and quantification across large sample cohorts by using wide selection windows to ensure that all precursor ions are fragmented in every sample. Ion mobility separation provides an additional dimension for separation of complex proteomics samples, that can also be used for alignment of precursor and fragment information. Here, we combine the PASEF technology (Meier et al., 2018) with a DIA approach and investigate the potential for complex proteomics samples using short gradients.

Methods

An in-house tryptic digest of HeLa was analyzed by coupling an Evosep One system (Evosep Biosystems) online to a trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro, Bruker Daltonics). A dia-PASEF scheme optimized for the short gradient methods has been used for targeting +2 and +3 ions in a three-window method covering an m/z range from 400 to 1000 with a total cycle time of 900ms per dia-PASEF cycle. Data processing was done using Spectronaut 14 (Biognosys).

Results

DIA workflows rely on spectral libraries for the correlation of quantitative data from fragment ion spectra with peptide identifications. We used PASEF in DDA mode and fractionated samples to assemble the resource-specific library using Spectronaut software. The library comprised 8,381 protein groups and 93,301 peptide sequences in ~10h of data acquisition. We applied the hybrid library approach supported in Spectronaut for shorter gradient data by combining the resource-specific library with a project-specific library. This workflow allows keeping retention time precision of the shorter gradients for targeted data extraction. Using the comprehensive libraries, we could identify and quantify on average 5,204 protein groups and 39,936 peptide sequences using 60 SPD method at a 1% FDR.

Conclusion

The dia-PASEF method in combination with short gradients enables high sample throughput without sacrificing depth and quantitative accuracy of analysis.

86

Evaluation of prm-PASEF for multiplexed targeted proteomics.

Adam Rainczuk[1]

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

Introduction

prm-PASEF is a targeted acquisition method that fully exploits the multiplexing capability and the high resolution of TIMS-Q-TOF mass spectrometers. Multiple peptides can be sequentially measured from a single ion mobility scan without compromising the sensitivity. We evaluated the reproducibility, sensitivity, accuracy and dynamic range of the method by using aqua peptides spiked in a Hela cell line digest and a depleted serum samples digest. Finally, we applied the method to quantify the mutations and isoforms of the Ras oncoproteins family in cancer cell lines.

SCIEX
The Power of Precision

Methods

The quantitative performance of prm-PASEF was evaluated with a tryptic digest of HeLa cells and depleted human serum samples spiked with 201 AQUA peptides and 15 light peptides. Serial dilutions ranging from 5.5 to 50,000 amole/µl were generated with 15 heavy/light peptides pairs. The other 186 AQUA peptides were spiked at the constant concentration of 2 fmole/µl. Ten cancer cell lines were analyzed for screening the Ras mutations. All acquisition were performed on a timsTOF Pro instrument.

Results :

The Aqua peptides calibration curves showed a signal response fitted by a linear regression over a concentration factor of 2900 (from 17.2 to 50,000 amole injected column), and an averaged RSD of 3% for the heavy/light peptide signal ratios. For label free quantification, the 2 fmole 186 AQUA peptides measured over 30 LC-MS runs showed a median RSD of 10.5% and a median of 25 data points across the chromatographic peaks. Finally, the G12 mutations of Ras protein family and the three isoforms (NRas, KRas and HRas) were identified and quantified in 10 colon and lung cancer cells lines using a 10 min chromatography separation.

Conclusions

We developed a new targeted acquisition method that takes advantages of trapped ion mobility and fast high-resolution Q-TOF. The method is well suited for clinical applications that require to measure high density of targets with fast chromatography separations.

87

OneOmics™ Suite – Proteomics Processing Pipeline in the Cloud

Jeremy Potriquet[1]

1. SCIEX, Mt Waverley, VIC, Australia

- Easy to use cloud processing pipeline developed for the speed and scale requirements of today's industrialized proteomics labs
- Identify proteins, perform iTRAQ® Reagent experiments and build spectral ion libraries with the **ID core**
- Process SWATH® acquisition data, normalize and compute protein fold change difference across sample set with the **QUANT core**
- Visualize protein expression data, perform cluster/trend analysis to find significant protein changes, compare datasets, and obtain ontology information using the **VISUALIZATION core**

SCIEX Cloud

Analyze Faster

- Up to 10x faster processing of DIA than desktop
- Up to 4x faster processing of DDA data than desktop
- User friendly app environment for large scale data processing

Securely store and manage

- AWS security standards, multi-site synchronous storage
- Standard 256-bit AES (Advanced Encryption Standard)
- Send data from the instrument to cloud environment automatically and securely
- Reduce need for onsite storage and computing power

Share and Collaborate

- Share projects (data, results) easily
- No more FTP sites or shipping of hard drives

TRADEMARKS/LICENSING

For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein are the property of AB Sciex Pte. Ltd., or their respective owners, in the United States and/or certain other countries.

AB SCIEX™ is being used under license. © 2019 DH Tech. Dev. Pte. Ltd. RUO-MKT-10-8463-A

SCIEX
The Power of Precision

88

Hexosaminidase disruption impedes the biosynthesis of paucimannosidic proteins in HL-60 cells

Julian Ugonotti[1], Ian Loke[2], Yuqi Zhu[3], Sayantani Chatterjee[1], Harry C Tjondro[1], Zeynep Sumer-Bayraktar[1], Meral Topçu[4], Sriram Neelamegham[3], Morten Thaysen-Andersen[1]

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

2. Cordlife Group Limited, Singapore

3. Department of Chemical and Biological Engineering, University at Buffalo, Buffalo, New York, United States of America

4. Department of Child Neurology, Hacettepe University, Ankara, Turkey

We have previously documented that human neutrophils are actively expressing immune-related glycoproteins carrying paucimannosidic glycans ($Man_{1-3}GlcNAc_2Fuc_{0-1}$), a functional yet under-explored class of *N*-glycosylation in the human glycobiology. Whilst inference from the paucimannose-rich invertebrates and plants points to a β-*N*-acetyl-hexosaminidase (Hex)-driven production of paucimannosidic proteins from immature β1,2-linked *N*-acetylglucosamine-terminating glycoprotein intermediates in the mammalian glycosylation machinery, this hypothesis remains experimentally unsupported in human cells including neutrophils. We therefore sought to obtain evidence for the involvement of the Hex isoenzymes, which in humans exist in homodimeric (αα and ββ) and heterodimeric (αβ) variants arising from two coding genes, *HEXA* (α) and *HEXB* (β), in the biosynthesis of paucimannosidic proteins in human neutrophils. Our approach was to perform quantitative PGC-LC-MS/MS-based *N*-glycomics and C18-LC-MS/MS-based glycoproteomics of several CRISPR-Cas9-edited Hex-deficient neutrophil-like HL-60 mutant cell lines (two *HEXA*-/- and two *HEXB*-/- mutant lines) relative to an unedited HL-60 control line. Accurate disruption of the target genes and absence of off-target genetic mutations were validated using next-generation sequencing and established Hex activity assays. Both *HEXA* (24.0%) and *HEXB* (22.8%) disruption led to a relatively weak yet still significant reduction of paucimannosidic *N*-glycans relative to unedited HL-60 (33.3%). In particular, $Man_{2-3}GlcNAc_2Fuc_{0-1}$ was reduced in the Hex-deficient mutants relative to unedited cells. Interestingly, an *N*-glycomics follow-up analysis of mature blood neutrophils from a patient diagnosed with early onset Sandhoff disease (*HEXB*-/-) recapitulated a reduced expression level of paucimannosidic *N*-glycans (20.7%) relative to levels found in mature blood neutrophils from a healthy age-matched donor (40.5%). This study provides evidence to support that several Hex isoenzyme variants encoded by *HEXA* and *HEXB* are directly responsible for the biosynthesis of paucimannosidic proteins in human neutrophils. These well-characterised Hex-deficient mutant cell lines displaying supressed levels of protein paucimannosylation will be useful to further our understanding of the functional roles of paucimannosidic proteins in neutrophil glycoimmunology.

89

Assessing an *E. coli* dataset by Top-down and Bottom-up proteomics using two-dimensional separation

Matthew J Fitzhenry[1], David I Cantor[1], Nima Sayyadi[2], Ardeshir Amirkhani[1]

1. Australian Proteome Analysis Facility, Macquarie Analytical and Workshop Facility, Macquarie University, North Ryde, NSW, Australia

2. Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Macquarie University, North Ryde, NSW, Australia

Top-down proteomics, the direct analysis of intact proteoforms by tandem mass spectrometry is becoming more widely available. The ready availability of high resolution mass spectrometers such as the Orbitrap has widely allowed more laboratories to perform top-down proteomics. The vast majority of proteomic analyses, being performed by bottom-up methods, are unable to directly measure proteoforms due to enzymatic digestion. Although top-down proteomics is currently constrained by technical limitations, these can be partially overcome by integrating top-down and bottom-up data generated in parallel. Here we describe the data processing of top-down and bottom-up datasets from *E. coli* K12 using the software programs Proteoform Suite and MetaMorpheus (Lloyd Smith group, University of Wisconsin). To expand the repertoire of top-down proteomics data processing programs, this top-down data is also processed by the programs TopPIC Suite (Xiaowen Liu group, University Indiana-Purdue UI) and MASH Explorer (Ying Ge group, University of Wisconsin).

E. coli K12 lysis was performed under denaturing conditions (1% SDS/PBS) then extracted proteins were separated in two dimensions. The first dimension separation was performed by Gel-Eluted Liquid Fractionation Entrapment Electrophoresis (GELFrEE) fractionation, the second dimension was separated by reverse-phase chromatography (RPC). Additionally, extracted proteins were digested with trypsin, desalted with S-trap (Protifi), then peptides fractionated by High pH (RPC). Top-down and Bottom-up workflows were processed in parallel using GELFrEE and S-trap methods, respectively. Top-down and bottom-up datasets were searched by MetaMorpheus then combined using Proteoform Suite. Top-down data was deconvoluted and database searched using TopPIC suite and validated with MASH Explorer.

90

Unexpected glycan signatures of the RCL of human corticosteroid-binding globulin impact its susceptibility to proteolytic cleavage

Anastasia Chernykh[1], Jodie L. Abrahams[1, 2], Zeynep Sumer-Bayraktar[3, 1], Rebeca Kawahara[1], Morten Thaysen-Andersen[4, 1]

1. Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia

2. Glycosciences Laboratory, Department of Metabolism, Digestion and Reproduction, Imperial College London, London W12 0NN, United Kingdom

3. Glycometabolic Biochemistry Team, Cluster of Pioneering Research, RIKEN, Wako, Saitama 351-0198, Japan

4. Biomolecular Discovery Research Centre, Macquarie University, Sydney, NSW 2109, Australia

The liver-derived corticosteroid-binding globulin (CBG) is a heavily *N*-glycosylated protein that transports anti-inflammatory cortisol in human plasma. We have previously demonstrated that Asn347-glycans positioned on the exposed reactive centre loop (RCL) of human CBG, impact the release of cortisol by modulating the RCL cleavage process catalysed by neutrophil elastase (NE). However, a comprehensive

SCIEX
The Power of Precision

structural and functional characterisation of the glycans decorating the RCL is still required to unravel the molecular mechanisms underpinning cortisol delivery to inflamed tissues. To this end, we have glycoprofiled human CBG isolated from pooled healthy donor sera (hCBG) using porous graphitised carbon-LC-MS/MS-based glycomics with parallel exoglycosidase digestion and in combination with C18-LC-MS/MS-based glycopeptide analysis employing CID-, HCD- and EThcD-MS/MS. Deep glycan and glycopeptide profiling was also performed on recombinant human CBG (rhCBG) produced in HEK293 cells. Longitudinal NE-based cleavage experiments with downstream glycopeptide detection were performed using rhCBG to study the impact of the RCL specific glycosylation on the NE-mediated cleavage process. Glycomics of the isolated RCL-glycopeptides of hCBG provided a quantitative profile of Asn347-*N*-glycans and their fine structural features including the assignment of core- and antenna-fucosylation, the elucidation of bi- and tri-antennary *N*-glycans and multiple α2,3- and α2,6-sialyl linkage isomers. Surprisingly, glycoproteomics of hCBG revealed a hitherto unknown presence of RCL *O*-glycosylation. The data confirmed that $NeuAc_{1-2}Gal_1GalNAc_1$*O*-glycan structures occupy, albeit at low levels, four out of six possible *O*-glycosylation sites within the RCL region including a strategically positioned Thr345 site in close proximity to the NE cleavage site. Quantitative analysis revealed that the hCBG RCL was ~18.7% non-glycosylated, ~80.0% *N*-glycosylated, ~0.5% *O*-glycosylated, and ~0.8% both *N*- and *O*-glycosylated. In contrast, the rhCBG glycoprofiling revealed a relatively high level of Thr345 *O*-glycosylation ($NeuAc_{1-2}Gal_1GalNAc_1$) (5.9%) while glycosylation of other possible *O*-glycosylation sites and Asn347 was negligible. Interestingly, cleavage experiments indicated that the Thr345-glycosylation of rhCBG strongly impedes NE-mediated proteolysis of the RCL. This suggests a functional significance of the RCL *O*-glycosylation, reported here for the first time. In conclusion, this is the first comprehensive structural and functional site-specific characterisation of the glycans decorating the RCL of native and recombinant human CBG. We present new functionally-important glycosylation features of the RCL that improve our understanding of the molecular mechanisms governing the timely and tissue-specific delivery of cortisol to inflammatory sites.

91

Sexual dimorphism in the liver's response to intermittent fasting

Dylan Harney[1], Michelle Cielesh[1], Mark Larance[1]

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

In natural environments animals normally experience periods of feast and famine due to the intermittent availability of food. Indeed, intermittent fasting (IF) provides beneficial effects relative to providing constant access to food, which shortens lifespan and produces undesirable metabolic outcomes. Strikingly, all previous molecular studies of the response to intermittent food availability have exclusively analysed male animals. However, it is axiomatic that males and females differ in many aspects of their physiology, and their response to intermittent feeding/starvation is no exception. For example, in humans females live longer, have more lipid-based energy reserves, and are more resistant to infection compared to males. Global mRNA studies in mice have also shown many genes are expressed in a sexually dimorphic pattern including those in the liver, a key nutrient-responsive organ. Here, we have employed proteome analysis of mouse liver to identify proteins whose abundance is regulated by IF, but are also sexually dimorphic in their response. From >6,800 proteins quantified, 663 proteins were significantly altered by the IF intervention. Of these IF-responsive proteins, 171 had a significant interaction with the gender of the animals indicating sexual dimorphism. Proteins showing the largest interaction included IFIT1 (innate immune response), ACOT2, (fatty acid degradation), SERPINA1 (major secreted protease inhibitor), and SELENBP1 (organosulfur degradation). Gene-set enrichment analysis showed that the interferon-alpha pathway was a major contributor to the sexual dimorphism with the transcription factor STAT1 being the likely master regulator of this response. These data correlate with the known differences in immune response between genders, where female liver is less likely to become infected by viral pathogens due to heightened interferon signaling. But how the gender difference in interferon signaling is generated for non-infected animals exposed to intermittent fasting is unknown. We hypothesise that the large flux in free fatty acids produced during intermittent fasting can lead to activation of STAT1 and thereby the entire downstream interferon responsive machinery, which we are currently testing by exogenous injection experiments across genders.

92

Using smartphones to explore protein structures in XR

Neblina Sikta[1], Stuart Anderson[2], Christian Stolte[1], Sandeep Kaur[1, 3], Bosco Ho[1], Nicola Bordin[4], Matt Adcock[2], Andrea Schafferhans[5, 6], Sean O'Donoghue[1, 2, 3]

1. Garvan Institute of Medical Research, Darlinghurst, NSW, Australia

2. CSIRO Data61, Sydney, Australia

3. School of Biotechnology and Biomolecular Sciences (UNSW), Sydney, Australia

4. Institute of Structural and Molecular Biology, University College London, UK

5. Department of Bioengineering Sciences, Weihenstephan-Tr. University of Applied Sciences, Freising, Germany

6. Department of Informatics, Bioinformatics & Computational Biology, Technical University of Munich, Germany

Proteins fold into intricate 3D shapes that can often be difficult to navigate and understand. Thus, it was recognized already in the 1970's that virtual reality (VR) has potential to help in protein research, leading to decades of research prototypes (Brooks 2014). Still today, however, the use of VR in molecular graphics is primarily limited to demonstration systems that require specialist hardware. However, this may be able to change, thanks to the advent of 'augmented reality' (AR), in which virtual objects are displayed interactively with the physical world. AR and VR are closely related technologies - here, we use the term 'extended reality' (XR) to encompass both. Most iOS and Android smartphones on the market today have impressive XR capabilities; in this project, we have created a production application that lets researchers use smartphones to explore protein structures in XR. Our application is a completely redesigned version of Aquaria (O'Donoghue et al. 2015), a web-based molecular graphics system with >100 million pre-calculated protein structure models, based on systematically matching all SwissProt sequences against all PDB structures. These 3D models can be mapped with features, which can be either user-defined, or predefined in CATH (Dawson et al. 2017), COSMIC (Tate et al. 2019), PredictProtein (Yachdav et al. 2014), SNAP2 (Hecht, Bromberg, and Rost 2015), or UniProt (The UniProt Consortium 2019). With the redesigned version (https://aquaria.app), it is now easy to use a smartphone to find structural models of interest, color them using mapped sequence features, then explore the colored models in XR.

SCIEX
The Power of Precision

All structures related to a protein can be found just by specifying a gene name (e.g., https://aquaria.app/**Human**/**WT1**); mutations can also be specified directly in the URL (e.g., https://aquaria.app/**Human**/**WT1**?**Arg370Leu**). Our goal was to make it easy for researchers to experience using XR to explore structures of direct relevance to their work - and even mapped with their own feature data. This experience is now possible with most smartphones currently in use. Our application also supports more specialist hardware, such as Microsoft HoloLens or other devices compatible with Windows Mixed Reality.

1. Brooks, Frederick P. 2014. “Impressions by a Dinosaur – Summary of Faraday Discussion 169: Molecular Simulations and Visualization.” Faraday Discussions 169 (September): 521–27. https://doi.org/10.1039/C4FD00130C.
2. Dawson, Natalie L., Tony E. Lewis, Sayoni Das, Jonathan G. Lees, David Lee, Paul Ashford, Christine A. Orengo, and Ian Sillitoe. 2017. “CATH: An Expanded Resource to Predict Protein Function through Structure and Sequence.” Nucleic Acids Research 45 (D1): D289–95. https://doi.org/10.1093/nar/gkw1098.
3. Hecht, Maximilian, Yana Bromberg, and Burkhard Rost. 2015. “Better Prediction of Functional Effects for Sequence Variants.” BMC Genomics 16 (8): S1. https://doi.org/10.1186/1471-2164-16-S8-S1.
4. O’Donoghue, Seán I., Kenneth S. Sabir, Maria Kalemanov, Christian Stolte, Benjamin Wellmann, Vivian Ho, Manfred Roos, et al. 2015. “Aquaria: Simplifying Discovery and Insight from Protein Structures.” Nature Methods 12 (2): 98–99. https://doi.org/10.1038/nmeth.3258.
5. Tate, John G, Sally Bamford, Harry C Jubb, Zbyslaw Sondka, David M Beare, Nidhi Bindal, Harry Boutselakis, et al. 2019. “COSMIC: The Catalogue Of Somatic Mutations In Cancer.” Nucleic Acids Research 47 (D1): D941–47. https://doi.org/10.1093/nar/gky1015.
6. The UniProt Consortium. 2019. “UniProt: A Worldwide Hub of Protein Knowledge.” Nucleic Acids Research 47 (D1): D506–15. https://doi.org/10.1093/nar/gky1049.
7. Yachdav, Guy, Edda Kloppmann, Laszlo Kajan, Maximilian Hecht, Tatyana Goldberg, Tobias Hamp, Peter Hönigschmid, et al. 2014. “PredictProtein—an Open Resource for Online Prediction of Protein Structural and Functional Features.” Nucleic Acids Research 42 (W1): W337–43. https://doi.org/10.1093/nar/gku366.

93

Proteomic characterisation of ancient Egyptian skin and bones

Dylan H Multari[1], Prathiba Ravishankar[1], Constance Lord[2], James Fraser[2], Ronika K Power[3], Paul A Haynes[1]

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

2. Chau Chak Wing Museum, University of Sydney, Sydney, NSW, Australia

3. Department of Ancient History, Macquarie University, Sydney, NSW, Australia

We were among the first to publish data concerning identification of proteins from 4200-year-old ancient Egyptian skin fragments, sampled from the collection of the Egyptian Museum in Turin [1]. Following on from this work, we are currently undertaking an analysis of ancient Egyptian skin and bone samples, as part of the University of Sydney interdisciplinary mummy project. Sample materials have been made available from the coffin and remains of Mer-Neith-it-es, a 26th dynasty (664-525 BCE) mummy excavated from Saqqara, Egypt. The remains are heavily fragmented, likely due to frequent incidences of grave robbing, and the transportation of the coffin by sea to Sydney in 1859.

One of the aims of our study was to develop a non-invasive technique for sampling ancient organic remains, using readily available equipment and materials. We set out to examine whether dermatology grade skin sampling strip tape could be successfully applied to the analysis of such materials. Successful development of a non-destructive sampling method would be a big step forward in bioarchaeological proteomics, because it would enable access to a much wider range of ancient materials housed in Museum collections.

Preliminary experiments have shown that we can identify intracellular protein components on the surface of skull fragments, which strongly suggests that they are indeed ancient remains rather than modern contamination. This presentation will include detailed results of proteomic analysis of skull and bone fragments using this novel non-destructive sampling approach.

1. [1] Jones J. et al. Philosophical Transactions of the Royal Society part A. 2016 Oct 28;374(2079).

94

Abiotic stress factors in combination change the proteome of rice plants

Fatemeh Habibpourmehraban[1], Yunqi Wu[1], Jemma Wu[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 remarkable part of the staple diet for half of the world’s population and is grown in every continent apart from Antarctica. Stresses are often unfavorable for rice growth. Abiotic stress is the primary cause of crop loss worldwide, reducing average yield for most major crops by more than 50%. 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. Salt, drought and temperature stresses are major environmental abiotic factors all of which negatively influence the yield of crops, limit plant productivity and threaten our food security. Elucidating the various mechanisms of plant response to stress, and their roles in acquired stress tolerance, is thus of great practical important. Plants respond to multiple abiotic stresses differently from individual stresses, activating a specific program of gene and protein regulation relating to the exact 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 at different time points.

SCIEX
The Power of Precision

We have investigated the proteomic response to multiple abiotic stresses in two varieties of rice (IAC1131 and Nipponbare) during their vegetative growing stage. The stress treatment was a combination of reducing the soil water content to 50% field capacity, with NaCl concentration increased to 50mM and temperature of 33/18 ºC for 2 and 4 days. Proteins were identified and quantified using TMT labeling, following protein extraction from 3-week leaf tissues combined with trypsin in-solution digestion. Peptides were separated and identified using a Q Exactive Orbitrap mass spectrometer coupled to an Easy-nLC 1000 nano-flow HPLC system. Proteome Discoverer v2.1 software was used to process raw files generated by Xcalibur software. Peptide to spectrum matching was performed using FASTA files of protein sequences from *Oryza sativa*. Relative quantitation of proteins was achieved by pairwise comparison of normalized TMT reporter ion intensities using TMTPrepPro. Deferentially expressed proteins in both genotypes under stress conditions were observed, with a greater number of proteins increase in abundance rather than decreased. Heat shock proteins and late embryogenesis abundant proteins were two of the most significantly altered protein groups, while hypothetical proteins with no specific function were also found to be deferentially expressed in response to stress.

95

Proteome changes in the brain and eye of APP/PS1 mice illustrate molecular mechanisms of Alzheimer's disease at the early stage

Liting Deng[1], **Stuart Graham** [2], **Vivek Gupta** [2], **Mehdi Mirzaei**[2], **Paul Haynes**[1]

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

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

Alzheimer's disease (AD) is the most common ageing-related neurodegenerative disease and contributes to 60-70% of dementia cases. Early diagnosis of AD is vital for timely disease stabilization and treatment. However, AD development follows distinctive patterns in different brain parts in the same individual, making treating this pathology tremendously difficult. Identifying the biochemical changes in specific brain regions is key to comprehending the neuropathological mechanisms in early pre-symptomatic phases of AD. Increased amyloid β (Aβ) aggregation is a hallmark feature of AD pathology, but the molecular mechanisms induced by Aβ toxicity, especially in the early stage of AD, have not been clearly demonstrated. To better understand the proteome alterations by early AD, and explore the brain-site specific protein regulation, quantitative and comparative proteomic analysis was performed on four brain regions (hippocampus, frontal and parietal cortices, and cerebellum) from 2.5 months old APP/PS1 double transgenic AD model mice and matched controls. The greatest proteome perturbation was detected in the hippocampus and frontal cortex (AD-susceptible brain regions), compared to fewer changed proteins in the cerebellum (less vulnerable region to AD). The expression of the majority of the other proteins between hippocampus and cortices was not similar, highlighting differential effects of the disease on specific brain regions, and the fundamental compositional and functional shifts even in early AD. In addition to brain abnormities, pathologies are also exhibited in AD eye, including reduced retinal function and other degenerative changes. Investigating molecular changes induced by Aβ in the brain and eye is an active area for the discovery of potential ocular biomarkers for AD diagnosis and treatment. However, the overall molecular effects of AD on the retina remain undetermined. Proteomic profiling of retinas from the same AD mice model was performed followed by biochemical pathway enrichment analysis. Protein changes in AD brain and retina were compared. Up-regulated App was found in all brain areas and retinas, indicating AD effects on both organs. A total of 50 common regulated proteins between two organs were found with similar yet distinct expression patterns, such as App processing related proteolytic enzymes cathepsin D (Ctsd) and Na(+)/K(+) ATPase alpha-1 subunit (Atp1a1), illustrating complicated responses in brain and eye to AD. This study highlights proteome and biochemical pathway alterations in the brain and eye that underlie the early stages of AD pathology. The generated molecular datasets will broaden our knowledge of AD, and provide a framework for future longitudinal studies.

96

Temporal ordering of omics and multiomic events inferred from time-series data

Sandeep Ms Kaur[1], **Timothy Dr Peters**[2], **Pengyi Dr Yang**[3], **Laurence Dr Luu**[4], **Jenny Dr Vuong**[2], **James Dr Krycer**[3], **Sean Dr O'Donoghue**[1]

1. UNSW and Garvan, Kogarah, NSW, Australia

2. Garvan Institute, Darlinghurst

3. USYD, Syndey

4. BABS, UNSW, Kensington

Temporal changes in omics events can now be routinely measured; however, current analysis methods are often inadequate, especially for multiomics experiments. We report a novel analysis method called Minardo-Model[1] - that can infer events (such as phosphorylation, dephosphorylation), and temporal ordering of events. The temporal ordering of events is inferred at a better temporal resolution than the experiment The identified events, and the temporal ordering are presented via two novel, concise and intuitive visualisation techniques called event maps and event sparklines. We tested Minardo-Model on two time series datasets, a phosphoproteomics dataset and a multiomics dataset consisting of transcriptomic, proteomic and phosphoproteomic measurements. The ordering revealed by our method correlated well with prior knowledge and indicated that our method streamlines analysis of time-series data.

1. Kaur, S., Peters, T.J., Yang, P. et al. Temporal ordering of omics and multiomic events inferred from time-series data. npj Syst Biol Appl 6, 22 (2020). https://doi.org/10.1038/s41540-020-0141-0

SCIEX
The Power of Precision

97

Personalised phosphoproteomics of insulin action potentiated by exercise

Elise J Needham[1], Janne R Hingst[2], Benjamin L Parker[3], Christian Pehmoller[4], Jorgen FP Wojtaszewski[2], Sean J Humphrey[1], David E James[1]

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. Internal Medicine Research Unit, Pfizer Global Research and Development, Cambridge, MA, USA

Human phenotypic variation underpins health outcomes, but defining its molecular basis represents a major biological challenge. Here we introduce “personalised phosphoproteomics”, which links longitudinal physiological measures with subject-specific dynamic phosphoproteome signatures to extract biologically relevant signalling events. We applied this to unravel how exercise potentiates muscle insulin sensitivity by sampling paired muscle biopsies from rested and exercised legs during an insulin infusion alongside muscle-specific in vivo glucose uptake measures. There was considerable individual variance in the phenotypic response to both exercise and insulin and this was mirrored at the level of individual phosphoproteomes. Associations exploiting this variance identified kinases and phosphosites functionally linked with the observed potentiation of insulin action by exercise. Among these were AMPK and mTORC1. Previously thought to be mutually exclusive, AMPK and mTORC1 were both active in the post-exercise state, exhibiting cooperativity through mTORC1 phosphorylation of the AMPK α2 subunit at Serine 377. This study highlights that protein phosphorylation viewed through the lens of dynamic human phenotypic variance can provide mechanistic insights into complex biology, including how exercise promotes insulin sensitivity.

98

Quantitative mass spectrometry of cysteine redox post translational modifications in the diabetic heart

Alexander Rookyard[1, 2], Desmond Li[3], Stuart Cordwell[1, 2, 4, 5], Melanie White[2, 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. Heart Research Institute , Sydney, NSW, Australia

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

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

The global rate of diabetes mellitus is projected to grow to 700 million by 2045, of which type 2 diabetes mellitus (T2DM) accounts for ~90% of incidences. T2DM develops in response to caloric excess and hyperlipidaemia which leads to peripheral insulin resistance, pancreatic insufficiency, and eventual hyperglycaemia. In addition to this, T2DM is an independent risk factor for cardiovascular disease (CVD). Increased oxidative stress via reactive oxygen species (ROS) and reactive nitrogen species (RNS) has been implicated with the aetiology of the diabetic heart, however, the targets of ROS/RNS are poorly defined. Reactive Cys can readily undergo redox post translation modifications (PTMs), however, with the prolonged exposure anticipated in the diabetic setting, irreversible PTM of Cys (Cys-SO_2H and Cys-SO_3H) are predicted to be a feature. We utilised our pre-clinical rodent model of T2DM (high fat diet and streptozotocin), to excise the hearts for Langendorff ex-vivo perfusion, where a progressive loss of function was observed over time. This was functional depression was prevented when perfused with the antioxidant intervention with *N*-2-mercaptopropoinylglycine (MPG). Enrichment of irreversible Cys PTMs using strong cation exchange followed by hydrophilic interaction chromatography and identification by LC-MS/MS permitted identification of close to 700 sites, of which 242 were not identified in the control hearts and 90 were unique to the diabetic phenotype. Modified Cys-SO_3H were increased, localised to the mitochondria and contractile filament. This correlative increase was further confirmed by quantitation of 160 sites by PRM-MS, of which 20 sites were modified in the combined HFD-STZ setting. Krebs cycle enzyme malate dehydrogenase (MDH) was shown to be highly modified by irreversible oxidation in the diabetic hearts. In the heart, MDH interacts with the malate-aspartate shuttle (MAS), which is the primary pathway responsible for the transfer of reducing equivalents between the cytosol and mitochondria, making it an attractive target for increased redox modifications.

99

Increasing the Ease of Use of Nanoflow with Plug and Play Low Flow Source

Keith Ashman[1]

1. AB Sciex Australia Pty Ltd, Galston, NSW, Australia

When performing electrospray ionization mass spectrometry, it is well understood that reducing chromatographic flow rates can improve sampling efficiency and therefore increase sensitivity.1 When operating in the nanoflow regime (50-500 nL/min) very high sampling efficiencies have been measured for favorable solvent conditions.2 This is particularly beneficial when sample amounts are limited, such as in proteomics applications, where nanoflow LC-MS is frequently used. One drawback to this approach is that typical run times are quite long; often 1-2 hour gradients are used to achieve optimal separations.

In other application areas, throughput and robustness are more important, such as with pharma bioanalysis studies. For the analysis of 100s of samples daily, analytical flow rates of 300-1000 µL/min are typically combined with gradients of less than 5-10 minutes. In recent years, there has been a resurgence in the interest in using microflow LC coupled to mass spectrometry as evolving research demands increasingly require more sensitivity while still maintaining good throughput and robustness.

SCIEX
The Power of Precision

This led SCIEX to develop the OptiFlow® Source, a single source that would cover the full spectrum of low flow rates for high sensitivity LC-MS analysis. Significant research was done to develop a low flow source that was sensitive, while maintaining high robustness and ease of use of higher flow sources. The ability to easily switch between flow regimes was also critical in the source design. Removing the barrier of switching helps researchers choose the right flow rate for every project.

REFERENCES

1.Covey, T. R., Schneider, B. B., Javaheri, H., Y., L. J. C., Ivosev, G., Corr, J. J., Kovarik, P., Electrospray and MALDI mass spectrometry: Fundamentals, instrumentation, practicalities, and biological applications. In ESI, APCI, and MALDI. A Comparison of the Central Analytical Figures of Merit: Sensitivity, Reproducibility, and Speed. Cole, R. B., Ed. Hoboken: John Wiley & Sons, Inc.: (2010); pp 443-490.

2.Schnieder BB, Javaheri H, Covey TR (2006) Ion Sampling Efficiency Under Conditions of Total Solvent Consumption. *Rapid Comm Mass Spectrometry* **20**, 1538-1544.

3.SWATH Acquisition Performance Kit Standard Operating Protocol.

4.OptiFlow™ Interface for TripleTOF 6600 System - Switch from Nanoflow LC to Microflow LC in Minutes. SCIEX Technical Note RUO-MKT-02-7219-B.

100

Quantitative lipidomic analysis of soil-transmitted nematode *Ascaris suum*

Tao Wang[1], Shuai Nie[2], Guangxu Ma[1], Johnny Vlaminck[3], Peter Geldhof[3], Nicholas Williamson[2], Gavin Reid[4, 5, 6], 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. Laboratory of Parasitology, Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium

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

Human ascariasis, caused by soil-transmitted nematode *Ascaris*, is one of the most important and commonest neglected tropical diseases. Approximately 0.8 billion people are affected with this worm worldwide, equating to 0.86 million disability-adjusted life-years (DALYs). Lipids are of vital importance in biology of parasitic worms, particularly in relation to cellular membranes, energy storage, and intra- and intercellular signalling. However, very little is known about the biology of lipids in parasitic nematodes. Using a high-throughput LC-MS/MS approach, we characterised the first global lipidome for *Ascaris*. Totally, we confidently identified and quantified (i.e. in precise molar amounts in relation to the dry weigh of worm material) nearly 600 lipid species across 18 lipid classes in five key developmental stages/sexes and three organ systems of adult male and female *Ascaris*. The results showed substantial differences in the composition and abundance of lipids with key roles in cellular processes and functions (e.g. energy storage regulation and membrane structure) among distinct stages and among organ systems, likely reflecting differing demands for lipids, depending on stage of growth and development as well as the need to adapt to constantly changing environments within and outside of the host animal. This work provides the first step toward understanding the biology of lipids in *Ascaris*, with possibilities to work toward designing new interventions against ascariasis.

101

Phosphoproteome profiling uncovers a key role for CDKs in TNF signaling

Maria C Tanzer[1], Isabell Bludau[1], Che A Stafford[2], Veit Hornung[2], Matthias Mann[1]

1. Max-Planck institute of biochemistry, Munich/Planegg, BAVARIA, Germany

2. Innate Immunity, gene center munich, LMU, Munich, Bavaria, Germany

Tumor necrosis factor (TNF) is one of the few cytokines successfully targeted by therapies against inflammatory diseases. However, blocking this well studied and pleiotropic ligand can cause dramatic side-effects. We reasoned that a systems-level proteomic analysis of TNF signaling could dissect its diverse functions and offer a base for developing more targeted therapies. Combining phosphoproteomics time course experiments with subcellular localization and kinase inhibitor analysis identifies functional modules of phosphorylations. The majority of regulated phosphorylations could be assigned to an upstream kinase by inhibiting master kinases. Spatial proteomics revealed phosphorylation-dependent translocations of hundreds of proteins upon TNF stimulation. Phosphoproteome analysis of TNF-induced apoptosis and necroptosis uncovered a key role for transcriptional cyclin-dependent kinase (CDK) activity to promote cytokine production and prevent excessive cell death downstream of the TNF signaling receptor.

SCIEX
The Power of Precision

102

Comparative study of excretory-secretory proteins released by *Schistosoma mansoni*-resistant, susceptible and naïve *Biomphalaria glabrata*

Conor Fogarty[1], **Tianfang Wang**[1], **Scott Cummins**[1], **Min Zhao**[1], **Mary Duke**[2], **Don McManus**[2]

1. University of the Sunshine Coast, Caloundra, QUEENSLAND, Australia

2. Molecular Parasitology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Schistosomiasis is a widespread neglected tropical disease caused by digenetic trematodes called *Schistosoma*. Due to the current lack of a viable vaccine and the inability of mass drug administration of praziquantel to eliminate the parasite from the ecosystem, alternative management methods are being considered. These include interfering with process of *Schistosoma* infection of their intermediate molluscan host. Chemical analyses of host-parasite interactions reveal that chemosensation is vital to molluscan host identification, including naïve host preference. Proteomic technique advances enable sophisticated comparative analyses between infected and naïve snail host proteins. This study compared resistant, susceptible and naïve *Biomphalaria glabrata* snail-conditioned water (SCW) to identify potential attractants and deterrents. Behavioural bioassays were conducted on *Schistosoma mansoni* miracidia to compare the effects of susceptible, F1 resistant and naïve *B. glabrata* SCW. The F1 resistant and susceptible *B. glabrata* SCW excretory–secretory proteins (ESPs) were fractionated using SDS-PAGE, identified with LC-MS/MS and compared to naïve snail ESPs. Protein-protein interaction (PPI) analyses based on published studies (including experiments, co-expression, text-mining and gene fusion) identified *S. mansoni* and *B. glabrata* protein interaction. A respective total of 291, 410 and 597 ESPs were detected in the susceptible, F1 resistant and naïve SCW. Less overlap in ESPs was identified between susceptible and naïve snails than F1 resistant and naïve snails. F1 resistant *B. glabrata* ESPs were predominately associated with anti-pathogen activity and detoxification, such as leukocyte elastase and peroxiredoxin. Several susceptible *B. glabrata* proteins correlated with immunity and anti-inflammation, such as glutathione S-transferase and zinc metalloproteinase, and *S. mansoni* sporocyst presence. PPI analyses found that uncharacterised *S. mansoni* protein Smp_142140.1 potentially interacts with numerous *B. glabrata* proteins. This study identified ESPs released by F1 resistant, susceptible and naïve *B. glabrata* to explain *S. mansoni* miracidia interplay. Susceptible *B. glabrata* ESPs shed light on potential *S. mansoni* miracidia deterrents. Further targeted research on specific ESPs identified in this study could help inhibit *B. glabrata* and *S. mansoni* interactions and stop human schistosomiasis.

103

Strategies and Methodologies for Quantification of Senescence-Derived Biomarkers in Human Plasma

Sandip K Patel[1], **Nathan Basisty**[1], **Toshiko Tanaka**[2], **Luigi Ferrucci**[2], **Judy Campisi**[1, 3], **Birgit Schilling**[1]

1. Buck Institute for Research on Aging, Novato, CA, USA

2. Translational Gerontology Branch, National Institute on Aging, NIH, Baltimore, MD, USA

3. Lawrence Berkeley Laboratory, University of California, Berkeley, Berkeley, CA, USA

Aging is a complex biological process associated with progressive loss of physiological function and susceptibility to several diseases, such as cancer and neurodegeneration. Cellular senescence and mitochondrial dysfunction have been defined as classical hallmarks of aging and neurodegenerative diseases, but non-invasive biomarkers of these processes, such as plasma proteins, are lacking. We hypothesized that senescent cells and their senescence-associated secretory phenotype (SASP) may qualify as biomarkers for aging and age-related diseases. Our studies thus far have demonstrated that selected SASP-derived biomarker candidates (e.g., GDF15, STC1, SERPINs, MMP1) serve as indicators for human aging as demonstrated in multiple human plasma cohort studies, such as The Baltimore Longitudinal Study of Aging (BLSA) and The Invecchiare in Chianti (InCHIANTI) study.

Another largely underexplored part of the SASP are exosomes – extracellular vesicles - that are involved in signaling from cell to cell, and that carry various types of cargo, such as proteins, RNA, DNAs, metabolites and lipids. We have thus started to not only investigate soluble SASP (SASPAtlas.com), but also to comprehensively characterize exosomes from senescent cells to develop senescence-derived exosome biomarkers. The latter also includes investigations of 'Mitochondrial Dysfunction-Associated Senescence' (MiDAS) in a tissue culture model to better understand the molecular links between mitochondrial dysfunction and cellular senescence, and likely these mechanisms may be relevant during aging and disease.

In order to efficiently analyze plasma exosomes we developed a high-throughput method to isolate plasma exosomes by sequential size-exclusion chromatography (SEC) and ultrafiltration (UF) to overcome challenges of exosome contamination with abundant soluble plasma proteins. Quality control analysis of the isolated exosomes confirmed high exosome 'purity'. We performed data-dependent acquisitions (DDA) from offline high-pH reversed-phase fractions of exosome lysate to generate a deep spectral library of ~2,300 proteins. Subsequently, in a pilot aging study we used comprehensive data-independent acquisitions (DIA), to compare plasma exosomes from young (n=5; 20–26 yrs) and old (n=5; 60–66 yrs) individuals. We were able to quantify 1,208 total exosome proteins, and 144 proteins significantly changed between young and old plasma groups (Q<0.05; >1.5-fold change). We also analyzed the exosome miRNA cargo and detected 331 miRNA, some of which changed significantly between young and old individuals. These translational workflows aim to develop 'Biomarkers of Aging' for prognostic and diagnostic applications.

In the future, we intend to perform multi-omics analysis of plasma exosomes obtained from human aging cohort studies as well as from dementia patients – specifically analyzing longitudinal plasma samples.

SCIEX
The Power of Precision

104

Band 3 Anion transport protein is up-regulated in high-grade serous ovarian cancer patients with acquired chemotherapy-resistant disease, as identified by MALDI-MSI

Tannith M Noye[1], Parul Mittal[2], Noor Alia A Lokman[1], Tara L Pakala[2], Peter Hoffmann[3], Martin K Oehler[4, 1], Carmela Ricciardelli[1]

1. Adelaide Medical School, Obstetrics and Gynaecology, University of Adelaide, Adelaide, South Australia, Australia

2. Adelaide Proteomics Centre, School of Biological Sciences, University of Adelaide, Adelaide, South Australia, Australia

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

4. Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

Ovarian cancer is the most lethal gynaecological disease, with poor 5-year survival rates and limited treatment options for patients who develop resistant disease. The majority of ovarian malignancies, up to 70% of cases, are high-grade serous carcinomas that have high chemosensitivity to first-line platinum-based therapies. However, 75% of patients will become chemoresistant, following relapse. The underlying mechanism for developing resistance to chemotherapy in ovarian cancer is poorly understood. In this study, we employed peptide matrix-assisted laser desorption/ionisation mass spectrometry imaging (MALDI-MSI) in formalin-fixed paraffin-embedded sections of ovarian cancer tissues at the time of diagnosis and following relapse from 4 patients with serous cancer. Using, MALDI-MSI we have identified *m/z* features that were present in relapsed tissues but absent in ovarian cancer tissues at diagnosis. One of these identified proteins is Band 3 anion transporter (SCL4A1) using LC-MS/MS and data dependant analysis on paraffin sections. SCL4A1 was validated by immunohistochemistry and was elevated in relapse tissues compared to matching ovarian cancer tissues at diagnosis in three of the four patients ($p<0.05$, Mann-Whitney U). SCL4A1 positivity was significantly increased in serous ovarian cancer tissues following chemotherapy treatment and relapse ($P=0.0093$) In addition using online databases, high *SCL4A1* expression was significantly associated with reduced progression-free survival ($p=0.0004$) and reduced overall survival ($p=0.004$). In conclusion, MALDI-MSI has the potential to identify proteins associated with chemotherapy resistance, which can be used as a novel therapeutic target.

105

Exploring albendazole resistance of human diarrheal parasite *Giardia duodenalis*

Qiao Su[1, 2], Samantha Emery-Corbin[1], Louise Baker[1, 3], Swapnil Tichkule[1, 2], Ern Lacey[4], Balu Balan[1, 3], 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. Microbial Screening Technologies Pty. Ltd., Smithfield, New South Wales, Australia

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

Giardia duodenalis is a gastrointestinal parasite causing ~1 billion infections and > 200 million cases of diarrheal disease (giardiasis) worldwide annually, with a significant contribution to malabsorption syndrome and post-infectious gastrointestinal disorders. Treatment of giardiasis is limited to nitroheterocyclic (e.g., metronidazole; MTZ) and benzimidazole (BZ, e.g. albendazole; ALB) drugs. However, treatment failure occurs in up to 20% of cases, and clinical resistance is reported for each drug class. MTZ acts as an oxidizing agent and kills by damaging DNA, proteins and membrane. Consistent with this, MTZ-resistance involves complex changes in oxidoreductive enzymes, metabolism, stress responses and major post-transcriptional and post-translational regulatory systems. In parasitic nematodes and other metazoans, ALB inhibits β-tubulin and prevents microtubule polymerization and ALB-resistance is caused by point mutations in the ALB-binding site of β-tubulin isoform 1, resulting in point mutations at amino acid positions F167, E198 and F200. In contrast, relatively little is known about ALB mode of action (MOA) or resistance in *Giardia*, and no point mutations in *Giardia* β-tubulin have been reported with ALB resistance to date.

Herein, we undertook *in vitro* drug-susceptibility screens comparing isogenic ALB susceptible (ALBS) and culture derived ALB resistant (ALBR) *Giardia duodenalis* for 13 BZ compounds and 7 ALB structural analogues. The cross-resistance of ALBR line to many other BZ compounds and the decreased potency of ALB analogues with longer thiol substituents indicated mutations in the β-tubulin binding-site for ALB as a core component of the resistance phenotype. This was supported by amplicon sequencing and molecular modelling. However, *Giardia* is intriguing among parasites treated with ALB in that it has lost myosin and its microtubule system is significantly expanded to a wide range of functional roles as a consequence. To further explore the consequences of perturbing this system, we undertook transcriptomic and proteomic analysis of these ALBS and ALBR isogenic lines using RNA and proteins co-purified from the same cell pellet. These data indicated changes in protein turnover, glycolysis, arginine metabolism, vesicular trafficking, cytoskeleton organization, and tubulin regulation, with ALBR, and pointed to potential involvement of post-transcriptional and post-translational regulation. This work provides new insight into benzimidazole resistance, provides a new avenue to develop clinical diagnostic tests for resistant parasites, and generates new data on β-tubulin drug-interactions that will inform novel microtubule-inhibitor development.

SCIEX
The Power of Precision

106

Effect of gas phase fractionation on protein identification rates in discovery proteomics of barley lines

Mahya Bahmani 1 , Angela Juhasz 1 , Hugh Dunn 2 , Ian B Edwards 3 , Mitchell G Nye-Wood 1 , Michelle L Colgrave 1 4

1. Food and Agriculture Proteomics, Edith Cowan University, Perth, Western Australia, Australia

2. Pilot Malting Australia (PMA) , School of Science, Edith Cowan University, Perth, Western Australia, Australia

3. Edstar Genetics Pty Ltd, Perth, Western Australia, Australia

4. Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Brisbane, Queensland, Australia

Barley is the most important material for malting and brewing, and the fourth major cereal being produced globally second in Australian cereal production. Plant breeding approaches aim to incorporate desired traits for industry benefit, including optimised seed germination and maturation, and end-use in malting and brewing.

Protein content is one of the crucial parameters that affects and determines barley's final application. In this research, a bottom-up mass spectrometry approach is used to investigate the proteome of twenty barley lines to inform the barley breeding program. Selected barley lines with different protein profiles have been studied to explore the relationship between proteins and traits. To identify proteins in these barley lines, a data-dependent acquisition (DDA) and a gas phase fractionation method was applied to a pooled biological quality control sample (PBQC). The results showed that by applying gas phase fractionation a higher number of proteins were identified (22%). This method was used to acquire data to generate the spectral library that will be used for global quantitative proteomics to identify those proteins that are linked with desirable malting characteristics and the high protein trait. The obtained result will assist in understanding how genetic and quantitative variability in protein content and composition affects the use of these experimental barley lines in food and beverage products.

107

Proteogenomic study to investigate the impact of a transcription factor mutation on nutritional value and grain protein composition in barley

Angela Juhasz[1], Utpal Bose[2], Ronald Yu[3], Mahya Bahmani[1], Keren Byrne[2], Malcolm Blundell[3], James A Broadbent[2], Crispin A Howitt[3], Michelle L Colgrave[1]

1. Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, School of Science, Edith Cowan University, Joondalup, WA, Australia

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

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

In cereals, lysine is the most limiting essential amino acid, and attempts have been made over the decades to improve the nutritional quality of these grains by increasing lysine content, while maintaining favourable agronomic traits. The high proportion of lysine-poor prolamin storage proteins in cereals are associated with the sub-optimal nutritional quality of cereal grains. To this end, the single *lys3* mutation caused by the mutation in the PBF transcription factor in barley has been shown to significantly increases lysine content but reduces grain size and the content of both starch and gluten. Yet, the low agronomic yield and germination defects have presented as barriers for the commercial uptake of these barley mutants.

Conventional breeding strategies were used to combine the malting barley cultivar Sloop with the high lysine/low gluten lines, Risø 56 and Risø 1508, or the Ethiopian genotype R118 with a decreased D-hordein content to generate the single null B-, C- and D-hordein lines, respectively. These single-null lines were intercrossed to produce double-null lines that further reduced the gluten content combined with elevated lysine content.

The mechanisms activated in the grains of these double-null mutant lines to compensate for the loss of these major gluten proteins and decreased starch content remain unknown. In the current study, data-independent acquisition (DIA) mass spectrometry analyses were performed in parallel with phenotypic characterisations. The aim was to study the large-scale quantitative changes in proteins within the hordein double null lines in comparison to their parent lines. Functional annotation and bioinformatic analyses were carried out to uncover the protein classes related to the increased lysine content and gluten reduction in the double null lines. Balanced changes in the induced and suppressed protein abundances confirms the dual regulatory mechanisms of nutrient accumulation and energy metabolism in the mutant lines, resulting in differences both in the hordein levels and composition as well as in the lysine content of the double nulls. This research serves as a framework for future proteomics-assisted crop development in order to study pleiotropic effects on safety and nutrition quality-related improvements.

SCIEX
The Power of Precision

108

Discovery proteomic analysis of malt barley proteome – Understanding the impact of genetic variation on key malt quality traits.

Clare E O'Lone[1], Mitchell Nye-Wood[1], David Moody[2], Hugh Dunn[3], Angela Juhasz[1], Jean-Philippe F Ral[4], Michelle L Colgrave[5, 1]

1. Food & Agriculture Proteomics, School of Science, Edith Cowan University, Joondalup, Western Australia, Australia

2. Cereal Breeding, InterGrain, Bibra Lake, Western Australia, Australia

3. Pilot Malting Australia, Edith Cowan University, Joondalup, Western Australia, Australia

4. Agriculture & Food, New Markets, CSIRO, Black Mountian, Canberra, Australia

5. Agriculture & Food, Future Protein, CSIRO, St Lucia, Queensland, Australia

Abstract

Malting is a controlled germination of the barley grain in preparation for brewing, distilling, or food manufacturer. Where exposure to moisture leads to the activation and synthesis of enzymes that in turn modifies the physical structure of the endosperm, converting stored starch into simple sugars, yielding a friable malt suitable for downstream use. Different barley varieties, defined by a distinct genotype, exhibit different malting characteristics that are attributable to both the way the grain responds to water, and the abundance of isoforms of key enzymes that are expressed. Identifying the proteins that control this response is important to barley breeders in developing varieties with desirable malting characteristics.

In this investigation, we used discovery proteomics (DDA) and compositional analysis to achieve an unbiased characterisation of the barley proteome before and after malting, in three barley varieties that differ in protein content and response to water. We focused on three InterGrain barley genotypes: (1) Flinders, an established premium variety; (2) Maximus, a high protein variety malted to unique specification, and (3) IGB1467, a trial breeding line with a unique phenotype of malting at lower water content.

High-performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS), and computational biology was used to define the proteins that are differentially expressed between barley grain and malt barley in the three InterGrain genotypes. The results will provide insight into how the expression of different protein isoforms or classes of proteins in the three genotypes influence key malt quality traits. Understanding the difference in protein expression and their impact on the level of modification will provide foundational knowledge for further research.

109

Quantifying gluten and allergens in a novel "low gluten" wheat variety

Mitchell G. Nye-Wood[1], Angela Juhasz[1], Michelle Colgrave[1, 2]

1. Food and Agricultural Proteomics, School of Science, Edith Cowan University, Perth, Western Australia, Australia

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

Gluten refers to a class of seed storage proteins present in wheat and related species. They contribute nutritional and functional attributes to pasta, bread, and other cereal products, but provoke an immune reaction in people susceptible to wheat allergy, coeliac disease, and non-coeliac gluten sensitivity. The market for gluten-free flour is largely supplied by flours from plant species that do not express gluten proteins, though there is also a 'low gluten' market for consumers who experience less severe reactions to dietary wheat products. In this vein, a non-GMO wheat flour that purports to contain '65% less allergenic gluten' was recently brought to market. We sought to understand this alteration in protein profile and test its claim.

We conducted a bottom-up proteomics approach on the low-gluten flour, as well as a mixed wheat flour as control, using independent chymotryptic and tryptic digests of gluten-enriched extracts. Liquid chromatography-mass spectrometry (LC-MS) was performed on a SCIEX TripleTOF 6600 with microflow LC to determine the identities of proteins present, altogether revealing 635 distinct proteins at a 1% global false discovery rate. Multiple reaction monitoring assays were developed for the targeted quantitation of tryptic and chymotryptic wheat gluten peptides on a SCIEX QTRAP 6500+ with ExionLC system. Of these, 359 peptides were used to measure the abundance of gluten and allergenic proteins. Due to a high degree of sequence homology and shared peptides, proteins were categorised into groups by the presence of known domains. Relative abundance of protein groups in the low-gluten wheat and control wheat were plotted, and proteins were mapped to the wheat genome to reveal the chromosome locations that were altered in the 'low gluten' wheat variety. These data were used to determine gluten proteins that were present or absent in the novel low-gluten wheat, and the genome locations selected against in this variety. Of note, several omega- and gamma-gliadins, and low-molecular weight glutenins mapping to the short arm of chromosome 1, as well as alpha-gliadins from the chromosome 6 short arm were absent or expressed at lower levels in the low-gluten variety. In contrast, the high-molecular weight glutenins and alpha-amylase/trypsin inhibitors were notably more abundant in this variety.

SCIEX
The Power of Precision

110

Confirmatory analysis of snake venoms by LC-HRMS for application to coronial toxicology

Sarah McCabe[1], Thao T. Le[1, 2], Leigh Boyd[1], Ashley Nolan[1], Joel P.A. Gummer[1, 2], Nicola Beckett[1], Bianca Douglas[1], Colin Priddis[1]

1. ChemCentre, Bentley, WA, Australia

2. School of Science, Edith Cowan University, Joondalup, Western Australia, Australia

Animal venoms contain a significant contribution of proteins and peptides. Co-evolution of venomous animals has resulted in a diversity of biologically active venom proteins, which are integral to an animal's defence and in subduing prey. Snake venoms more specifically contain a variety of bioactive proteins and peptides possessing pharmacological and toxicological activities; most dominantly in the form of phospholipases A2, metalloproteinases, serine proteases and three-finger peptides. Due to the composition, exposure to venom by snake bite can induce hemotoxic, neurotoxic and cytotoxic effects in the bitten animal. The successful diagnosis of snake bite, including the definitive determination of species of origin, and thus successful path to treatment, can be erroneous. Forensic and clinical toxicology laboratories have historically relied upon immunoassay to determine the presence or absence of snake venom.

Proteomics techniques, employing high resolution mass spectrometry (MS) can be used to achieve selectivity and specificity for the detection of target venom proteins, therefore enabling the unequivocal confirmation of snake bites, and the originating species possible. Here we report the development of a novel methodology using bottom-up proteomics, for species identification from venom proteins. Preparation of venom proteins by reduction/alkylation and tryptic digestion was achieved before analysis by liquid chromatography mass spectrometry (LC-MS) using a Thermo Fisher QE Plus Orbitrap MS. Data were acquired using the TraceFinder™ software and Parallel Reaction Monitoring (PRM), with protein identification achieved using a Serpentes fasta subset of the UniProtKB database, processed through Proteome Discover software (version 2.2).

The established methodology permits the definitive confirmation of a snake bite, uniquely detecting and identifying nine snake venoms using LC-MS. This work reports the first case of the unequivocal identification of the Eastern Brown snake venom in a swab taken from a person who had succumbed to the snake bite. The method provides an advancement in forensic coronial toxicology and highlights the potential of proteomics assays to routine forensics investigations.

111

C18ORF25 is a novel exercise-regulated AMPK substrate regulating metabolic adaptations

Yaan Kit Ng[1], Benjamin L Parker[1], Ronnie Blazev[1], Vanessa R Haynes[1], Andrew J Kueh[2], Marco J Herold[2], Thomas E Jensen[3]

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

2. The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia

3. Department of Nutrition, Exercise and Sports, University of Copenhagen, Copenhagen, Denmark

Exercise regulates a diverse array of phosphorylation networks which are thought to promote numerous health benefits. Functionally characterising these networks hold great promise to identify new therapeutic targets for a range of diseases including type 2 diabetes, cancer and neurological disorders. We have recently shown that the uncharacterised protein, C18ORF25 is significantly phosphorylated at Ser-67 across human, mouse and rat exercise models in skeletal muscle. *In silico* analysis reveals that it is a homolog of ARKadia, a regulator of TGF-B/BMP signalling, and is conserved down to bone-jawed fish. Furthermore, machine learning to predict the upstream kinase(s) mediating this phosphorylation event revealed C18ORF25 to be a highly significant hit as an AMPK substrate. Given the well described role of AMPK in metabolic adaptations during exercise, we hypothesise C18ORF25 is a novel regulator of exercise metabolism. Here, we validate C18ORF25 as a novel exercise-regulated AMPK substrate. To characterise the functional role of C18ORF25, we generated a whole-body knockout (KO) mouse model. Our preliminary data reveal KO mice gained similar weight on a chow diet compared to WT littermates but we observed a striking increase in adiposity and subtle decrease in lean mass from 6 weeks of age. These young KO mice showed a slight glucose intolerance and a significant reduction in energy expenditure with no difference in food intake, and they are also not able to run as far following forced treadmill exercise. Furthermore, functional analysis of ex vivo isolated muscles revealed those from KO mice have reduced contractile force and fatigue quicker. To investigate the functional effects of phosphorylation, we overexpressed a wild-type, a Ser-66/67-Ala mutant (to inhibit phosphorylation) or Ser-66/67-Asp mutant (to mimic phosphorylation) in HEK293 cells and performed a proteomic analysis. We observed a surprising remodelling of the proteome induced by these mutations including significant regulation of BMP signalling. Our preliminary data suggest C18ORF25 plays a vital role in AMPK-mediated regulation of BMP-signalling to allow metabolic adaptations to exercise. My project aims to perform further proteomic and phospho-proteomic analysis on C18ORF25 KO tissue to explore upstream and downstream signalling of C18ORF25 and the effect that phosphorylation has on it. We also aim to replicate experiments in metabolically insulted (high fat diet) C18ORF25 KO mice.

1. Nelson, M., Parker, B., Burchfield, J., Hoffman, N., Needham, E., Cooke, K., . . . James, D. (2019). Phosphoproteomics reveals conserved exercise-stimulated signalling and AMPK regulation of store-operated calcium entry. The Embo Journal, 38(24), 102578.

112

Development of quantitative assay for the detection of Proteinase A in beer using mass spectrometry

Jan Degenhardt[1], Brooke A Dilmetz[1], Chris Desire[1], Clifford Young[1], Jon Meneses[2], Mark R Condina[1], Peter Hoffmann[1]

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

SCIEX
The Power of Precision

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

A long lasting and stable foam head is a key quality indicator for beer around the world. The distribution of beer with inferior foaming properties can lead to substantial financial losses and damage to the brand reputation. The determination of beer foam stability throughout the brewing process is essential for breweries as a key quality indicator. A stable foam head is achieved through the interplay of various proteins, metal ions and small molecules1. Proteinase A, an aspartic yeast proteinase, was shown to degrade foam promoting proteins and may be responsible for inferior foaming properties2. Here, we demonstrate the applicability of liquid chromatography mass spectrometry (LC-MS) methods to quantitate Proteinase A in commercial beer samples. Label-free quantitation was performed on beer samples to determine the relative amount of Proteinase A in beer and its associated peptides. A targeted LC-MS/MS method was developed for the quantification of Proteinase A in commercial beer types. The targeted assay enables breweries to assess beer foam stability and improve current quality control processes.

1. 1. C. W. Bamforth, Hg., Beer: A quality perspective, Handbook of alcoholic beverages series, Academic, Burlington, MA 2009.
2. 2. S. E. Brey, S. Costa, P. J. Rogers, J. H. Bryce, P. C. Morris, W. J. Mitchell, G. G. Stewart, Journal of the Institute of Brewing 2003, 109, 194–202.

113

Whole-cell proteomics of *Streptomyces hygroscopicus* NRRL 30439 reveals detailed temporal dynamics of secondary metabolism during submerged cell culture

Craig P Barry[1], Gert H Talbo[1], Esteban Marcellin[1]

1. Australian Institute for Bioengineering and Nanotechnology, Fairfield, QLD, Australia

Continued evolution of multidrug resistant pathogenic bacteria is sustaining a need for novel antibiotics with mechanisms distinct from those currently used. *Streptomyces* is a genus of filamentous soil dwelling bacteria with incredible genomic diversity and is a repertoire for biosynthesized antimicrobial compounds. Naturally, biosynthetic gene cluster (BGC) expression in *Streptomyces* is reserved for a secondary metabolic phase where antimicrobial compounds are synthesized as a biological warfare strategy during nutrient competition. Identifying novel BGCs and characterising their product is often constrained by the initial challenge of eliciting their expression. Here, we undertook a detail temporal proteome characterisation of *Streptomyces hygroscopicus* NRRL 30439 during submerged culture fermentation with the aim of elucidating the capabilities of its secondary metabolome.

PacBio RSII sequencing of *S. hygroscopicus* NRRL 30439 allowed us to construct a single 9.7 Mb linear genome and identify 24 biosynthetic regions using antiSMASH v5.1.0. A chemically defined media was developed to elicit a diverse nutrient stress response, which was captured by temporal intracellular proteomics. Data was acquired using a Thermo Fisher Q Exactive HF-X in DIA mode, and analysed with Spectronaut. Of the 24 predicted BGCs, 6 were found to express all their core BGC proteins where most of these proteins exhibited expression profiles coincident with nutrient depletion. Well-studied regulators of morphological differentiation (*bld* cascade) were detected in the proteome, suggesting that *S. hygroscopicus* NRRL 30439 regulates facets of secondary metabolism by selective expression of $tRNA^{leu}UUA$. This is further supported in the proteome and exemplified in cases where TTA codon-containing BGC genes exhibiting stress-induced expression, whilst adjacent genes of the same cluster exhibit an earlier onset of expression. While many of the product molecules of predicted BGCs remain unknown, our work represents a method of eliciting secondary metabolism in *Streptomyces* and identification of likely-expressed novel candidate molecules. This allows for a narrowed focus when proceeding to characterize novel antimicrobial candidates.

114

Australian BioCommons communities: engaging researchers at the national scale to understand bioinformatics challenges and deliver solutions

Tiff Nelson[1], Andrew Lonie[1], Johan Gustafsson[2], Jeff Christiansen[1]

1. Australian BioCommons, Melbourne

2. Australian BioCommons / Bioplatforms Australia, Melbourne

The Australian BioCommons integrates the tools, methods and training required to advance bioinformatics capacity for Australia's life scientists. This is achieved through a combination of integration and development across the BioCommons infrastructure partners, the provision of community supported tool and workflow sets on both command line and graphical user interface platforms, as well as a national training effort.

To achieve a fit-for-purpose outcome, we need to identify the requirements of many thousands of geographically dispersed researchers, and use this to deliver useful infrastructure. Strong user engagement is paramount to understand community needs and direct the deployment and resourcing of appropriate infrastructure to ensure maximum impact.

The BioCommons have developed a five step process to maximise interaction with the community, from initial consultation, to deployment of solutions:

1. Identify existing communities of manageable scope;

2. Analyse the community area, in consultation with its members, to understand its broad needs and challenges;

3. Communicate with the broad community, inclusive of everyone from any expertise level or any institution, to identify issues, roadblocks and solutions/suggestions through electronic surveys, shared discussion boards and virtual meetings;

4. Document and distill the challenges, and in discussion with infrastructure specialists, detail conceptual solutions supported by a subset of the community;

5. Deploy and implement solutions with testing and feedback from the community.

SCIEX
The Power of Precision

Through this engagement process, the Australian BioCommons has identified and then coordinated work to address the absence of essential infrastructure to support critical communities (e.g. those undertaking genome annotation). The method is now being applied to engage a diverse range of communities, including proteomics.

Visit https://www.biocommons.org.au/proteomics for more information or to join the conversation.

115

The cancer stem cell marker DCLK1 influences extracellular vesicle biogenesis and cargo-selection in a kinase dependent manner promoting migration and adhesion processes in gastric cancer.

Annalisa Carli[1, 2], Shoukat Sterle[2], Alin Rai[3], Fleur M Ferguson[4, 5], Nathanael S Grey[4, 5], Matthias Ernst[1, 2], David W Greening[1, 3], Michael Buchert[1, 2]

1. La Trobe University, Heidelberg, VIC, Australia

2. Cancer and Inflammation Program, Olivia Newton-John Cancer Research Institute, HEIDELBERG, VIC, Australia

3. The Baker Institute, Melbourne, VIC, Australia

4. Department of Cancer Biology, Dana-Faber Cancer Institute, Boston, MA, USA

5. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA

Doublecortin-like kinase 1 (DCLK1) is a putative cancer stem cell marker, a promising diagnostic and prognostic maker for malignant tumors and a proposed driver gene for gastric cancer. The overexpression of DCLK1 correlates with advanced disease and overall poor-prognosis. In cancer cells, DCLK1 expression has been shown to promote an epithelial-to-mesenchymal transition (EMT), loss of cell-cell adhesion, and promoting cell migration and invasion. Here, we report that DCLK1 influences extracellular vesicle (EV) biogenesis in a kinase-dependent manner and EVs originating from DCLK1 overexpressing MKN1 ($MKN1^{OE}$) gastric cancer cells can induce cell migration of parental isogenic cells ($MKN1^{PAR}$). The quantitative proteomics analysis of $MKN1^{OE}$-sEVs revealed enrichment in migratory and adhesion processes. Moreover, using a specific small molecule inhibitor of DCLK1, we reversed the observed increase in EV size and concentration, as well as kinase dependent cargo selection of proteins involved in EV biogenesis and migration and adhesion processes. Our findings highlight a specific role of DCLK1-kinase dependent cargo selection for EVs and shed new light on its role as a regulator of signaling in gastric tumorigenesis.

116

Tasmanian devil facial tumor-derived extracellular vesicles reveal mesenchymal transition markers and adhesion molecules related to metastasis

Camila Espejo[1], Richard Wilson[1], Greg Woods[1], Eduard Willms[2], Andrew Hill[2], Bruce Lyons[1]

1. University of Tasmania, Hobart, TAS, Australia

2. College of Science, Health and Engineering , La Trobe University, Melbourne, Victoria, Australia

Tasmanian devils are threatened with extinction by Devil Facial Tumor Disease (DFTD), which consists of two genetically independent transmissible cancers (DFT1 and DFT2). Both cancers typically cause death due to metastases. However, the mechanisms underpinning DFTD metastasis are not well understood. The nano-sized, membrane-enclosed extracellular vesicles (EVs) released by cancer cells have been implicated in metastasis, thus EVs may yield insights into DFTD metastasis. Here, we characterized EVs derived from cultured DFT1, DFT2, and devil fibroblast cells. The proteome of EVs was determined using data-independent acquisition mass spectrometry and an in-house spectral library of >1,500 proteins. Relative to EVs from fibroblast cells, EVs from both DFT1 and DFT2 cell lines expressed higher levels of proteins associated with cell adhesion and focal adhesion functions. Furthermore, hallmark proteins of epithelial-mesenchymal transition, which are associated with increased metastatic features in some cancers, were enriched in DFT2 EVs relative to DFT1 EVs, suggesting differential aggressiveness between the cancers and a target for novel differential diagnosis biomarkers. This first EV-based investigation of DFTD increases our understanding of the cancers' EVs and their possible involvement in the metastatic process. As EVs are found in body fluids, these results offer potential for non-invasive biomarkers for DFTD.

117

Combining correlation network and enrichment analysis to help extract biological insights from complex proteomics experiments

Jemma X Wu[1], Dana Pascovici[1], Yunqi Wu[1], Adam Walker[2], Mehdi Mirzaei[1]

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

2. Neurodegeneration Pathobiology Laboratory, Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia

Discovery proteomics experiments generate large amounts of data, but there is an unmet need for tools and workflows capable of extracting useful biological signals in complex multi-condition proteomics experiments, and particularly for proposing a small number of key targets to prioritise for follow-up experiments. We present a useful workflow for characterizing proteomics experiments that incorporate many conditions and abundance data, which incorporates the popular approach weighted gene correlation network analysis (WGCNA) and functional enrichment analysis with the PloGO2 R package. In this workflow we have extended the PloGO2 R package and made it available to Bioconductor, the open source repository of R software packages for the analysis of high-throughput omics data. The approach can use quantitative data from labelled or label-free experiments, and was designed and developed to handle multiple files stemming from data

SCIEX
The Power of Precision

partition or multiple pairwise comparisons. Enrichment analysis will identify clusters or subsets of proteins of interest, and the WGCNA network topology scores will produce a ranking of proteins within these clusters or subsets. This can naturally lead to prioritized proteins to be considered for further analysis or as candidates of interest for validation in the context of complex experiments. We demonstrate our approach and the application of the workflow on two previously published datasets. In both, the automated workflow recapitulates key insights or observations of the published papers, and provides additional suggestions for further investigation. These findings indicate that dataset analysis using WGCNA combined with the updated PloGO2 package is a powerful method to gain biological insights from complex multi-condition proteomics experiments and provide information that may guide future investigations.

119

Proteomics to understand bio-therapeutic production

Aiden Beauglehole[1], Ben Schulz[1], Esteban Marcellin[1], Susann Wudtke[1]

1. Australian Institute for Bioengineering and Nanotechnology, AIBN, Brisbane, Queensland, Australia

Bio-therapeutics are drugs or vaccines derived from living systems used to treat a seemingly limitless amount of illnesses and many diverse diseases, including cancers, diabetes, haemophilia, and arthritis. In 2019, almost 50% of new drugs approved by the Food and Drug Administration were bio-therapeutics. Of those new bio-therapeutics, 70% were made from in Chinese Hamster Ovary (CHO) cells making CHO cells the producer of over a quarter of all new drugs. Due to the heavy reliance on CHO cells, considerable amounts of research has gone into understanding how they express bio-therapeutics. Past research has focused on improving culture processes through media, bio-reactors, vector optimisation, and targeted engineering. Whilst these improvements have been highly successful, a ceiling has been hit limiting useful amounts of production improvements. Systems biology offers a way to break through the ceiling to gain larger production titres in CHO cells. Systems biology treats CHO cells as a complex interactive systems gathering results on computational and mathematical modelling/analysis from experimental data obtained from various omics techniques. By using proteomics as a part of systems biology, we want to understand why the model bio-therapeutic Factor IX is expressed in relatively small quantities compared to other bio-therapeutics.

120

Automated proteomics workflows for translational medicine: myocardial infarction and early origins of heart disease

Selvam Paramasivan[1], Roberto Barrero Gumiel[2], Paul Millls[1], Janna Morrison[3], Pawel Sadowski[4]

1. University of Queensland, Brisbane, QLD, Australia

2. Division of Research and Innovation, Queensland University of Technology, Brisbane, Queensland, Australia

3. UniSA clinical and Health Sciences, University of South Australia, Adelaide, South Australia, Australia

4. Central Analytical Research Facility, Queensland University of Technology, Brisbane, QUEENSLAND, Australia

Myocardial infraction (MI) is a leading cause of cardiovascular disease-related deaths globally. In Australia, more than 500,000 persons suffer from heart attack annually. Studying sheep (*Ovis aries*) model, where the heart structure and development closely resemble that of humans, can contribute to a better understanding of cardiac repair mechanisms and factors associated with increased risk of heart failure. We have previously introduced a high-throughput and fully automated proteomics workflows which allowed us to accelerate spectral library generation and SWATH-MS (Sequential Window Acquisition of all Theoretical Mass Spectra)-based protein quantitation in bovids. The pipeline utilizes an advanced robotic sample preparation system (PerkinElmer JANUS G3) to enable protein digestion, desalting and fractionation in a 96 format. Here we have deployed this pipeline to process sheep heart tissue samples from control and animals with failing hearts to generate highly comprehensive proteome maps of heart disease. Quantified proteins allowed us for a clear distinction between healthy and infarct tissue, and revealed pathways associated with cardiomyopathy, abnormal cardiovascular system physiology, dilated cardiomyopathy and ventricular arrhythmia. Furthermore, using bioinformatics predictions, we have correlated this next generation proteomics dataset with miRNA profiling to better understand the early origins of a heart disease.

121

Secreted midbody remnants are a class of extracellular vesicles molecularly distinct from exosomes and microparticles

Alin Rai[1, 2], David W Greening[1, 2], Rong Xu[1], Maoshan Chen[1, 3], Wittaya Suwakulsiri[1], Richard J Simpson[1]

1. La Trobe University, Melbourne, VIC, Australia

2. Baker Heart and Diabetes Institute, Melbourne, VIC, Australia

3. Myeloma Research Group, Australian Centre for Blood Diseases, Monash University/The Alfred Hospital, Melbourne, VIC, Australia

During the final stages of cell division, newly-formed daughter cells remain connected by a thin intercellular bridge containing the midbody (MB), a microtubule-rich organelle responsible for cytokinetic abscission. Following cell division the MB is asymmetrically inherited by one daughter cell where it persists as a midbody remnant (MB-R). Accumulating evidence shows MB-Rs are secreted (sMB-Rs) into the extracellular medium and engulfed by neighbouring non-sister cells. While much is known about intracellular MB-Rs, sMB-Rs are poorly understood. Here, we report the large-scale purification and biochemical characterisation of sMB-Rs released from colon cancer cells, including profiling of their proteome using mass spectrometry. We show sMB-Rs are an abundant class of membrane-encapsulated extracellular vesicle (200-600 nm) enriched in core cytokinetic proteins and molecularly distinct from exosomes and microparticles. Functional dissection of sMB-Rs demonstrated that they are engulfed by, and accumulate in, quiescent fibroblasts where they promote MAPK-signalling, cellular transformation and an invasive phenotype.

SCIEX
The Power of Precision

122

Expanding proteomic coverage of barley via SWATH-MS with multiple proteolytic enzymes to assist malting accreditation of new varieties and brewing performance

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

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

Barley is the third largest agricultural crop grown in Australia averaging over 9 million tonnes per year. Barley is germinated under controlled conditions in a process known as malting, to allow endogenous enzymes to begin the process of cellular degradation for use in the brewing industry. 30-40% of Australia's national barley crop is malting grade, generating a price premium to feed grade barley and driving the constant development of new malting barley varieties. The recognition of new malting barley varieties is a long process based on pilot and commercial scale malting taking over at least two growing seasons. Once harvested, barley crops are graded as either malting or feed grade based on total protein content, an important marker for barley breeders, growers, maltsters and brewers. A limited number of protein characteristics are regularly assessed, such as diastatic power, but limited regard is given to the overall composition of the barley proteome. Challenges in the proteomic assessment of barley are due to a large number of storage proteins, known as hordeins, representing around 40% of total barley protein content. Hordeins are largely insoluble and contain large regions of glutamine rich repeats which can reduce their detection using standard SWATH-MS workflows. The use of alternative proteolytic enzymes can be utilised to improve proteomic coverage of hordeins but at the expense the total proteomic coverage. Methods have been developed to combine DIA/SWATH-MS data across samples digested with multiple enzymes to improve the proteomic coverage of barley and malted barley. The application of miniaturised methods previously developed for proteomic assessment of different stages of the brewing process[1] have also been combined along with newly developed single seed methodologies for malting and brewing. The increased proteomic coverage of the barley and malted barley proteome has been applied to multiple malting barley varieties across malting and brewing with a goal to be implemented to improve the speed of malting accreditation of new barley varieties.

1. 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

123

Proteomic analysis of root tissue from two different varieties of rice with contrasting root architecture phenotypes

Yunqi Wu[1], Gene Hart-Smith[1], Hosseini Salekdeh[2], Somayeh Abdirad[2], Zahra Ghorbanzadeh[2], Mehbano Kazemi[2], Mehdi Mirzaei[1], Brian J. Atwell[3], Paul A. Haynes[1]

1. Department of Molecular Sciences, Macquarie University, North Ryde, NSW 2109, Australia

2. Agricultural Biotechnology Research Institute of Iran, Tehran, Iran

3. Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia

Plant roots are the primary mechanism for absorbing water and nutrients from the soil. Breeding for deeper and more branched root systems would enhance the extractive properties of roots. This involves first identifying key genes and proteins responsible for optimal root phenotypes, and subsequently altering their expression in order to determine the roles of regulatory elements in root development. By manipulating rice root architecture, we aim to maximize water and nutrient use efficiency in high-intensity, modern agricultural systems.

Two rice varieties with contrasting root architecture phenotypes - a lowland rice with shallow roots (IR64) and an upland rice with deeper roots (Azucena) - were grown under control and water deficit conditions. In order to focus of the development of new root tissues, three sections distal to the root apex were sampled: 0 – 5 mm (Z1), 5 – 10 mm (Z2) and 10 – 15 mm (Z3); three variables were therefore analysed - genotypes, water supply and developmental zones. Quantitative proteomic analysis using TMT labelling identified and quantified 7,509 proteins from Azucena and 8,011 proteins from IR64. Differentially abundant proteins were calculated from normalized TMT label peak area intensities. Hierarchical clustering and heat maps showed that the proteomics profiles of Z1 samples clustered together irrespective of genotype and conditions, and were highly distinctive compared with Z2 and Z3, which clustered closely together.

Proteins associated with glutathione metabolism were found to be increased in abundance in stress conditions in all zones examined, including several isoforms of glutathione S-transferase which were observed at greatest abundance in Z3 in Azucena. Data analysis is ongoing, including characterization of functional marker proteins to verify expected metabolic processes overrepresented in each zone: cell division in Z1, cell wall expansion and vacuolarisation in Z2, and xylogenesis in Z3. Further details of all the differentially expressed proteins identified in each of the comparisons will be presented.

124

The multiple factors driving the protein composition of bee venom

Daniela Scaccabarozzi[1], Thao Le[2, 3], Joel P.A. Gummer[2, 3], Michele Lussu[4], Lynne Milne[5], Ashley Nolan[2], Ken Dods[2], Colin Priddis[2]

1. School of Molecular and Life Sciences, Curtin University, Bentley, WA, Australia

2. ChemCentre, Bentley, WA, Australia

3. School of Science, Edith Cowan University, Joondalup, Western Australia, Australia

4. Regional Institute for floriculture (IRF), Viale Carducci, San Remo, Italy

5. School of Earth and Planetary Sciences, Curtin University, Bentley, WA, Australia

SCIEX
The Power of Precision

Publish consent withheld

SCIEX
The Power of Precision