ABSTRACT BOOK

US HUPO 2022

ONE WORLD

February 26 – March 2

Charleston, South Carolina

www.ushupo.org

Table of Contents

P10: Metabolomics and Imaging Mass Spectrometry

Parallel Session 01: Advances in Single Cell Measurements

IS01.01 Intelligent Data Acquisition for Multiplexed Single Cell Analysis

Christopher Rose, Genentech Inc.

Single cell proteomics methods often utilize a carrier proteome added at 20-500x single cell channels to enable peptide detection, selection, and quantification. Adding excess carrier proteome may increase measurement variability and decrease measurement accuracy. We introduce an advanced data acquisition approach that utilizes a carrier proteome labeled with an offset mass chemical tag (i.e., TMT-SH) that separates the carrier proteome peptide cluster from the multiplexed single cell proteomes and enables quantification of single cells in the absence of carrier proteome reporter ions. This approach was implemented with inSeqAPI, an API program that enables facile construction of real-time data analysis methods, such that sequencing of the TMT-SH labeled carrier proteome triggers quantitative MS2/SPS-MS3 on TMT labeled single cell proteomes.

The recent introduction of real-time database searches (RTS) have drastically improved the number of peptides that can be quantified during SPS-MS3 quantification of isobaric labeled samples. Here we introduce inSeqAPI - an API based Windows application that enables the construction of custom acquisition methods on Orbitrap Eclipse mass spectrometers. inSeqAPI is unique in that it enables modular method construction complete with a number of custom available real-time filters as well as '.json' method files that can be shared across instruments. Furthermore, inSeqAPI produces a full audit trail of real-time filter results for each raw file, enabling optimization and troubleshooting during method development.

We apply inSeqAPI to the analysis of single cells by implementing an offset mass triggering method at proteome scale. This approach enables a deep proteomic analysis and sensitive quantification of single cell proteomes by quantifying peptides only after they are identified. Utilizing TMT-SH as a carrier proteome removes carrier proteome reporter ions from analysis during quantitative scans. We demonstrate that this enables accurate quantification of proteins even when carrier proteome levels exceed 500x.

IS01.02 Si Wu Presentation

Si Wu, University of Oklahoma

Presentation coming soon.

OA01.01 Improved sensitivity and depth of bottom-up and top-down proteomic and PTM profiling of ng- and sub-ng-level samples

Alexander R. Ivanov (1)

(1) Northeastern University, Boston, Massachusetts, United States

Informative proteomic characterization of limited samples (e.g., small populations of rare cells, microneedle biopsies, extracellular vesicles (EVs) isolated from minute volumes of physiological fluids, or even single cells) and especially, profiling of post-translational modifications, e.g., glycosylation and phosphorylation, of such specimens have been a major

challenge because of very low abundance and high heterogeneity in biological matrices. Alterations of proteomes, glycomes, phosphoproteomes, and acetylomes may be associated with a number of pathologies, including immune, cardiovascular and oncological diseases, as well as other biological phenomena. In this study, we evaluated a combination of advanced sample preparation, ultra-low flow high-efficiency liquid phase separation coupled to mass spectrometry (MS) via a high-field asymmetric waveform ion mobility spectrometry (FAIMS) interface to evaluate the potential applicability for high sensitivity, robust and reproducible proteomic profiling of low ng- and sub-ng-level complex biological samples. We explored several electric field- and pressure-driven separation approaches, including ultra-narrow bore monolithic and porous layer open tubular (PLOT) liquid chromatography (LC) columns and capillary electrophoresis (CE) interfaced with mass spectrometry to enhance the sensitivity and depth of proteomic, glycomic, acetylomic, and phosphoproteomic profiling of several types of limited biological specimens. Both bottom-up and top-down MS-based proteomic strategies were evaluated using cultured human cells, blood isolates, and commercial digest standards. The acquired results demonstrate the potential applicability of the developed techniques in single-cell proteomic studies and analyses of minute amounts of biological and clinical samples.

Contact: Alexander Ivanov, a.ivanov@northeastern.edu

OA01.02 Infrared Photoactivation Boosts Reporter Ion Yield in Isobaric Tagging

Trenton M. Peters-Clarke (1), Kenneth W. Lee (1), Keaton L. Mertz (1), Graeme C. McAlister (2), John E.P. Syka (3), Michael S. Westphall (1), Joshua J. Coon (1)
(1) University of Wisconsin-Madison, Madison, Wisconsin, United States
(2) Thermo Fisher Scientific, San Jose, Wisconsin, United States
(3) Thermo Fisher Scientific, San Jose, California, United States

Isobaric tagging facilitates multiplexed experiments that can determine sequences and relative amounts of peptides in biological samples using tandem mass spectrometry. Limited reporter ion generation limits quantitative accuracy and precision. Further, unintended co-isolation of contaminating peaks in MS^2 experiments distorts reporter intensities. MS^3 experiments address contamination by generating reporter ions via collisional activation (HCD) of peptide product ions rather than the precursor ion. Because HCD performance is related to m/z, activation of synchronously isolated products generates suboptimal reporter intensities. We supposed that a combination of IRMPD and ion parking would maximize reporter generation from tagged peptides by eliminating the m/z dependence inherent in HCD. Peptide ISQAVHAAHAEINEAGR was labeled with 10-plex TMT reagents. Pierce TMT11plex Yeast Digest Standard was obtained. Experiments were performed on a quadrupole-Orbitrap-QLT MS system modified to include a 40 W CO_2 continuous wave laser which allowed for infrared photoactivation of ions within the QLT. Adjustments to the instrument control code allowed for broadband ion parking at reporter ion m/z (126–131) during IRMPD experiments. Data-dependent MS^2 scans were followed by back-to-back HCD and IRMPD MS^3 scans of the same MultiNotch isolation precursors, allowing direct comparisons. As reporter ions are susceptible to unintended fragmentation and scattering by high-energy collisions, we photoactivated product ions and prevented successive dissociation of generated reporter ions with ion parking, which altogether boosted reporter ion yield by up to 55%. We show that using IRMPD to generate reporter ions from ten synchronously isolated product ions results in a 2.4-fold increase in reporter intensities, significantly enhancing the sensitivity and dynamic range of quantitation via isobaric tagging. We describe the inefficiency of HCD when activating individual peptide fragments and when synchronously activating multiple precursors and demonstrate IRMPD boosts reporter generation by up to an average of 2.4-fold, without perturbing expected quantitative ratios.

Contact: Trenton Peters-Clarke, petersclarke@wisc.edu

Parallel Session 02: Automation and Large-Scale Proteomics

IS02.01 Accelerating quantitative proteomics with Zeno MS/MS and microflow chromatography
Christie Hunter, SCIEX

In both data dependent (DDA) and data independent acquisition (DIA) workflows, the ability to collect high quality MS/MS at fast acquisition rates is key to maximizing peptide and protein identifications and obtaining highly reproducible quantification. Novel Zeno trap functionality can greatly improve duty cycle in the orthogonal pulsing region of a QTOF system, providing large gains in MS/MS sensitivity. This work investigates the impact of Zeno trap MS/MS sensitivity increases on protein and peptide workflows, at the very fast LC-MS run times offered by microflow chromatography.

IS01.02 Automated Proteomic Sample Preparation: The Key Component for High Throughput and Quantitative Mass Spectrometry Analysis
Qin Fu, Smidt Heart Institute

There is a great need in high-throughput workflows for reproducibility and precision during proteomic sample preparation for mass spectrometry-based protein quantification, and large-scale LC-MS analysis. Proteomic sample preparation consists of protein denaturation, cystine reduction and alkylation, and trypsin digestion carried out in a temperature-controlled incubator, with salts and contaminants removed prior to MS analysis. In particular, the steps around the digestion of protein samples are prone to inconsistency. One route for reliable sample processing is the development and optimization of a workflow utilizing an automated liquid handling workstation. When combining robust liquid chromatography-mass spectrometry with either discovery or targeted methods, automated sample preparation facilitates increased throughput and reproducible quantitation of biomarker candidates.

OA02.01 'Think Big, Act Bigger": Opportunity and challenges of high throughput plasma proteomics
Benoit Fatou (1), Hanno Steen (1)
(1) Boston Children's Hospital, Boston, Massachusetts, United States

Blood is widely used to diagnose disease or monitor health because blood is an ideal body fluid: it is a systemic body fluid, easily and readily accessible using minimally invasive collection procedures. However, wide variations in measured blood biomarkers across populations require the need of large number of subjects at different states of the immune response for meaningful clinical proteomics studies. In addition, as multiple measurements are being made from the same samples including metabolome and transcriptome, sample amounts can be limited particularly for very young population such as newborns. Given these constraints, we have been developing robust sample-sparing high throughput sample processing methods in combination with state-of-the-art instrumentation to enable large scale plasma/serum proteomic studies using submicroliter volumes of sample. We have developed a targeted and multiplexed LC/MS method, on a Shimadzu LCMS8060 triple quadrupole mass spectrometer to monitor the human classical plasma proteome in a high throughput manner (<15 min per sample). We are focusing on those classical plasma proteins needed for a comprehensive description

of the immunophenotype of humans under a wide range of immune challenges such as infection, inflammation and/or vaccination. Our current plasma proteomics pipeline allows for the complete processing and analysis of up to 600 plasma samples within a week. This method has been thoroughly tested by application to large-scale studies within the Human Immunology Project Consortium (HIPC) in the context of newborn immune development, providing unique insights into the maturation of the immune system during the first week of life (DOI: 10.3389/fped.2020.610461i) and the IMmuno Phenotyping Assessment in a COVID-19 Cohort (IMPACC; DOI: 10.1126/sciimunol.abf3733) study, allowing us to better understand the immune determinants associated with positive and negative outcomes from COVID-19. These studies have resulted in the successful processing and analysis of >5,000 plasma samples (and counting).

Contact: Benoit Fatou, benoit.fatou@childrens.harvard.edu

OA02.02 Fully Automated Proteomic Sample Preparation for the iPSC Neurodegenerative Disease Initiative (iNDI): A Large-Scale Effort at the NIH

Luke Reilly (1), Erika Lara (1), Lirong Peng (1), Daniel Ramos (1), Caroline Pantazis (1), Julia Stadler (1), Marianita Santiana (1), Faraz Faghri (2), Mike Nalls (3), Steven Coon (4), Priyanka Narayan (1), Andrew Singleton (1), Mark Cookson (1), Michael Ward (1), Yue A. Qi (1)
(1) NIH, Bethesda, Maryland, United States
(2) Data Tecnica International, Glen Echo, Maryland
(3) Data Tecnica International, Glen Echo, Maryland, United States
(4) NIH, Bethesda, Maryland

The induced pluripotent stem cell (iPSC) Neurodegenerative Disease Initiative (iNDI) is a large-scale effort to characterize iPSC-derived neurons harboring one of 132 Alzheimer's-related variants across 73 genes using multi-omics approaches, openly providing both the lines and associated data to the scientific community. Such a large-scale and long-term project requires consistent and efficient workflows in sample preparation and data acquisition for proteomics analyses. Here, though combining liquid- and bead-handling systems, we report a fully automated proteomic sample preparation pipeline in 96-well format. This involves direct protein collection and reduction in 96-well tissue culture plates using high-percentage detergent SP3 lysis buffer, automated protein and peptide concentration assays and normalization using an Agilent Bravo liquid handler, and a customized 30-minute SP3 paramagnetic bead protocol on the Kingfisher APEX platform.

To test this pipeline, we compared three neuron differentiation protocols: overexpression of transcriptional regulator neurogenin-2 (NGN2), co-expression of NGN2 with homeobox protein EMX1, and the NGN2-EMX1co-expression strategy with CultureOne supplement. The iPSCs were differentiated into mature neurons using the three protocols (n = 6) and processed via our automated pipeline. We consistently identified more than 45,000 peptides and 7,500 protein groups from as few as 3000 cells per replicate using single-shot DIA analysis. The CVs of biological replicates were less than 15%, and more than 90% of proteins were quantified across all replications, indicating high reproducibility and low missing values. Our results indicated upregulated cortical neuron markers (e.g., BCL11A) and downregulated cholinergic neuron markers (e.g., CHAT, SCL18A3) when comparing NGN2-EXM1 to the NGN2 condition. The CultureOne condition resulted in increased synaptic markers (e.g., MAP2 and MAPT), but higher expression of cholinergic markers. Importantly, we validated these observations using single-cell RNA sequencing datasets from the same cell lines. Together, we provide an automated MS sample preparation protocol from sample collection to LC-MS/MS.

Contact: Luke Reilly, luke.reilly@nih.gov

Parallel Session 03: System Suitability in Mass Spectrometry

IS03.01 Lindsay Pino Presentation
Lindsay Pino, Talus Bio

Presentation coming soon.

IS03.02 HeLa Digest Mixed with Synthetic Standard Peptide Reference Is Used to Monitor System Suitability in a Data-Dependent Acquisition Proteomics Workflow
Leonard Collins (1), Taufika Islam Williams (1), Jaclyn Kalmar (1), Michael Bereman (1), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States

The goal of this study was to design a procedure to define system suitability for discovery proteomics that uses mathematically derived thresholds to judge LC-MS system performance instead of relying on arbitrary limits. We currently evaluate commercially available HeLa digest using data-dependent acquisition (DDA) and bovine serum albumin digest using parallel reaction monitoring (PRM). We set an arbitrary benchmark for non-specific, total system performance by noting the number of protein identifications (IDs) obtained with Proteome Discoverer 2.4 (PD) for HeLa injections and track chromatography and mass spectrometer output changes with BSA injections. There is a need for a single mechanism specific to DDA that provides relevant system suitability information. We designed an experiment to reveal the point at which deteriorating system performance would unacceptably impact final results generated by PD. Standard HeLa digest was combined with Promega 6 x 5 Peptide Reference Mix and serially diluted to mimic a gradual instrument failure. Data were collected using DDA-only experiments, and reference peptide peak intensities were correlated to the number of peptide spectral matches (PSMs) and protein identifications found by PD. A logarithmic decay relationship of the type $y=A(1-e^{-kx})$ explained decreasing PSMs and IDs with decreasing peak intensities. Reference peptide peak intensity thresholds were calculated from these relationships and new quality control plots were created that visually identify failure points. The 6 x 5 reference peptides were also used to monitor spectrometer mass accuracy, signal dynamic range, retention time stability, and chromatographic peak widths. Combining HeLa digest and Promega 6 x 5 peptides provided a mechanism to assess chromatographic and mass spectrometric performance and predict expected PD outputs in a single DDA experiment.

Contact: Leonard Collins, lbcolli2@ncsu.edu

OA03.01 System suitability utilizing Skyline-Daily and AutoQC Loader in Metabolomics
Yu-Chun Chiu (1), Jeff Enders (1), Whitney Stutts (1), Allison Stewart (1)
(1) North Carolina State University, Raleigh, North Carolina, United States

Skyline introduced metabolomics and small molecules features into the software in 2017 and since then has shown to be a powerful tool for targeted metabolomics. However, system suitability monitoring in metabolomics and small molecules application by Skyline is relatively limited compared to proteomics applications. Therefore, we would like to share a workflow utilizing Skyline and AutoQC for monitoring system suitability for a LC-MS/MS metabolomics platform. One challenging aspect of system suitability in metabolomics applications is that the analytical column and the mobile phase system change so frequently, especially in laboratory settings where multiple classes of compounds are analyzed. A dual LC column setup

is used in this scenario as one column is designated to only analyze the specific standard mixture and the other column is for sample analysis. This setup eliminates the variables introduced from samples and makes system suitability in metabolomics more consistent within the laboratory. The raw file of system suitability mixtures is saved in a folder that is automatically scraped by AutoQC and uploaded to an online portal called Panorama. Using this web portal and an accompanying Skyline document, the user can check attributes such as mass accuracy, retention time, and peak area to evaluate if the sample analysis can be continued in this system. AutoQC can also be used as a tracking tool to monitor the system and column performance over time, especially when multiple sets of experiments are completed using the same column. Though mainly demonstrated previously for proteomics applications, Skyline and AutoQC together also serve as a powerful open-source tool to perform system suitability of metabolomics LC-MS/MS applications.

Contact: Yu-Chun Chiu, ychiu2@ncsu.edu

OA03.02 Global multi-site assessment of LC-MS proteomics precision and accuracy with high-throughput analysis

Oleksandr Boychenko (1), Runsheng Zheng (1), Tabiwang Arrey (2), Amirmansoor Hakimi (3), Christopher Pynn (1), Alec Valenta (4)
(1) Thermo Fisher Scientific, Germering, Germany
(2) Thermo Fisher Scientific, Bremen, Germany
(3) Thermo Fisher Scientific, San Jose, California, United States
(4) Thermo Fisher Scientific, Michigan, Michigan, United States

Throughput limitations and often preclude the adoption of nanoLC-MS methods for translational proteomics applications such as biomarker validation because the large sample cohorts need to be analyzed to reveal changes that stand out of biological, analytical, and sample preparation variation. The throughput (the number of samples per day) and MS utilization (the ratio of peptide elution window to total cycle time) can be boosted by utilizing LC instrumentation with the extended flow and pressure capabilities and minimal delay volumes. Optimized LC-MS methods operated at flow rates in the range from 1.3 to 0.3 µL/min were created to maximize MS utilization (calculated as the ratio of peptide elution window to cycle time). The MS utilization gradually increased from 68% for the shortest method to 95% for the longest 60 min method. The length of the methods is also linked to the increased FWHM from ca. 3 sec for 8 min method to ca. 10 sec for 60 min method. To comprehensively estimate the analytical variability were compared the quantification and identification results collect (i) on the same low-flow LCMS instrument using the same separation column; (ii) on the same instrument using three separation columns of the same dimensions; (iii) on different instruments, with different separation and trap columns, located at three different laboratories in the US and Europe. The label-free quantification showed that 72% of identified proteins have less than 25% variation of abundance for 100 continuous injections. The obtained results allow setting realistic estimates for abundance variation that are required to confidently detect targets above the multi-site analytical variation.

Contact: Alexander Boychenko, oleksandr.boychenko@thermofisher.com

Parallel Session 04: Functional Glycomics

IS04.01 Heparan Sulfate 6-O-endosulfatases in HNSCC and Other Malignancies
Radoslav Goldman, Georgetown University

Post synthetic editing of heparan 6-O-sulfation by the human endosulfatases SULF1 and SULF2 regulates extracellular matrix remodeling, growth factor signaling, or chemokine distribution. These pathways modulate outcomes of cancer diseases and patient responses to therapy. In this study, we carried out a translational study of multiple cancer diseases which confirms the impact of the 6-O-endosulfatases on the progression of head and neck squamous cell carcinoma (HNSCC). We optimized expression of the SULF enzymes in mammalian cell lines and began characterization of their structure, activities, and function in tumor biology. We expect that the advances will facilitate detection and targeting of the enzymes in HNSCC and other malignancies.

IS04.02 Biomarkers of hepatocellular carcinoma- from tissue to blood
Anand Mehta, Medical University of South Carolina

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

Methods: Formalin-fixed human liver tissue from patients with healthy livers, cirrhotic livers, or hepatocellular carcinoma (HCC) were evaluated using N-glycan MALDI imaging mass spectrometry. In total, 239 HCC tissue samples, many which were genetically subtyped, and 145 control tissue samples were analyzed. Subsequently, serum glycoproteomics was performed and proteins identified as containing the same glycans observed in HCC tissue were further examined in five independent and cohorts consisting of 1,071 patients, 492 with liver cirrhosis and 472 with hepatocellular carcinoma, including 270 with early-stage HCC, and 51 patients treated for the HCC, either by transplant or resection.

Results: Through tissue based glycan imaging, increased levels of fucosylation was observed in 96% of the HCC tissues examined. Glycan alterations in tissue correlated with the underlying genetic subtype of the cancer. The glycan most often observed in the tissue of aggressive HCC was a tetra-antennary glycan with one to three fucose residues. Using a recombinant AAL lectin with increased affinity towards fucosylated and branched glycan, we identified low molecular weight kininogen (LMWK) as a serum protein that contains tetra-antennary glycan with one to three fucose residues. A plate based lectin ELISA assay was used to test the performance of fucosylated LMWK as a biomarker of HCC. In five independent studies, a fucosylated LMWK based biomarker algorithm had a median AUC of 0.9575 in the detection of all HCC and 0.945 for early stage HCC, far exceeding the performance obtained currently used markers.

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

OA04.01 Lupus Nephritis-Associated Alterations in Serum and Urine N-Glycan Profiles
Calvin Blaschke (1), Bryan Weselman (1), Tamara Nowling (1), Richard Drake (1)
(1) Medical University of South Carolina, Charleston, South Carolina, United States

Detecting changes in the N-glycan profiles of biofluids as a reflection of disease state can potentially identify new biomarker candidates. The goal of this study was to detect alterations in the serum and urine N-glycan profiles of lupus nephritis (LN) patients, as well as differences between responders and non-responders of LN treatment. Forty donor matched urine and serum samples from LN and non-autoimmune control subjects were used. Urine samples were concentrated, and buffer exchanged with molecular weight centrifuge filters. An amine-reactive slide was used to immobilize serum and urine glycoproteins from samples spotted on to the slide. The samples were delipidated and desalted with washes of Carnoy's solution and water. Peptide N-glycosidase F was sprayed on the samples to cleave the N-glycans from the captured glycoproteins, and a layer of matrix is applied. The released N-glycans were detected across the samples by MALDI-QTOF imaging mass spectrometry (IMS). The intensities of the detected N-glycans and structural classes, such as fucosylation, sialylation, and sulfation, were compared using SCiLS Lab software. Linear regression modelling was used for detection of significant differences between sample groups. This recently developed workflow allowed for the detection of 70 N-glycans in serum and 98 N-glycans in urine. Urine samples from LN patients displayed significant glycan differences from non-LN individuals. Specifically, high-mannose glycans were lower in LN patients, while a triantennary sialylated N-glycan was more abundant. Several of these alterations were also detected in the serum samples. While there were no differences in the intensities of the N-glycans across the responder and non-responders to LN treatment, there were a series of hybrid N-glycans that were present in nearly all responders but largely absent in the non-responders. Overall, serum and urine N-glycan profiling revealed significant differences between LN patients and non-autoimmune patients, as well as between treatment responders and non-responders.

Contact: Calvin Blaschke, blaschkc@musc.edu

OA04.02 Developing isotopically labeled calibrants for quantification analysis of heparan sulfate from biological samples

Zhangjie Wang (1), Katelyn Arnold (1), Yongmei Xu (1), Jian Liu (1)
(1) University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States

Heparan sulfate (HS) is a sulfated polysaccharide that displays essential physiological functions. The amount of HS and its structural information are critically important for investigating the functions of HS in vivo. Disaccharide compositional analysis is the most widely used technique for the conformation of HS structure and total content determination as the analysis is capable of handling saccharide sequence heterogeneity. In addition to eight common disaccharide compositions, the 3-O-sulfated glucosamine in HS is also a key saccharide unit for the biological activities of HS. Here, two types of ^{13}C-labeled calibrant standards were prepared using chemosynthetic method: eight ^{13}C-labeled disaccharide calibrants for the disaccharide composition analysis, five ^{13}C-labeled 3-O-sulfated oligosaccharides for 3-O-sulfation analysis of HS. We performed a LC-MS/MS-based method to conduct disaccharide quantitative analysis of HS from biological sources including fluids, tissues, and tissue slides. Using the ^{13}C-labeled 3-O-sulfated oligosaccharides as internal standard, we determined the levels of 3-O-sulfation in the pharmaceutical heparin, analyzed the clearance rate of enoxaparin from a mouse model, and probed and quantified the specific 3-O-sulfated tetrasaccharides in the Alzheimer's disease patients. The crucial innovation in this method is to utilize a set of eight ^{13}C-labeled disaccharide calibrants and oligosaccharides as internal standards. The combination of both disaccharide calibrants and 3-O-sulfated oligosaccharide calibrants provides the sensitivity and throughput to study dynamic changes of heparan sulfate in a wide array of biological samples

Contact: Zhangjie Wang, zhangjie@ad.unc.edu

Parallel Session 05: Biomarkers and Precision Medicine

IS05.01 Tony Hollingsworth Presentation
Tony Hollingsworth, University of Nebraska Medical Center

Presentation coming soon.

IS05.02 Glycomics and Glycoproteomics Enabling the identification of Glycan and Glycopeptide isomers as Biomarkers for Cancer and Neurodegenerative Diseases
Yehia Mechref, Texas Tech University

Cancers and neurodegenerative diseases have an essential impact on human lives with high mortality and poor prognosis in the early stage. Several biomarkers such as CA125 and tau have been utilized to improve the prognosis of these diseases. However, the limited prediction accuracy is demanding to explore more reliable markers. Glycosylation is one of the most common post-translational modifications in mammalian cells and influences a variety of essential biological functions. Aberrant glycosylation and its isomeric distribution have been associated with many neurodegenerative diseases such as neurological disorders, Alzheimer's disease, Parkinson's disease, and various cancers. However, the glycomics and glycoproteomics methods for the isomeric glycans and glycopeptides for clinical samples remain a great challenge due to the complex microheterogeneity of glycoproteins. Over decades, different methods have been developed to facilitate the identification and quantitation of glycan and glycopeptide isomers. Coupling to a variety of separation techniques, mass spectrometry has been considered the most powerful tool in isomeric glycomics and glycoproteomics due to its high resolution, high sensitivity, and the ability to acquire sufficient structural information. Additionally, multiple sample preparation strategies and automated software were developed to improve the identification and quantitation accuracy and throughput for large clinical cohorts. During the past ten years, we have developed efficient methods for both glycomics and glycoproteomics isomeric analysis using C18-, porous graphitized carbon (PGC)-, mesoporous graphitized carbon (MGC)-, or micropillar array column (μPAC)-LC-MS/MS under different acquisition techniques, combined with several reliable sample derivatization and multiplexing strategies. Together, these advanced technologies and approaches enhance the pace of isomeric glycan and glycopeptide biomarker discovery. Here, we review recent glycomics and glycoproteomics methods for better isomeric characterization and their applications in cancers and neurodegenerative diseases.

OA05.01 Urinary Extracellular Vesicle Phosphoproteome Profiling by DIA Mass Spectrometry for Renal Cell Carcinoma Subtype Differentiation
Marco Hadisurya (1), Zheng-Chi Lee (2), Anton Iliuk (3), Ronald Boris (4), W. Andy Tao (1)
(1) Purdue University, West Lafayette, Indiana, United States
(2) West Lafayette Junior/Senior Highschool, West Lafayette, Indiana, United States
(3) Tymora Analytical Operations, West Lafayette, Indiana, United States
(4) Indiana University School of Medicine, West Lafayette, Indiana, United States

Extracellular Vesicles (EVs) are promising sources to develop disease phosphoprotein markers, based on the knowledge that phosphorylation events are directly responsible for many signaling pathways at the cellular level. This study focused on optimizing and applying robust data-independent acquisition using mass spectrometry for phosphoproteomic EVs profiling that could differentiate renal cell cancer (RCC) subtypes in a non-invasive way. We compared several methods for DIA

quantification, such as using the gas-phase fractionated library, direct DIA, forbidden zones, and several windowing schemes. After finding the optimized DIA mass spectrometry method, we applied the strategy to identify and quantify urinary EV phosphoproteome from 57 individuals representing four different urine samples from low-grade RCC, high-grade RCC, chronic kidney disease (CKD), and healthy control (HC) individuals. For these samples, we utilized the EVtrap capture method for EV enrichment and PolyMAC for phosphopeptides enrichment. In addition, we will use a bigger sample cohort with the help of automation for EV isolation to further substantiate the results. After the optimization, we quantified around 1,400 phosphoproteins. Several RNA-related pathways, such as RNA splicing, mRNA processing, and RNA binding, were significantly upregulated in CKD. Meanwhile, several common cancer-related pathways are upregulated in high-grade RCC only, but not in CKD and low-grade RCC. Overall, based on our urinary EV phosphoproteins, CKD, low-grade RCC, and high-grade RCC have unique upregulated phosphoproteins, signaling pathways, upstream kinases, and transcription factors. These results show that EV phosphoproteome quantification utilizing our optimized EV isolation, phosphopeptides enrichment, and DIA method could potentially be further developed for RCC subtype differentiation to benefit many newly diagnosed RCC patients.

Contact: Marco Hadisurya, mhadisur@purdue.edu

OA05.02 Quantification of Brain-Derived CSF Protein Biomarkers by Targeted Mass Spectrometry to Stage Alzheimer's Disease Progression

Caroline M. Watson (1), Lingyan Ping (1), E. Kathleen Carter (1), James J. Lah (1), Allan I. Levey (1), Blaine R. Robert Roberts (1), Nicholas T. Seyfried (1)
(1) Emory University School of Medicine, Atlanta, Georgia, United States

Alzheimer's disease (AD) is the most common form of dementia, with cerebrospinal fluid (CSF) amyloid-beta42 (Aβ), Tau, and phosphorylated Tau providing the most sensitive and specific biomarkers for diagnosis. However, these diagnostic biomarkers are not dynamic throughout disease progression, and do not reflect the complex changes in AD brain beyond plaque and tangle pathology. We recently developed an integrative brain-CSF proteomics platform and discovered CSF proteomic changes reflective of diverse brain pathophysiologies and disease progression. We report a sensitive, quantitative, and scalable targeted proteomics assay for novel biomarkers. We analyzed CSF from two separate Emory Goizueta AD Research Center (ADRC) cohorts using targeted mass spectrometry. As quality controls (QCs), we pooled CSF from controls having normal Aβ and Tau levels, and AD samples having low Aβ and high Tau levels. QCs were processed and analyzed identically to the individual CSF samples. Collectively, 280 controls (cognitively normal), 131 asymptomatic (AsymAD) and 282 symptomatic AD, and 68 QC samples were reduced, alkylated, denatured, and enzymatically digested with endopeptidase LysC and trypsin. Samples were analyzed as a single replicate over approximately 11 days using a 1290 Infinity II system (Agilent) coupled with a TSQ Altis Triple Quadrupole mass spectrometer (Thermo Fisher Scientific). Isotopically labeled peptide standards were added for relative quantification by reporting the area ratios for each targeted peptide in Skyline. We reproducibly detected six proteins previously identified as brain-based CSF AD biomarkers, representing a diverse range of cell types and pathophysiology. The protein targets are consistent with previous CSF proteomic studies in their direction of change with CHI3L1, SPP1, YWHAZ, and SMOC1 increased, and NPTX2 and VGF decreased in AD. We also detected changes in protein abundance for SMOC1 and other targets from control to AsymAD and AD states suggesting these proteins could serve as useful biomarkers for staging AD progression.

Contact: Caroline Watson, caroline.watson@emory.edu

Parallel Session 06: Proteoform Biology

IS06.01 Toxic and Protective Proteoforms of SOD1 Inspire a Novel ALS Therapeutic
Jeff Agar, Northeastern University

Affinity capture top-down MS was used to identify prevalent SOD1 proteoforms in ALS patients and controls. The degree of toxicity or protection and their structural correlates were assessed using neurotoxicological assays and global- and peptide-based HDX-MS. These results indicate that oxidative modification of cysteine was toxic, that dissociation of the SOD1 dimer initiates toxic aggregation, and that biology prevents this toxic oxidation by protective S-thiolation. Together these inspired the development of a novel class of pharmacological chaperones, cyclic thiosulfinates, which employ S-thiolation-mediated crosslinking. Cyclic thiosulfinates have 1) drug-like properties, 2) blood-brain-barrier penetration as measured by a novel top-down pharmacodynamics assay, and 3) stabilize proteins by unprecedented amounts.

IS06.02 New Frontiers in Proteomics – Proteoforms, Proteoform Families, and the Human Proteoform Project
Lloyd Smith, University of Wisconsin-Madison

Proteins are the primary effectors of function in biology, and thus complete knowledge of their structure and behavior is needed to decipher function. However the richness of protein structure and function goes far beyond the linear amino acid sequence dictated by the genetic code. Multigene families, alternative splicing, coding polymorphisms, and post-translational modifications, work together to create a rich variety of proteoforms, whose chemical diversity is the foundation of the biological complexes and networks that control biology. The term “proteoforms” refers to the specific molecular forms in which proteins are present in biological systems; only direct analysis of the proteoforms themselves can reveal their structures, dynamics, and localizations in biological systems.

Remarkably, the current paradigm of proteomics research, “bottom-up” proteomics, is unable to identify proteoforms – rather, proteins are enzymatically digested into peptides, which are then identified, serving as surrogates for the likely presence of their parent proteins in the sample. This strategy destroys the information as to what form of the protein the peptide represents, and thus the critical information needed to identify proteoforms is lost. The entire field of Biology is thus attempting to understand life in the absence of the ability to understand the molecules that define life. This limitation of todays technology provides a “grand challenge” to the scientific community, to devise new strategies and approaches that are able to comprehensively and quantitatively reveal the full breadth of the proteome at the proteoform level. Developing the technology to accomplish this, building a comprehensive atlas of proteoforms present in human systems, and eventually deciphering the functional roles they play in normal and disease biology, comprise central elements in the quest to understand human biology.

OA06.01 Uncovering Dehydroamino Acids within the Capsid and Matrix HIV-1 Proteins
Rachel Miller (1), Rachel Knoener (1), Bayleigh Benner (1), Brian Frey (1), Mark Scalf (1), Michael Shortreed (1), Nathan Sherer (1), Lloyd Smith (1)
(1) University of Wisconsin-Madison, Madison, Wisconsin, United States

Human immunodeficiency virus type 1 (HIV-1), despite being heavily studied, remains a leading cause of death worldwide. The development of new and effective therapeutics is limited by an incomplete understanding of the specific mechanisms of

viral infectivity. Through in-depth proteomic analysis of HIV-1 virions, we discovered the novel post-translational modification of highly conserved serine, cysteine, and threonine residues within the viral Matrix and Capsid proteins to the dehydroamino acids, dehydroalanine (DHA) and dehydrobutyrine (DHB). Dehydroamino acids are rare, understudied, and until now have not been observed in HIV viral proteins. We identified and confirmed, via chemical derivatization with glutathione, the presence of nine dehydroamino acids: two DHA residues within the Matrix protein, as well as three DHA and four DHB residues within the Capsid protein. We hypothesize these residues are important in viral particle maturation and could provide valuable insight into HIV infectivity mechanisms. Dehydroamino acids contain an electrophilic alkene moiety which can form covalent inter- or intramolecular crosslinks with the nucleophilic side chains of lysine, histidine, and cysteine residues. The ability to form these crosslinks may cause dehydroamino acids to impact HIV-1 function. Existing literature was investigated to determine the impact of mutagenesis, at the sites of confirmed dehydroamino acid modifications, on virion production or viral infectivity. For eight of the nine confirmed dehydroamino acid sites, mutagenesis had a detrimental effect on virion production or infectivity. Preliminary top-down proteomics of HIV-1 virions revealed 11 Capsid proteoforms containing DHA and DHB residues. Going forward, we will investigate the origin of these dehydroamino acids within the HIV proteome, further characterize DHA and DHB containing Matrix and Capsid proteoforms, while exploring their functional roles.

Contact: Rachel Miller, rmmiller22@wisc.edu

OA06.02 Application of Top-Down Thermal Proteome Profiling to the Analysis of Staurosporine Treated HeLa Proteome Stability

Kellye Cupp-Sutton (1), Yanting Guo (1), Trishika Chowdhury (1), Ji Kang (1), Eric Moore (1), Si Wu (1)
(1) University of Oklahoma, Norman, Oklahoma, United States

Proteoforms produced as a result of post-translational modifications (PTMs), such as phosphorylation, are critical for cellular processes. In eukaryotic cells, phosphorylated proteins make up a substantial portion of the proteome with many proteins demonstrating multiple functional phosphorylated sites. These proteoforms often demonstrate unique functionality in biological systems and the high throughput study of the functionality of these proteoforms in their native or native-like environment is valuable in the study of biological pathways and disease. We developed a high throughput top-down thermal proteome profiling (TD-TPP) method for the direct determination of proteoform thermal stability and examination of proteoform functionality. We have applied this platform for the thermal profiling of the HeLa proteome and, furthermore, analyzed the difference in stability of differentially phosphorylated/dephosphorylated HeLa proteoforms produced in the presence of the broad-spectrum kinase inhibitor Staurosporine. HeLa cell cultures were incubated with 5 uM Staurosporine (or DMSO vehicle) for 6 hours and were subsequently lysed. Lysates were aliquoted, heated (37-67°C), and centrifuged to remove the aggregate protein. Label-free quantitative UHPLC-MS/MS analysis was conducted on the remaining soluble portions. We deconvoluted and examined the thermal stability of more than 300 HeLa proteoforms with a focus on the effect of PTMs on thermal stability. For example, we identified the human Peptidyl-prolyl cis-trans isomerase protein and observed increased thermal stability of the biologically active N-terminal acetylated proteoform when compared to the unmodified proteoform. Furthermore, we observed 71 proteoforms that were differentially expressed between Staurosporine treated and DMSO vehicle treated HeLa cell cultures. These proteoforms were analyzed for shifts in thermal stability using single temperature thermal analysis. Overall, we have applied our TD-TPP for the analysis of intact HeLa proteoforms to observe the effect of PTMs on proteoform stability. Furthermore, we observed the effect differential phosphorylation on thermal stability of HeLa proteins via treatment with Staurosporine.

Parallel Session 07: Glycoproteomics: New Applications

IS07.01 New Technologies for Exploiting the Human Glycoproteome for Personalized Medicine

Rebekah Gundry, University of Nebraska Medical Center

Cell surface glycoproteins and glycans play critical roles in a range of physiological functions and disease processes, are valuable drug targets, and may be exploited as biomarkers for precision medicine. Despite their biological relevance and utility, glycoproteins and glycans are often understudied largely due to technical challenges. This presentation will describe CellSurfer and glyPAQ, new analytical platforms that enable rapid identification and quantification of cell surface glycoproteins and glycans from small sample sizes. The application of these new methodologies to address outstanding questions in cardiac physiology and disease, with an emphasis on precision medicine, will be described. CellSurfer enables routine discovery of cell surface N-glycoproteins from samples with limited availability (100-1000 µg or 300,000 – 10 million cells) with >80% specificity. Innovative bioinformatic tools expedite discovery, analysis, annotation, and candidate prioritization for downstream validation. We applied CellSurfer to cells isolated from cardiac tissue to develop the first cell type-specific maps of the cell surface N-glycoproteome of adult human cardiomyocytes from five myocardial chambers. Primary explanted cardiac fibroblasts, cardiac microvascular endothelial and coronary artery smooth muscle cells were also profiled. New cardiac cell-type specific markers emerged and novel insights into cardiac fibroblast surfaceome dynamics due to extended culturing were revealed. glyPAQ enables standardized processing of biological samples for quantitative profiling of native, reduced N- and O-glycan structures by mass spectrometry. glyPAQ is suitable for the preparation of a broad range of sample complexities, including monoclonal antibodies, cells, tissues, serum, plasma, and urine. We applied glyPAQ to human serum and tissue to inform new precision medicine strategies for patients at increased risk for cardiovascular disease due to rheumatoid arthritis and generate new insights into the impact of COVID-19 infection on the human heart.

IS07.02 A Role for Glycoproteomic Analyses in Understanding Virus-Receptor Interactions

Lance Wells, University of Georgia, Complex Carbohydrate Research Center

Viral spike proteins are in general heavily glycosylated proteins that bind host receptors to facilitate host cell invasion. These spike proteins are often the primary immunogens utilized to develop neutralizing antibodies and vaccines. Furthermore, the host cell surface receptors themselves are also glycoproteins. Thus, a molecular level understanding of the glycosylation of the spike protein protein of viruses and of the cell surface receptors of hosts are key to understanding initial steps of the infection cycle as well as developing therapeutics. In this talk, we will primarily focus on the SARS-CoV-2 Spike protein and the human angiotensin converting enzyme 2 (ACE2) receptor but will also refer to other spike glycoproteins such as that found on HIV-1.

The SARS-CoV-2 virus is responsible for the COVID-19 pandemic that has ravaged the world population for the last 2+ years. This betacoronavirus utilizes a heavy glycosylated trimer spike protein to bind to the ACE2 glycoprotein to facilitate host cell entry. We utilized a glycomics-informed glycoproteomic approach to determine site-specific microheterogeneity at all 22 sites of N-linked glycosylation for a stabilized recombinant trimer Spike mimetic immunogen and for a soluble version of human ACE2. When combined with bioinformatics and molecular dynamic simulations, our results illuminate roles for glycans in masking viral epitopes as well as directly participating in modulating the viral spike-host receptor interaction.

OA07.01 Characterizing Aberrant IgG N-glycans During Acute and Post-treatment Lyme Disease

Ben Haslund-Gourley (1), Mary Ann Comunale (1), Anand Mehta (2), Stéphane Grauzam (2)
(1) Drexel University College of Medicine, Philadelphia, Pennsylvania, United States
(2) GlycoPath, Charleston, South Carolina, United States

There are an estimated 476,000 cases of Lyme disease (LD) in the USA each year. Acute LD is confirmed through a two-tier serological assay using Borrelia burgdorferi (Bb) antigens. This testing system has a low sensitivity (46%) and LD patients remain antibody reactive which negates the ability to differentiate between current and past infections serologically. When LD is not treated within the acute phase, the spirochete can disseminate into synovial, cardiac, and neuronal tissue resulting in long-term disability and possibly death. A subset of patients suffer from Post-Treatment Lyme Disease Syndrome (PTLDS) which is poorly elucidated due to a lack of diagnostic tools. Glycosylation, a known regulator of antibody response, remains uninvestigated in LD. We hypothesize that characterizing the N-glycans of IgG can improve upon the current two-tier assay and offer insight into mechanisms of Bb immunological evasion. HPLC analysis of acute LD IgG N-glycans reveals statistically significant decreased GlcNAcylation and increased galactosylation and sialylation. This result was unexpected; many inflammatory diseases induce the opposite trend of IgG N-glycans during acute IgG responses promoting pathogen clearance. The N-glycans marking the acute phase of infection are slow to return to baseline levels over 70 days post-antibiotic treatment. To further interrogate our findings, total IgG from serum was captured via an antibody array and analyzed using a novel high-throughput MALDI-FT-ICR method under development. This parallel analysis of total IgG via MS aligned with the trends and significant alterations observed in HPLC in a fraction of the time previously required and is a promising clinical diagnostic technique. This report of altered IgG N-glycans in LD is the first in the field. Future experiments will examine the ability of IgG glycovariants to improve testing sensitivity and as a biomarker for differentiating past infections and diagnosing PTLDS using the high-throughput MALDI-FT-ICR and Lectin ELISA.

Contact: Ben Haslund-Gourley, bsh62@drexel.edu

OA07.02 Elucidating the Structural Variations and O-Glycoform Heterogeneity of SARS-CoV-2 S-RBD Variants by Top-Down Mass Spectrometry

David Roberts (1), Morgan Mann (1), Allan Brasier (1), Song Jin (1), Ying Ge (1)
(1) University of Wisconsin-Madison, Madison, Wisconsin, United States

The novel zoonotic 2019 coronavirus disease (COVID-19) global pandemic has led to over 5 million deaths as of November 2021. SARS-CoV-2 utilizes an extensively glycosylated spike (S)-protein that protrudes from the viral surface to bind receptor angiotensin-converting enzyme 2 (ACE2) for cell entry. The S glycoprotein initiates pathogenesis and is the main target for therapeutic design. S-protein glycans alter viral function and infectivity, and enormous attention has been recently focused on deciphering these complex glycan structures. However, the structures and glycan heterogeneity of the S protein regional-binding domain (S-RBD) and ACE2 O-glycans remain cryptic because of the immense challenges in intact O-glycoform analysis by conventional “bottom-up” glycoproteomic approaches, which cannot reveal their structural diversity. Here we report a multipronged top-down mass spectrometry (MS) approach for the structural elucidation of intact S-RBD/ACE2 O-glycan proteoforms for the first time. This hybrid top-down MS approach is performed by combining trapped ion mobility spectrometry (TIMS), using a Bruker timsTOF Pro, to reveal and characterize native structural heterogeneity, with ultrahigh-resolution Fourier transform ion cyclotron resonance (FTICR)-MS, using a 12T solariX FTICR, to elucidate the intact O-glycoforms. We removed the N-glycans from the S-RBD/ACE2 using PNGase F treatment to minimize the interference posed by N-glycan heterogeneity. We next used TIMS to separate and analyze the various O-glycoforms. This

native TIMS-MS method will be used to reveal the conformational heterogeneity of S-RBD/ACE2 O-glycan structures and probe the influence of various O-glycoforms on TIMS collision cross section ($^{TIMS}CCS_{N2}$). A 12T FTICR-MS equipped with collisionally activated dissociation (CAD) and electron capture dissociation (ECD) fragmentation modes will be used to isotopically resolve the intact O-glycoforms for MS/MS characterization. Finally, we will further combine this hybrid top-down MS strategy with functional studies to reveal new insights into the structure-function roles of S-RBD/ACE2 O-glycoforms in viral interactions.

Contact: David Roberts, dsroberts@wisc.edu

Parallel Session 08: Innovation in Mass Spec Imaging Omics

IS08.01 Ambient Mass Spectrometry Imaging Omics Using Nanospray Desorption Electrospray Ionization (nano-DESI) Mass Spectrometry

Julia Laskin, Purdue University

Mass spectrometry imaging (MSI) is a powerful technique for studying the localization of lipids, metabolites, and proteins in biological samples. We have developed nanospray desorption electrospray ionization (nano-DESI), an ambient ionization technique that relies on localized liquid extraction of analyte molecules from the sample into a liquid bridge formed between two glass capillaries. The extracted analytes are transferred to a mass spectrometer inlet and ionized by electrospray ionization. Nano-DESI enables quantitative imaging of biomolecules in fully hydrated samples with minimal or no sample pre-treatment. Recent developments in the nano-DESI MSI instrumentation have enabled quantitative imaging of lipids, metabolites, and proteins in tissues with high sensitivity and spatial resolution down to 8-10 microns using finely pulled capillaries. Furthermore, we have developed a microfluidic nano-DESI probe, which greatly simplifies the experimental setup and demonstrates similar performance to the capillary probe. Correlative omics imaging of small biomolecules and proteins provides unique insights into biochemical pathways in biological systems.

IS08.02 Highly multiplexed, multiomic and multimodal tissue imaging using novel photocleavable mass-tags

Mark Lim, AmberGen, Inc.

While the advent of mass spectrometric imaging (MSI) has extended the metabolomic/proteomic capabilities of mass spectrometry to the spatial dimension, it is generally limited to untargeted analysis of small molecules and proteolytic protein fragments. We have developed a novel spatialomic approach based on photocleavable mass-tags (PC-MTs) for facile labeling of probes including antibodies, lectins and nucleic acids which enables highly multiplexed MSI of targeted macromolecules in tissues. This approach, termed MSI-IHC, significantly exceeds the multiplexity of both fluorescence and previous cleavable mass-tag based methods. In addition, when combined with direct label-free MSI of small biomolecules, it provides an integrated workflow to study the spatial distribution and interaction of small molecules and larger macromolecules, such as intact proteins, in a single tissue specimen. High-plex MSI-IHC has been achieved on brain, tonsil and cancer tissues using a variety of probe classes, reflecting the known molecular composition, anatomy and pathology of the targeted biomarkers. Novel dual-labeled fluorescent-PC-MT-probes extend the utility of this new approach to multimodal imaging. Overall, MSI-IHC holds significant promise for use in the fields of tissue pathology, tissue diagnostics, therapeutics and precision medicine as well as in research aimed at understanding the mechanisms of human disease.

OA08.01 Combined sialyl-Lewis tumor antigen immunohistochemistry, chemical labeling and N-glycan imaging mass spectrometry for targeted tissue glycoproteomics

Richard Drake (1), Xiaowei Lu (1), Colin McDowell (1), Grace Grimsley (1), Sarah Pippen (1), Peggi Angel (1)
(1) MUSC, Charleston, South Carolina, United States

Carbohydrate tumor antigen immunohistochemistry (IHC), represented by sialyl Lewis A/CA19-9 and sialyl Lewis X, has been evaluated by pathologists for decades for diagnostic tumor biomarkers in tissues. Our goal was to develop an integrated workflow that combines the molecular localization features of IHC and N-glycan imaging mass spectrometry (IMS) on the same slide with targeted chemical labeling approaches for identification of tumor associated

glycoproteins. Using commercially available antibodies to sialyl Lewis A and sialyl Lewis X antigens, a series of formalin-fixed paraffin embedded tumor tissues from prostate, colon, and pancreas were processed for standard IHC analysis. Each slide was scanned by high resolution microscopy, followed by removal of the coverslip. Next, an antigen retrieval heating step was done to dissociate bound antibodies, followed by spraying of a molecular coating of PNGase F PRIME for release of N-glycans. N-glycan IMS was done using a Bruker MALDI timsTOFfleX instrument, and the IHC-stained slide is already co-localized to the resulting glycan images. For both sialyl Lewis antigens, co-localized glycans in the stained tumor regions are multi-sialylated and multi-fucosylated branched structures. This workflow can further be extended by targeting the alpha-2,3 linked sialic acids that comprise the sialyl Lewis A/X motif. Prior to PNGase F digestion, the tissues can be treated chemically with an amidation reaction for alpha 2,6 isomers, followed by use of a biorthogonal amine-azide that specifically targets alpha-2,3 sialic acids. The N-glycan and IHC stained images can be used to scrape off regions of interest directly from the tissue. The lysate can be used for chemical affinity enrichment of the alpha-2,3 sialylated glycoproteins using a sialic acid bead enrichment reaction (SABER) by conjugation to an alkyne-bead. Alternatively, a sialic acid fluorescence enhancement reaction (SAFER) can be done in intact tissues to obtain sialic acid distributions by fluorescence at single cell resolution.

Contact: Richard Drake, draker@musc.edu

OA08.02 Mapping Glutathione Biosynthesis Heterogeneity via Multiple Infusion Start Times (MIST) and IR-MALDESI-MSI

Allyson Mellinger (1), Kenneth Garrard (2), Zahid Rabbani (3), Michael Gamcsik (3), David Muddiman (2)
(1) FTMS Laboratory for Human Health Research, Department of Chemistry, North Carolina State University, Raleigh, North Carolina, United States
(2) FTMS Laboratory for Human Health Research; Molecular Education, Technology and Research Innovation Center (METRIC); Department of Chemistry, North Carolina State University, Raleigh, North Carolina, United States
(3) University of North Carolina/North Carolina State University Joint Department of Biomedical Engineering, North Carolina State University, Raleigh, North Carolina, United States

Elevated levels of glutathione are characteristic to drug resistant tumor tissues. These tissues are highly heterogeneous and contain many different cellular microenvironments that may affect the observed disruption in glutathione abundances. These effects and the mechanisms involved require further elucidation. To map this heterogeneity two-dimensional sections of tissue, we developed a technique combining a stable isotope labeling tracer method and mass spectrometry imaging. While tracer studies are not uncommon to study biosynthesis kinetics, we uniquely employ a multiple infusion start time (MIST) method to monitor the incorporation of three glycine isotopologues into glutathione over the course of infusion time. We also use infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) coupled with a high resolution/accurate mass Orbitrap mass spectrometer to measure the degree of incorporation of the isotopologues across tissue sections. This platform allows us to ionize and fully resolve signals from the glutathione isotopologues without chemical derivatization. We then quantify the degree of incorporation in each analyzed tissue voxel using the newly developed percent isotope enrichment (PIE) tool within MSiReader. This tool first computes the percent isotope enrichment after correcting abundances for overlapping isotopic distributions within each voxel. It then generates heatmaps for visualization of flux heterogeneity across the tissues. This method has been used successfully to observe approximately 2-6% total isotopic enrichment within mouse liver tissue. Current efforts are focused towards adapting this protocol for studying tumor tissues and towards further developing the PIE tool for the computation of kinetic curves. Preliminary results in studying tumor

tissues reveal visible differences in isotopic enrichment across 4T1 mammary tumors in mice, which provide more dynamic information than heat maps of single time point measurements of glutathione molecules.

Contact: Allyson Mellinger, almellin@ncsu.edu

Parallel Session 09: Microbiome Multi-omics

IS09.01 Ran Blekhman Presentation
Ran Blekhman, University of Minnesota

Presentation coming soon.

IS09.02 Metaproteomics to investigate functional diet-microbiota interactions
Manuel Kleiner, North Carolina State University

Presentation coming soon.

OA09.01 NIST Stool Reference Material for Multi'omic Analyses
Sandra Da Silva (1), Stephanie Servetas (1)
(1) NIST, Gaithersburg, Maryland, United States

Researchers have established substantial evidence regarding the impact of the gut microbiome on human health and the potential for its use as a diagnostic tool and target for therapeutic interventions. Many gut microbiome studies depend on the analysis of stool samples, which present a complex matrix composed of microbes, protein, undigested plant matter and fat content that is strongly influenced by diet. Given the complexity of the gut microbiome (microbes and metabolites) and its importance for human health, the gut microbiome community has recognized a need to develop standards (e.g. reference material) to validate analytical workflows, improve reproducibility, and allow comparability among different efforts. To this end, the National Institute of Standards and Technology (NIST) has undertaken the effort to develop a human stool reference material to support this community. The development of a reference material, particularly a complex material such as human stool, occurs in several stages. We will present a detailed overview of the development process including key considerations for production and characterization of the material. Notably, this material will consist of pooled material, from two distinct cohorts, and the preparation will be suitable for multi-omic analysis. In addition to the detailed overview of the material development, we will present preliminary data on the metabolomic and metagenomic characterization of pilot materials and the results from an interlaboratory study assessing the current state of the field in metabolomic analysis of a standard fecal sample.

Contact: Sandra Da Silva, sdasilva@nist.gov

OA09.02 Next-gen Sequencing and Targeted Proteomics Reveal Global Binding and Suppression of the Herpesvirus Genome by the DNA Sensor IFI16
Timothy Howard (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Human cells express a wide array of pattern recognition receptors (PRRs) to identify invading pathogens and activate immune signaling pathways. The interferon-inducible protein 16 (IFI16) is a nuclear PRR that recognizes double-stranded DNA (dsDNA) from a number of viral sources, including genomes of nuclear-replicating viruses (e.g., herpesviruses and

papillomaviruses) and dsDNA intermediates formed during retrovirus (e.g., HIV) replication. Our lab and others have found that upon binding to the dsDNA genome of herpes simplex virus 1 (HSV-1), IFI16 induces antiviral cytokine expression and suppresses virus gene expression. Though structural studies have determined IFI16 binds dsDNA in a sequence-independent manner, the interaction between IFI16 and the 152 kilobase-pair HSV-1 genome remains uncharacterized. Further, it has been hypothesized that IFI16 promotes the heterochromatinization of the HSV-1 genome to achieve suppression of viral genes, but the direct effect of IFI16 binding to the virus DNA has not been investigated. Here, we utilized a multi-omics approach of DNA sequencing techniques paired with targeted mass spectrometry proteomics to obtain a complete view of the interaction between IFI16 and the HSV-1 genome, and how this binding affects virus DNA structure and protein expression. Specifically, we used chromatin immunoaffinity purification (ChIP)-sequencing and identified regions of IFI16 enrichment on the virus genome at loci that are crucial for virus replication. We next examined infected wild-type and IFI16-knockout cells for differences in virus genome structure and protein expression, using the assay for transposase accessible chromatin with sequencing (ATAC-seq) to probe IFI16-mediated changes in virus DNA conformation, along with parallel reaction monitoring-mass spectrometry to measure concomitant changes in the expression of more than 60 HSV-1 proteins. Altogether, we present the most complete view of a nuclear DNA sensor binding viral dsDNA, and we provide insights into how IFI16 represses HSV-1 replication by modifying the 3-dimensional structure of the virus genome.

Contact: Tim Howard, th12@princeton.edu

Parallel Session 10: Computational Proteomics: From Machine Learning to Human Insight

IS10.01 Olga Vitek Presentation

Olga Vitek, Northeastern University

Presentation coming soon.

IS10.02 David Fenyo Presentation

David Fenyo, NYU

Presentation coming soon.

OA10.01 Presentation

Presentation coming soon.

OA10.02 Creating a global map of cell structure from protein interactions and images

Leah V. Schaffer (1), Yue Qin (1), Edward L. Huttlin (2), Casper F. Winsnes (3), Gege Qian (1), Laura Pontano Vaites (2), J. Wade Harper (2), Steven P. Gygi (4), Emma Lundberg (3), Trey Ideker (5)
(1) University of California, San Diego, La Jolla, California, United States
(2) Harvard Medical School, Boston, Massachusetts, United States
(3) KTH-Royal Institute of Technology, Stockholm, Sweden
(4) Harvard Medical School, Boston, Massachusetts
(5) University of California, San Diego, La Jolla, California

Introduction Determining the structure of biological systems and its relation to function is one of the ultimate goals of the biological sciences. There has been a recent push to produce large-sized data sets from two of the different major approaches for mapping cell structure including the Human Protein Atlas (HPA), a major imaging effort consisting of >10,000 immunofluorescence (IF) images, and the BioPlex networks, derived from systematic affinity-purification mass spectrometry (AP-MS) experiments. Integrating the resulting datasets from these two different approaches provides an opportunity to generate a more complete map of cell structure across scales, from individual protein interactions to subcellular location of whole complexes. Methods Towards this goal, a recent collaborative effort developed an approach for creating Multi-Scale Integrated Cell (MuSIC) maps of cellular structure by integrating HPA IF images and the BioPlex protein interaction network. Pairs of protein embeddings from the two measurement types were integrated using a machine learning algorithm trained with Gene Ontology similarities to determine pairwise distances between proteins, and subsequently protein communities were detected from the network of these integrated protein-protein distances. Results This approach resulted in a MuSIC map for the 661 proteins in HEK-293 present in both the HPA and BioPlex data sets consisting of 69 subcellular systems, approximately half of which were novel. This map of the cell determined roles for poorly characterized proteins and identified new protein assemblies in processes such as ribosomal biogenesis and RNA

splicing. We will describe the initial developments to build MuSIC map v2.0 in the U2OS cell line consisting of the ~5,000 proteins that have HPA IF images and are present in the new U2OS BioPlex AP-MS network, representing a ten-fold expansion of the first MuSIC map in the HEK-293 cell line and thus providing a substantially more global view of cell structure.

Contact: Leah Schaffer, lvschaffer@health.ucsd.edu

Parallel Session 11: Proteomics and Biomimicry

IS11.01 Comparative Mammalian Proteomics: Increasing the Biomimicry Space alongside Genomes
Michael Janech, College of Charleston/Grice Marine Lab

With over 240 mammalian genome assemblies available through open-source repositories like DNAzoo, Zoonomia, and the Earth Biogenome Project, scientists are now able to interrogate chromosomal rearrangements, positive gene selection, copy number, and deletions across mammalian taxa like never before. Differences in genomic architecture due to evolutionary pressure result in much of the phenotypic diversity witnessed today and give rise to unique physiological adaptations that have become the topic of study for comparative physiologists. Some of these unique physiological adaptations hold promise for medical researchers interested in understanding ischemia/reperfusion injury (e.g. dive response in marine mammals) and uremia tolerance (hibernation in bears) amongst others. Although genome information is useful for understanding evolutionary relationships and physiological adaptations, differences in gene copy number or positive gene selection is difficult to translate into protein abundance, especially in biofluids. For this reason, parallel reaction monitoring was utilized to quantify the serum protein, pantetheinase, in 44 different mammals to determine whether a high abundance of serum pantetheinase was unique to diving marine mammals. Although levels could not be predicted from comparative genomics, diving marine mammals in general had relatively high levels of pantetheinase. In addition, the high pantetheinase phenotype also defined the ungulates (odd and even-toed), a previously unreported observation. To expedite hypothesis testing across taxa the Comparative Protein Aggregator Resource (CoMPARe) was initiated in 2018 to examine differences between mammalian taxa. This ongoing venture is collecting proteomic data from serum across 52 mammalian species. Ranked comparisons are being made following “humanization” of the proteomes to capture information that is not predicted from genomes. Exploration of the proteome adds another layer of information to describe interrelationships amongst mammals and will help to define the landscape of comparative mammalian physiology from which biomimicry can be employed to extract novel solutions for humans and advance industry.

IS11.02 Insights into human pathology from proteome studies of extreme adaptations in mammals
Jane Khudyakov, University of the Pacific

Many wild mammals are adapted to tolerate "extreme" environmental and physiological challenges, such as prolonged fasting, hypoxia, and rapid skin regeneration, among others. Understanding the molecular basis of natural tolerance in wild mammals may provide insights into the treatment of human pathologies such as obesity and skin disorders. The northern elephant seal is an excellent non-model "model system" for studies of natural adaptation to prolonged fasting, catastrophic molting, and hypoxia in mammals. We examined changes in skeletal muscle, adipose tissue, skin, and plasma proteomes of elephant seals over 5 weeks of fasting associated with rapid molting and identified enzymes, signaling molecules, and transcriptional regulators of these processes, which are potential targets for further studies of human disease. Our findings include changes in apolipoprotein composition that may underlie maintenance of naturally high HDL levels and resistance to oxidative stress during fasting and alterations in extracellular matrix composition and abundance of metabolic enzymes that may support rapid skin regeneration during catastrophic molting.

OA11.01 Leveraging large-scale comparative proteomics across the tree of life to improve human disease genetics
Rachael Cox (1), Ophelia Papoulas (1), Edward Marcotte (1)
(1) University of Texas at Austin, Austin, Texas, United States

Proteins interacting in the same biochemical complex are often linked to similar genetic traits. Moreover, previous studies have shown that evolutionarily conserved (ancient) proteins are enriched for disease traits and are abundant across human cell types and tissues. A significant portion of these deeply conserved genes are known to be responsible for a large and diverse subset of major human diseases, spanning developmental disorders (e.g., Leigh syndrome, microcephaly, neural tube defects), cancers (e.g., leukemia, breast cancer, colorectal cancer), chronic respiratory diseases (e.g., ciliary dyskinesia, asthma), neurological disorders (e.g., encephalopathy, schizophrenia, autism) and motor dysfunction (e.g., dystonia, spastic paraplegia). Thus, we sought to better characterize the ancient eukaryotic protein complexes that likely contribute to human disease by computationally integrating thousands of of mass spectrometry proteomics experiments spanning an evolutionarily diverse set of 31 eukaryotes. We combine this data with phylostratigraphy and machine learning to reconstruct a likely set of ancestral eukaryotic protein assemblies. We will present our progress on the analysis and application of these data to the discovery of disease associations and new functions for poorly characterized human genes.

Contact: Rachael Cox, rachaelcox@utexas.edu

OA11.02 Are Mouse Models of Alzheimer's Disease a Surrogate for Human Disease? – Using quantitative proteomics to monitor time resolved formation of pathology

Kathrin Wenger (1), Arthur Viode (1), Kinga Smolen (1), Christoph Schlaffner (1), Patrick van Zalm (1), Tammy Delovade (2), Xavier Langlois (2), Anthony Bannon (3), Juri Rappsilber (4), Hanno Steen (5), Judith Steen (1)
(1) Boston Children's Hospital, Boston, Massachusetts, United States
(2) Abbvie, Cambridge, Massachusetts, United States
(3) Abbvie, Cambridge, Massachusetts
(4) Technische Universität Berlin, Berlin, Germany
(5) Boston Children's Hospital, Boston, Massachusetts

While Alzheimer's Disease (AD) was described and characterized in 1901 by Alois Alzheimer, a cure has remained elusive 121 years later. Most if not all therapeutics are tested preclinically in mice and as such understanding if these mouse models recapitulate human disease is important to finding therapeutics. Transgene mouse models that overexpress human Tau with known risk mutations such as P301S or P301L are a widely used tool for Tau drug development in AD as these models exhibit robust Tau aggregation during disease progression. It is, however, not clear to which extent these models reflect human Tau pathology and thus AD at the molecular level. An in-depth, mass spectrometry-based proteomic analysis was applied to study pathological and functional Tau, derived from affected brain tissue of two commonly used mouse models of AD: the P301S (hTau.P301S)) and P301L (rTg(tauP301L)4510) model expressing mutant Tau forms. Identification of post translational modifications (PTMs) on Tau and quantification of the extent and changes thereof during Tau aggregation provides molecular insights into the drivers during the appearance of the pathology. Comparison to human AD (Wesseling, Mair et al, Cell., 2020) shows similarities, but also highlights the dramatic differences between the human disease and the preclinical animal models. Our data show that the two mouse models reflect a subset of modifications of pathological Tau which are observed in human AD. However, many other disease and disease stage associated PTMs that are important in symptomatic human AD, are not represented in these mouse models. This analysis provides guidance regarding the extent to which mouse models can be used in preclinical studies of therapeutics. Thus, such in depth molecular characterization and comparison of mouse models and human disease are crucial in designing mechanistic studies and preclinical testing of therapeutics for application in human disease.

Contact: Kathrin Wenger, kathrin.wenger@childrens.harvard.edu

Parallel Session 12: Proteomics to Advance Equity

IS12.01 Differences in Stromal Patterns from Breast Cancer Metastatic Lymph Between Black Women and White Women in the South Carolina Sea Islander Population

Peggi Angel, Medical University of South Carolina

In South Carolina (SC), there is a disproportionate increase in female breast cancer death rates in black women (BW) compared to white women (WW) (30.1 and 21.2/100,000 respectively, years 2000-2019). Historical accounts of the SC slave trade report that the majority of the enslaved African originated from the sub-Saharan West African Coastal regions. After slavery, these Africans remained in relative isolation within communities along the south eastern coastal islands, called Sea Islands (SIs). The West African origins of the Sea Islanders predispose this diasporic population to higher breast cancer risk and development of more aggressive breast cancers. Contemporary literature reports that stromal collagen differences are predictive of breast cancer progression and survival. We hypothesized that collagen stroma variations between SI BW and WW may be involved in disproportionate SC breast cancer outcomes. The study investigated collagen stromal variations by targeted collagen tissue imaging proteomics in breast cancer tumor, normal adjacent tissue, normal adjacent lymph, and metastatic lymph tissue. Newly diagnosed patients were from documented SI geographic regions (BW n=10; WW n=21). Collagen peptide peak intensities were analyzed using Area Under the Receiving Operating Curve, p-value <0.01 to determine differentiating signatures between BW and WW. Intriguingly, the largest variation when comparing by race occurred in the lymph nodes. Normal lymph tissue showed 83/1377 significantly different collagen peptides while the metastatic lymph tissue showed significant changes in 74 collagen peptides, t-test p-value ≤0.001. Certain peptides were significant both metastatic and normal lymph tissue, whereas others were uniquely significant in either metastatic or normal lymph. This study suggests that there may be an ancestry-dependent immune involvement to metastatic breast cancer that could contribute to higher breast cancer mortality rates in BW from SC SI regions. Additional studies are in progress investigating collagen stroma and immune system involvement in West African-origin breast cancer.

IS12.02 In-Cell Protein Footprinting Coupled with Mass Spectrometry for Structural Biology Across the Proteome

Lisa Jones, University of Maryland

In recent years, protein footprinting coupled with mass spectrometry has been extensively used to analyze the higher order structure (HOS) of proteins. These methods have been successfully used to identify protein-protein interaction sites and regions of conformational change through modification of solvent accessible sites in proteins. The footprinting method, fast photochemical oxidation of proteins (FPOP), utilizes hydroxyl radicals to modify these solvent accessible sites. To date, FPOP has been used in vitro on relatively pure protein systems. We have further extended the FPOP method for in-cell and in vivo analysis of proteins. This will allow for study of proteins in their native cellular environment and be especially useful for the study of membrane proteins which can be difficult to purify for in vitro studies. A major application of these methods is for proteome-wide structural biology. In one such application, we used in-cell FPOP (IC-FPOP) to identify on and off targets of the cancer drug methotrexate in leukemia cells. By obtaining structural information on proteins across the proteome, we were able to distinguish structural changes that occur in response to drug treatment. We have further extended the FPOP method for analysis in C. elegans, a member of the nematode family. This allows us to study protein structure directly in animal model for human disease. These methods have the potential to become a powerful tool in the structural biology toolbox.

OA12.01 Proteomic Changes Associated with Survival Outcomes of Sepsis Patients with Primary Intra-Abdominal Infection

Kathryn Kapp (1), Albert Arul (1), Kevin Zhang (2), Liping Du (2), Sachin Yende (3), John Kellum (3), Derek Angus (3), Octavia Peck-Palmer (4), Renã Robinson (1)
(1) Vanderbilt University, Nashville, Tennessee, United States
(2) Vanderbilt University Medical Center, Nashville, Tennessee, United States
(3) University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, United States
(4) University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

The plasma proteome can distinguish patient survival outcomes as early as hospital admission. Patient response to sepsis is highly heterogeneous and depends on primary source of infection and potentially the patient's racial background. For example, African American/Black patients have higher mortality rates of sepsis than non-Hispanic Whites. To better understand acute changes in sepsis that impact survival outcomes, we analyzed plasma samples from a racially diverse cohort of sepsis patients with primary intra-abdominal (IA) infection. Blood plasma samples (N = 107) from non-Hispanic White and African American/Black survivors and non-survivors of sepsis secondary to IA infection were obtained from the Protocolized Care for Early Septic Shock cohort at emergency department admission. Survivorship was defined as a patient's status at 60 days post-admission. Plasma samples were immunodepleted, digested using trypsin/Lys-C, tagged with tandem mass tag reagents, fractionated, and analyzed with high-resolution liquid chromatography-tandem mass spectrometry (LC-MS/MS). Linear regression models assessed the effects of survival on protein outcome with and without stratification based on the patients' self-identified race. In total, 2,540 plasma proteins were identified, and from regression analysis, 224 of these proteins were differentially-expressed between non-survivors (N = 45) and survivors (N = 62). Differentially-expressed proteins are primarily involved in the hepatic fibrosis/stellate cell activation pathway, acute phase response signaling, and the coagulation and complement systems. In race-stratified comparisons, 151 proteins were differentially-expressed between non-survivors and survivors, 103 of which were unique to one racial group. We identified eight proteins with a significant race x survival outcome interaction (FDR adjusted $p < 0.05$) that are involved in the acute phase response, complement system, and liver X receptor/retinoid X receptor and farnesoid X receptor/retinoid X receptor activation pathways. This presentation will discuss study findings and implications of the plasma proteome toward identifying acute changes related to patient survival and racial background.

Contact: Kathryn Kapp, kathryn.l.kapp@vanderbilt.edu

OA12.02 Resolving racial differences in the brain proteome of African Americans with Alzheimer's disease

Erica Modeste (1), Lingyan Ping (1), Duc M. Duong (1), Marla Gearing (1), James J. Lah (1), Allan I. Levey (1), Nicholas T. Seyfried (1)
(1) Emory University, Atlanta, Georgia, United States

Chronic health conditions including cardiovascular disease and diabetes, lower quality and level of education, higher rates of poverty, and greater exposure to discrimination disproportionally affect African Americans, putting this population at a heightened risk for Alzheimer's disease (AD) and related dementias. Neuropathological hallmarks of AD include accumulations of amyloid-beta (Aβ) plaques and tangles of hyperphosphorylated tau in the brain. Levels of Aβ and tau in cerebrospinal fluid (CSF) can act as proxies for AD pathology in the brain. Emerging biomarker studies in CSF show,

however, that African Americans have lower levels of tau, compared to Caucasians. This suggests that race may influence tau and related AD pathophysiology in African Americans. Whether race impacts tau levels or other pathophysiologies in the brain is not well established. Here, we sought to characterize the brain proteome of Caucasians and African Americans with AD. Using tandem mass tag mass spectrometry-based proteomics coupled with off-line fractionation, we quantified ~8,000 proteins across 129 matched African-American and Caucasian AD and control postmortem brain tissues. Notably, characteristic hallmarks of AD, including increases in Aβ peptide and total tau levels in brain, did not differ by race. Following phosphopeptide enrichment using iron-chelated magnetic beads, approximately 80 tau phosphopetides were identified and quantified. Four of those tau phosphopeptides were statistically different between Caucasians and Africans Americans. None, however, fell within the microtubule-binding region most associated with disease and tau tangles. Furthermore, co-expression network analysis revealed pathways of proteins specific to neuronal, microtubule binding, vascular, and RNA splicing / binding biology that differed between African Americans and Caucasians, particularly in the controls. Together, this data suggests that race plays a role in AD, particularly at earlier stages and these differences exist outside of the pathological hallmark proteins, Aβ and tau.

Contact: Erica Modeste, esmodes@emory.edu

Parallel Session 13: Proteomics in Drug Discovery

IS13.01 Profiling the Diversity of Agonist Selective Effects on the Proximal Proteome Environment of GPCRs
Ruth Huttenhain, University of California, San Francisco

G protein-coupled receptors (GPCRs) represent the largest family of signaling receptors and drug targets. Activation of GPCRs can initiate complex signaling networks from different cellular compartments. Depending on the cellular location and the signaling pathways activated, GPCRs can elicit different cellular effects. This capacity may be exploited in drug discovery by designing ligands selectively activating specific receptor-mediated signaling pathways or target receptors at selected cellular locations. However, the efforts to create pathway or location specific ligands is limited by our understanding of GPCR-based signaling and trafficking mechanisms. One reason for this is the lack of technologies that can capture the full complexity of GPCR activity.

Recently, we developed a proximity biotin labeling method for GPCRs combining the engineered peroxidase APEX with quantitative mass spectrometry (GPCR-APEX). GPCR-APEX simultaneously captures proximal protein interaction networks and cellular location of the activated receptor with high temporal resolution. While this represents a powerful unbiased method to profile agonist-selective effects on receptor interaction networks and cellular localization, a major limitation of its application has been to deconvolve this information from the proximity labeling datasets.

Here, we present a novel computational framework which predicts time and agonist dependent subcellular location of the receptor and quantitatively deconvolutes the effect of receptor location and proximal interactors in the proximity labeling data. As our model system for developing the framework, we selected the mu-opioid receptor (MOR). Activating MOR with chemically distinct agonists, we show pronounced differences in the proximal proteome of MOR elicited by their distinct capacity to initiate receptor trafficking. Moreover, we discovered two novel G protein associated proteins–KCTD12 and EYA4. We show that recruitment of KCTD12 and EYA4 is dependent on Gαi activity and suggest that KCTD12 and EYA4 form a buffer system for G protein signaling by interacting directly with the Gβ/γ and Gα subunits, respectively.

IS13.02 Enabling Drug Discovery with Practical Proteomics
Jon Williams, AbbVie

Mass spectrometry based proteomic technologies are indispensable for small molecule drug discovery programs. For example, proteomics and chemical proteomics are used, along with other technologies, to identify protein targets from well-defined phenotypic readouts. Intact protein mass spectrometry (MS) is performed to ensure the appropriate protein constructs are made, define PTM status, and can be used in covalent inhibitor programs as a screening method. Chemical proteomic experiments are conducted to assess both on-target and off-target engagement. This presentation will present practical examples and insights into how proteomic technologies are used throughout the drug discovery pipeline. Also, featured will be the development and incorporation of a new technology for intact protein and protein characterization studies, infrared matrix assisted laser desorption ionization mass spectrometry IR-MALDESI-MS.

OA13.01 Unveiling of RF-amide family GPCRs interaction networks and signaling pathways with proteomics for analgesics development
Qiongyu Li (1), Benjamin Polacco (1), Trupti Patil (1), Aaron Marley (1), Aliza Ehrlich (1), Jiewei Xu (1), Xiaofang Zhong (1),

Nevan Krogan (2), Mark Von Zastrow (1), Ruth Huttenhain (2)
(1) UCSF, San Francisco, California, United States
(2) UCSF, San Francisco, California

Studies on a group of GPCRs, the RF-amide peptide receptors (RFPR), demonstrated their involvement in the regulation of nociception, which further revealed the potential of the development of analgesics targeting RFPRs. In this study, two receptors of RFPR family, including NPFFR1 and QRFPR have been studied for their interaction networks and signaling pathways with mass spectrometry (MS)-based proteomics. To delineate interaction networks of NPFFR1 and QRFPR, an APEX-based proximity labeling strategy was employed on HEK293T cells stably expressing either receptor. Receptors were activated with their corresponding agonists for different time courses ranging from 1 to 30 min. To understand further the signaling pathways downstream of activated QRFPR, the global quantitative phosphoproteomics was conducted with QRFPR activated by its agonist. For NPFFR1, further study of the regulation of its interaction network by its agonist was achieved by conducting APEX-based proximity labeling on wild and phosphosite-mutant types of NPFFR1 in parallel. The APEX-labeling data revealed different interaction networks of NPFFR1 and QRFPR. The data demonstrated that the activation of QRFPR lead to the labeling of proteins known as subunits of complexes involved in actin filament polymerization. For NPFFR1, several components of the CUL1-E3 ligase were labeled with the APEX-proximity labeling. The phosphoproteomics for QRFPR activation generated 4000 significantly regulated phosphorylation sites, with more than 200 sites being regulated in a Gq-dependent manner. Furthermore, the comparative APEX-labeling data of wild and phosphosite-mutant NPFFR1 revealed the dependence of the recruitment of CUL1-E3 ligase on the phosphorylation of phosphodegron on NPFFR1. With novel proteomics approaches combining with molecular biology tools, we were able to reveal the molecular changes downstream of two RFPR family receptors. This leads to the further investigation of molecular mechanisms of their activation, which will guide the design of novel molecules targeting them for the development of analgesics.

Contact: Qiongyu Li, qiongyu.li@ucsf.edu

OA13.02 Identification of Host Cell Proteins Critical to the Hepatitis B Viral Life Cycle

Isabella Whitworth (1), Sofia Romero (1), Abena Kissi-Twum (1), Rachel Knoener (1), Mark Scalf (1), Daniel Loeb (1), Nathan Sherer (1), Lloyd Smith (2)
(1) University of Wisconsin-Madison, Madison, Wisconsin, United States
(2) University of Wisconsin-Madison, Madison, Wisconsin

Hepatitis B Virus (HBV), a leading cause of liver cancer, chronically infects over 200 million people worldwide, has no reliable cure, and lacks targetable host factors for development of therapeutics. Identification of new drug targets will require moving beyond just protein characterization. Physical interactions between biomolecules, including protein and RNA, provide essential information about viral life cycles. Hybridization-Purification of RNA-protein complexes followed by Mass Spectrometry (HyPR-MS) is an RNA-centered approach to characterize RNA-protein complexes by purifying target RNA and performing mass spectrometric analysis to identify bound proteins. Previous research focused on the packaging of HBV pregenomic RNA (pgRNA) has suggested unknown factors are required for the specific packaging of the genome into capsids. By looking at the differentially expressed proteins bound to pgRNA in cells transfected with a packaging competent HBV construct (ε+) and with a packaging incompetent HBV construct (ε-), host proteins specifically associated with genome packaging can be identified. Hybridization and capture of pgRNA as well as scrambled and mRNA controls was performed on 4 bioreplicates of cell transfected with ε+ plasmids and 4 bioreplicates of cells transfected with ε- plasmids. Pairwise T-

tests identified 267 proteins with a p-value less than 0.05 that were at least 2X more abundant in cells transfected with the ε+ construct. This includes proteins that are known to promote HBV capsid formation, catalyze HBV viral protein dephosphorylation, post-transcriptionally modify the ε loop of pgRNA, as well as proteins known to participate more broadly in viral regulation. 30 of these proteins have been selected for further evaluation via siRNA knockdowns based on their fold-change differences and previous identifications in the literature. After host protein knockdowns, the impact on packaging will be evaluated by comparing the pgRNA packaging efficiency of knockdown and wildtype cells.

Contact: Isabella Whitworth, iwhitworth@wisc.edu

Parallel Session 14: Enabling Pandemic Proteomics

IS14.01 Douglas Fraser Presentation

Douglas Fraser, London Health Sciences Centre

Presentation coming soon.

IS14.02 Revealing and targeting host proteostasis to inhibit virus infections

Lars Plate, Vanderbilt University

The ongoing COVID-19 pandemic caused by SARS-CoV-2 has brought fresh attention to the enormous global health risks posed by RNA viruses and the need for broad antiviral strategies. Our focus is on host-targeted therapeutics by expanding our understanding of common host processes critical for virus infections. Comparative mass spectrometry-based interactomics is a powerful approach to identify and sensitively compare the enrichment of shared and unique host cell dependencies exploited by virus proteins. We coupled affinity purification to TMTpro-multiplexed quantitative proteomics to profile the host cell interactions of several nonstructural proteins from SARS, non-pathogenic coronavirus strains, as well as newly emerging SARS-CoV-2 variants. We identified pan-strain interactions with mitochondria-associated endoplasmic reticulum (ER) membrane factors and ER proteostasis pathways. As RNA virus replication and assembly frequently occurs at the ER membrane, viruses extensively remodel this organelle and the ER proteostasis network requiring precise modulation of the unfolded protein response (UPR). We found divergent roles for two nonstructural proteins, nsp4 and nsp3, in activating and suppressing the UPR to finetune its activity. To disrupt these host ER proteostasis dependencies during viral propagation, we can take advantage of recently developed small-molecule ER proteostasis regulator compounds. We discovered that such compounds are broadly effective at inhibiting infections with flaviruses (dengue and Zika) and we are currently exploring their use against human and mouse coronavirus strains. These results open up a broadly applicable therapeutic strategy to target essential host ER protein folding processes to impair viral infections.

OA14.01 Comparative Multiplexed Interactomics of SARS-CoV-2 and Homologous Coronavirus Nonstructural Proteins

Jonathan Davies (1), Katherine Almasy (1), Lars Plate (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

Human coronaviruses (CoVs) are a threat to global health and society, as evident from the SARS outbreak in 2002, the MERS outbreaks in 2012 and 2014, and the most recent COVID-19 pandemic. Despite the sequence similarity between severe disease-causing CoVs, each strain has distinctive virulence. A better understanding of the basic molecular mechanisms mediating changes in virulence is needed. Here, we profile the virus-host protein-protein interactions of three CoV nonstructural proteins (nsps) that are critical for virus replication. We use tandem mass tag-multiplexed quantitative proteomics to sensitively compare and contrast the interactomes of nsp2, nsp3, and nsp4 from three severe disease-causing CoV strains (SARS-CoV-2, SARS-CoV, MERS-CoV) and two common-cold CoVs (hCoV-OC43, hCoV-229E). This approach enabled us to identify both unique and shared host cell protein interactors and quantitatively compare the enrichment of interactions between homologs. Both nsp2 and nsp4 common interactors are strongly enriched for ER-mitochondria contact site proteins, suggesting a new functional role for modulating host processes such as calcium homeostasis and mitochondrial function. Furthermore, using functional genetic screens we show many of these interactors

are pro-viral for CoV infection. Nsp3 homologs showed more strain-specific interactor profiles, including interactions with nuclear import machinery for hCoV-229E and ribosomal RNA processing for MERS-CoV. Lastly, we find that SARS-CoV-2 nsp3 interacts with ATF6, a regulator of the Unfolded Protein Response, and can suppress the ATF6 stress response. Our results shed light on the role these CoV proteins play in the infection cycle, as well as host factors that may mediate the divergent pathogenesis of common cold CoVs from SARS/MERS strains. Our mass spectrometry workflow enables rapid, robust comparisons of multiple bait proteins and identification of common CoV-host dependencies to be targeted by host-directed anti-viral therapeutics.

Contact: Jonathan Davies, jonathan.p.davies@vanderbilt.edu

OA14.02 Plasma proteomics of human Ebola virus infection distinguishes and defines outcomes

Arthur Viode (1), Kinga K Smolen (1), Benoit Fatou (1), Zainab Wurie (1), Patrick van Zalm (1), Mandy Kader Konde (2), Balla Moussa Keita (3), Richard Amento Ablam (4), Eleanor N. Fish (5), Hanno Steen (6)
(1) Boston Children's Hospital, Boston, Massachusetts, United States
(2) Sustainable Health Foundation (FOSAD), Center of Excellence for Training on Research & Priority Diseases (CEFORPAG), Conakry, Guinea
(3) Sustainable Health Foundation (FOSAD), Conakry, Guinea
(4) Sustainable Health Foundation (FOSAD), Conakry
(5) Toronto General Hospital Research Institute, University Health Network, Toronto, Canada
(6) Boston Children's Hospital, Boston, Massachusetts

The 2013-2016 Ebola Zaire virus (EBV) outbreak in West Africa resulted in over 28,000 cases and 11,000 deaths. Ebola virus disease (EVD) is a highly virulent systemic disease with a high case fatality rate of ~50%. EVD results in hemorrhagic fever marked by an exaggerated systemic inflammatory response, and impaired vascular and coagulation systems. The immune response of patients who either survived or died is characterized by pronounced differences. Notably, fatalities showed a diminished capacity to mount an appropriate immune response, resulting in high viremia and increased pro-inflammatory cytokine production. In this study, we analyzed 38 sequential samples collected from 12 patients: 8 survivors and 4 fatalities. Our analytical strategy combined three analytical platforms covering three different fractions of the plasma proteome: the undepleted classical plasma proteome, the depleted plasma proteome, and cytokines/chemokines, using LC/MS- and antibody-based assays. This analysis resulted in over 1000 quantified host and pathogen proteins. For depletion of the most abundant plasma proteins, we advanced a perchloric acid-based precipitation method. This method is low cost, high-throughput and has proven its robustness for thousands of plasma samples in the Immunophenotyping Assessment in a COVID-19 Cohort study (IMPACC; Sci Immunol. 2021), which bring significant advancement in the field of high-throughput proteomics. We first characterized the early host response to EBV infection and identified distinct protein signatures that could differentiate between fatalities and survivors. Next, we focused on proteins whose abundance levels changed based on viral load and disease course. Our data revealed distinct protein signatures and identification of a prognostic 4-protein biomarker panel that performs better than the widely used viral load, comprising a histone, a ribosomal protein, moesin, and kininogen. Given the easy implementation of protein assays, this 4-protein panel may help direct care in emergency settings in resource poor regions.

Contact: Arthur VIODE, arthur.viode@childrens.harvard.edu

P01: Advances in Technology

P01.01 Evaluating Impacts of Nanoflow Column Length and Core-shell Particle Size for Proteomics DDA and SWATH Workflows

Alexandra Antonoplis (1), Jason Anspach (2), Christie Hunter (1)
(1) SCIEX, Redwood City, California, United States
(2) Phenomenex, Torrance, California, United States

Introduction: Nanoflow chromatography is often used in proteomics DDA and SWATH acquisition workflows in order to obtain the highest sensitivity on digested samples. Peak shape and resolution are critical to allow the MS system to sample as many peptides as possible. Here, new long nanoflow columns packed with core-shell chromatographic phase were evaluated in DDA and SWATH workflows with digested K562 protein extracts. Methods: In the current study, three different nanoflow column lengths (15, 25, and 50 cm) were packed with core-shell Kinetex C18 phase. For the 50 cm column, both 2.6 µm and 5.0 µm particle sizes were tested. Gradient durations ranging from 60-180 minutes were used across the column types for DDA workflows. K562 cell digest was evaluated at several loads and a trap-elute workflow was implemented with an analytical flow rate of 300 nL/min. The TripleTOF 6600+ system and the ZenoTOF 7600 system were used for all data acquisition, and data was processed using the ProteinPilot App in OneOmics suite. Results: Here, the impacts of column length, column particle size, and gradient duration on protein and peptide identifications were assessed. As expected, protein and peptide identifications increased as the gradient length increased and column length increased, with optimal identifications observed with a 180 min gradient and 50 cm column length. The impacts of higher sensitivity MS/MS (Zeno MS/MS) on protein ID gains were evaluated using the final nanoflow chromatographic conditions. Optimized DDA and SWATH acquisition methods were also determined for nanoflow system suitability analysis. Conclusions: Using 50 cm nanoflow columns packed with core-shell chromatographic phase improved protein and peptide identifications in DDA and SWATH workflows.

Contact: Alexandra Antonoplis, alexandra.antonoplis@sciex.com

P01.02 Infrared Matrix-Assisted Laser Desorption Electrospray Ionization and Quadrupole Time-of-Flight Mass Spectrometer: Design, Coupling and Challenges.

Cristina Arciniega (1), Ken Garrard (1), Jacob Guymon (1), Alex Apffel (2), John Fjeldsted (2), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States
(2) Agilent Technologies, Santa Clara, California, United States

Infrared Matrix-Assisted Laser Desorption Electrospray Ionization (IR-MALDESI) is an ambient ionization source that combines matrix-assisted laser desorption with electrospray ionization. This source has been widely employed for mass spectrometry imaging (MSI) and high-throughput screening (HTS) applications. It has been previously coupled to FT-ICR and Orbitrap mass spectrometers due to their high resolving power and mass accuracy. These are pulsed instruments that require tens to hundreds of milliseconds between scan, consequently limiting throughput. This work focuses on the integration and characterization of IR-MALDESI source to a commercial Q-TOF mass spectrometer for the analysis of biological samples. For this purpose, an IR-MALDESI source has been designed as a proof of concept to perform analyses of samples directly from microliter well plates. Also, an extended metal capillary has been engineered for the front end of the mass spectrometer. The ionization source is equipped with a 2970nm IR laser which operates at 100Hz using 1 pulse-per-

burst (PPB) producing a quasi-continuous ion generation. Since the Q-TOF is a continuous mass analyzer, IR-MALDESI was adapted to properly couple the source with the instrument. In this study, various samples were analyzed to prove the functionality of the IR-MALDESI source paired with a Q-TOF mass spectrometer.

Contact: Cristina Arciniega, cvarcini@ncsu.edu

P01.03 Highly multiplexed, spatially resolved protein and gene expression profiling in FFPE tumor tissues

Jennifer Chew (1), Cedric Uytingco (1), Naishitha Anaparthy (1), Jun Chiang (1), Christina Galonska (2), Karthik Ganapathy (1), Ryo Hatori (1), Alexander Hermes (3), Layla Katiraee (1), Anna-Maria Katsori (4), William Nitsch (1), Patrick Roelli (2), Joe Shuga (1), Rapolas Spalinkas (2), Benton Veire (1), Dan Walker (1), Neil Weisenfeld (1), Stephen Williams (1), Zachary Bent (1), Marlon Stoeckius (2)
(1) 10X Genomics, Pleasanton, California, United States
(2) 10X Genomics, Stockholm, Sweden
(3) 10X Genomics, Pleasanton, California
(4) 10X Genomics, Stockholm

The tumor microenvironment (TME) is composed of highly heterogeneous structures and cell types that dynamically influence and communicate with each other. Protein interactions within these cell types often play a critical role in how the cancer develops, progresses, and responds to therapies. Traditionally, immunofluorescence or immunohistochemistry has been used to annotate and detect proteins and other antigens in tissue sections. More recent higher throughput approaches capture phenotypes using a limited number of protein markers while spatially interrogating targeted or transcriptome-wide expression of RNA in tissue sections. However, for a more comprehensive understanding of the unique characteristics of cells within the TME, multiple layers of information must be studied together. Here we demonstrate a novel, streamlined multiomic spatial assay that integrates histological staining and imaging with simultaneous transcriptome-wide gene expression and highly multiplexed protein expression profiling from the same formalin-fixed paraffin embedded (FFPE) tissue section. In short, tissue sections from archived FFPE samples were placed on slides containing arrayed capture oligos with unique positional barcodes. The tissues were stained with hematoxylin and eosin, imaged and then incubated with transcriptome-wide probes and a high-plex DNA-barcoded antibody panel containing intra- and extracellular markers. Transcriptome probes and antibody-barcodes were then spatially captured on the slide and converted into sequencing-ready libraries. We apply this method to simultaneously measure gene and protein expression within the TME of human breast cancer FFPE samples using whole transcriptome probes and an immune-oncology antibody panel. This data enables comparison and correlation of multiple analytes and their patterns within the same sample section and demonstrates that a spatially resolved, multiomic approach provides a more comprehensive understanding of cellular behavior in and around tumors, yielding new insights into disease progression, predictive biomarkers, drug response and resistance, and therapeutic development.

Contact: Jennifer Chew, jennifer.chew@10xgenomics.com

P01.04 Deep Plasma Proteomics at Scale with Proteograph Product Suite: A Performance Evaluation with Label-free and TMT Multiplexing Methods

Ryan Benz (1), Xiaoju Zhang (1), Tianyu Wang (1), Mathew Ellenberger (1), Gabriel Castro (1), Ray Schmidt (1), Veder Garcia (1), Qiu Yang (2), Chenyi Zhao (1), Purvi Tandel (2), Michelle Dubuke (1), Ramón Díaz Peña (3), Alex Campos (3),

Khatereh Motamedchaboki (1)
(1) Seer, Redwood City, California, United States
(2) Seer, Redwood City, California
(3) Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California, United States

Human blood plasma is a widely accessible sample for assessing individual health status. However, the large dynamic range of circulating proteins combined with the diversities of proteoforms present in plasma have limited the comprehensive characterization of the plasma proteome in a high throughput manner. To address such challenges, current plasma proteomics workflows combine immunodepletion of high abundance proteins, peptide fractionation and sample multiplexing approaches such as tandem mass tags (TMT). Recent advancement in sample preparation (Seer's Proteograph Product Suite), coupled with improved mass spectrometry instrument sensitivity and speed, enable the quantification of thousands of proteins from plasma without compromising throughput or reproducibility, creating a unique opportunity to detect robust protein biomarkers for complex diseases. Here we evaluate the performance of label-free and TMT multiplexing methods with a set of control plasma samples processed with Proteograph Product Suite for deep plasma proteomic analysis. Pooled control human plasma samples were processed with Proteograph in 4 batches prepared on 4 different days. Tryptic peptides were either directly analyzed by LC-MS/MS or labeled with one of the TMTpro reagents followed by peptide fractionation (high pH RP) and LC-MS/MS analysis. LC-MS analysis were performed with 40-48-hours workflow (LFQ and TMTpro 16plex). TMT with peptide fractionation, resulted in ~3,000 protein groups. Approximately 86% of the peptide features are detected across all 4 batches, with a median CV < 20%, (PSM level). Label-free performance across four plates run on two different instrument and two different days were evaluated with protein group intensity for most NPs with CVs < 20%, enabling large scale plasma proteomics without compromising depth or precision. We detected plasma proteins spanning 9 orders of magnitude including 40 cytokine activity proteins and several members of TNF superfamily. This study evaluates performance of the Proteograph platform combined with label-free or TMT technology mass spectrometry analysis.

Contact: Ryan Benz, rbenz@seer.bio

P01.05 Optimization of an Automated 96-well Plate Protein Depletion Protocol for Plasma

Min Choi (1), Albert Arul (1), Rena Robinson (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

Depletion of high-abundant proteins in plasma increases the coverage of the plasma proteome detectable by mass spectrometry (MS). Traditional methods of protein depletion can require processing one sample at a time using an immunoaffinity-based depletion column, which presents a bottleneck in a high-throughput proteomics workflow. Single-use spin columns can shorten the processing time through sample parallelization but then becomes limited to centrifugal rotor capacities. The goal of this study was to optimize an automated platform for depleting multiple plasma samples simultaneously on a 96-well plate packed with High Select Top 14 resin (Thermo Fisher). In order to benchmark our results, we compared depletions of human plasma standard on the 96-well plate method with multiple injections on an Agilent Human 14 Multiple Affinity Removal column (MARS14). Plasma standards were depleted by the 96-well plate method using an automated liquid handling device. Different plasma-to-resin ratios and incubation methods were tested to optimize depletion efficiency calculated using bicinchoninic acid (BCA) protein assays of the automated and column methods. The binding efficiencies of the top 14 abundant proteins captured by the High Select Top 14 resin and MARS14 column were evaluated by bottom-up proteomics. Compared to the MARS14 column, the automated 96-well method using the High

Select Top14 resin resulted in decreased depletion and binding efficiencies. However, the depletion and binding efficiencies of the automated platform were reproducible across the technical replicates. Overall, the automated platform can reduce processing time by depleting up to 96 samples simultaneously in less than an hour.

Contact: Min Choi, min.j.choi@vanderbilt.edu

P01.06 Optimization of urinary proteomics workflow for the rapid analysis

H. Sophia Chung (1), Junmin Wang (1), Shao Huan Samuel Weng (1), Chelsea Boo (1), Sonja Hess (1), Ventzislava Hristova (1)
(1) AstraZeneca, Gaithersburg, Maryland, United States

Purpose Urine has tremendous value for non-invasive biomarker identification and routine detection, but intra- and inter-individual variability present a challenge with interpretation. Here, we aimed to standardize urine collection and disease-specific processing and analysis. Our end-to-end workflow enables rapid detection of proteins with minimum sample requirements to support large cohort studies. Methods Eight pooled urine samples, 10 donors per pool, were digested and fractionated by HPLC into 12 fractions. Each fraction was analyzed with an optimized data-dependent acquisition (DDA) method and data-independent acquisition (DIA) method on an Orbitrap Exploris 480. DDA and DIA data were used to generate a spectral library. Subsequently, individual urine specimens were analyzed without fractionation and searched against a library or analyzed using directDIA for large cohorts. High abundance proteins in chronic kidney disease (CKD) urine, such as albumin, presented a challenge and thus, strategies to reduce albumin interference were introduced for disease-specific sample processing to improve proteome coverage. Results The impact of processing variables including centrifugation and storage temperature on the true urine proteome were evaluated. We assessed a range of LC gradient lengths, chromatography columns and ionization sources and determined that a 30 minute gradient with an ES906 column and Easy-Spraysource on the Exploris 480 was optimal for rapid and comprehensive analysis. For DIA analysis, multiple spectral libraries were generated and evaluated based on size and proteome coverage, with current libraries containing >3400 protein groups. Using this rapid DIA workflow, single shot analysis of individual healthy urine samples detected >2000 protein groups without fractionation. Currently, the workflow is further being optimized and applied to the analysis of CKD urine, to address numerous analytic challenges due to extensive proteinuria and high levels of albumin. Conclusion We have optimized and standardized an end-to-end high-throughput workflow that enables rapid urine analysis.

Contact: Sophia Chung, sophia.chung@astrazeneca.com

P01.07 Assessing the Precision of Detergent-based Workflows: Cartridge-Based Proteome Precipitation for Robust Sample Preparation

Alan Doucette (1), Jessica Nickerson (1), Victoria Miller (2), Sara Little (2), Ziheng Dang (1)
(1) Dalhousie University, Halifax, Nova Scotia, Canada
(2) Proteoform Scientific, Halifax, Nova Scotia, Canada

Introduction: Multiple advances in MS instrumentation have improved the rate, depth and accuracy of proteome profiling. Particularly for comparative and quantitative proteome characterization, a reliable analytical platform is essential for accurate and precise measurements. Inconsistencies during proteome sample preparation can contribute even greater inconsistencies than the MS platform, and thereby presents a bottleneck to delivering repeatable results. Here we evaluate

the analytical precision and robustness of an SDS-based workflow for bottom-up proteome analysis. The ProTrap XG is a disposable cartridge designed to automate protein precipitation for SDS depletion, together with in-cartridge digestion and subsequent peptide enrichment. Each component of the workflow (proteome extraction, purification & recovery through precipitation, digestion efficiency and peptide isolation) will be evaluated relative to the resulting MS precision. Methods: Samples of bovine liver were extracted with varying levels of SDS (0, 1, 5%), and independently processed in triplicate using the ProTrap XG cartridge. Protein recovery, residual SDS, and peptide yield following digestion are reported. Reproducibility was evaluated following a bottom-up workflow across technical replicates vs. replicate MS injections, comparing ID counts and intensity. Preliminary Results: Inclusion of SDS is favorable for complete proteome extraction of bovine liver. However, acetone precipitation of samples with 1% SDS increases recovery, with 45% more proteins (avg 8103) and 23% more peptides (29478) vs extraction at 5%. More importantly in 1%, RSD drops from >50% to 9% at the peptide level, and from 28 to 4% at the protein level. MS replicates of an identical sample provided 7% absolute difference in PSMs between runs, with 78 ± 0.5% overlap in identified proteins. Adding the complete ProTrap-based workflow, the agreement in individual peptide PSMs across runs remained within 9%, with 76 ± 0.5% overlapping identification. Conclusions: The ProTrap XG provides a highly reproducible workflow for automated protein sample preparation for bottom-up analysis.

Contact: Alan Doucette, alan.doucette@dal.ca

P01.08 Automated S-Trap Sample Preparation on an Integrated Liquid Handling / Pneumatic Sample Processing Workstation for MAM Applications

Heather Eastwood (1), John Laycock (1), Claudia Martelli (2), Manuel Bauer (2), John Wilson (3), Xin Zhang (4), James Teuscher (1), Shang Tsai (5)
(1) Tecan, Baldwin Park, California, United States
(2) Tecan, Männedorf, Switzerland
(3) ProtiFi, LLC, Farmingdale, New York, United States
(4) Sciex, Redwood City, California, United States
(5) Tecan, Baldwin Park, California

Purpose Multiple Attribute Monitoring (MAM) is an LC-MS technique designed to simultaneously and directly monitor critical quality attributes (CQAs) of biologics including impurities, post-translational modifications (PTMs) and primary sequence. The highly varied nature of biotherapeutic manufacturing from bioreactors containing surfactants to final formulations presents analytical challenges. Here we present a universal, automatable MAM kit capable of removing contaminants from salts to surfactants to excipients to dyes in a streamlined workflow usable without change at all stages from bioreactor to final product. Methods A MAM-specific kit and automated workflow based on the S-Trap 96-well plate was optimized to the specific needs and requirements of MAM at all stages of biotherapeutic manufacturing. The kit was stress tested using the NISTmAb RM 8671 monoclonal antibody spiked into various PEG or surfactant containing solutions. MAM sample preparation was performed with full automation on the Tecan Freedom EVO® liquid handler with an integrated A200. Methods and scripts for (semi-)automated analyses are included in the ProtiFi/Tecan MAM workstation configuration. Samples were analyzed on a Sciex ZenoTOF 7600 in CID and EAD modes for comparative analysis. Results The MAM-sample processing kit was effective without protocol modification at removing all contaminant challenges including salts, surfactants, emulsifiers, dyes and other small molecules; this performance matches previous studies in the analysis of surfactant-containing bioreactor supernatants. Excellent reproducibility was observed for both unmodified and PTM-modified peptides with detection of the expected glycosylations of NISTmAb RM 8671. Residual host cell proteins (HCPs) were also

detected. Conclusion Quantification variability from sample preparation and especially automated sample preparation was similar to levels of variability observed in technical replicate injections. The MAM automation workflow and the corresponding kit has utility at all levels of therapeutics manufacturing, from initial R&D and characterization through process control, formulations, and QC and release testing.

Contact: Heather Eastwood, heather.eastwood@tecan.com

P01.09 Multi-nanoparticle Workflow Enables Deep Plasma Proteomics at Scale, with Enhanced Precision, and Depths of Coverage.

Shadi Ferdosi (1), Tianyu Wang (1), Renata Blatnik (2), Moaraj Hasan (3), Jessica Chu (3), Eltaher M. Elgierari (3), Xiaoyan Zhao (3), Martin Goldberg (4), Asim Siddiqui (3), Serafim Batzoglou (4), Michael Krawitzky (2), Nagarjuna Nagaraj (2), Omid C. Farokhzad (3), Daniel Hornburg (3)
(1) Seer, Redwood City, California, United States
(2) Bruker Daltonics, Bremen, Germany
(3) Seer, Redwood City, United States
(4) Seer, Redwood City

To overcome limitation of deep plasma proteomics in large cohorts, we have developed a fast and scalable technology that employs intricate protein-nano interactions. Introducing a nanoparticle (NP) into a biofluid such as blood plasma leads to the formation of a selective, specific, and reproducible protein corona at the nano-bio interface driven by the relationship between protein-NP affinity, protein abundance and protein-protein interactions. We previously demonstrated that this process, incorporated within the Seer Proteograph™ Product Suite, offers superior performance in terms of depth, breadth, precision, and throughput compared to conventional deep workflows. The ratio of plasma-to-nanoparticles determines the competition between proteins for binding surface, which plays an important role in protein corona composition and can be optimized to enhance and differentiate protein selectivity. Here we investigate effects of different conditions on protein corona composition enabling enhanced performance of Proteograph. We have investigated compositional changes of protein coronas from 5 NPs with blood plasma at different ratios. Samples were analyzed with timsTOF Pro mass spectrometry and UltiMate3000 Dionex LC system using 30min DIA runs. We evaluated depth, dynamic range, coverage, and precision of quantification at a wide range of concentrations for each NP. By limiting the available binding surface of NPs and increasing the binding competition, we are able to identify 20 – 60% more proteins on the surface of each NP. Moreover, by increasing the competition the proteins are more reproducibly identified and quantified across the replicates of the same NP. In addition, protein selectivity was enhanced, leading to improved coverage of plasma proteome when using multiple physicochemically distinct NPs. In summary, NP panels with optimized workflow, capture a large and diverse set of proteins and biological pathways based on their specific physicochemical makeup.

Contact: Shadi Ferdosi, sferdosi@seer.bio

P01.10 The Next Generation Infrared Matrix-Assisted Laser Desorption Electrospray Ionization (NextGen IR-MALDESI) Source

Kevan Knizner (1), Jacob Guymon (2), Kenneth Garrard (3), Guy Bouvree (4), Jan-Peter Hauschild (5), Kerstin Strupat (5), Kyle Fort (5), Lee Earley (6), Eloy Wouters (7), Fan Pu (8), Nathaniel Elsen (9), Jon Williams (9), Mark Pankow (2), David Muddiman (10)

(1) FTMS Laboratory for Human Health Research, Department of Chemistry, North Carolina State University, Raleigh, North Carolina, United States
(2) Precision Engineering Consortium, Department of Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, North Carolina, United States
(3) FTMS Laboratory for Human Health Research, Department of Chemistry, North Carolina State University; Precision Engineering Consortium, Department of Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, North Carolina, United States
(4) GB Conseil & Services, Brie-Comte-Robert, France
(5) Thermo Fisher Scientific, Bremen, Germany
(6) Thermo Fisher Scientific, San Jose, California
(7) Thermo Fisher Scientific, San Jose, California, United States
(8) Drug Discovery Science and Technology, Abbvie Inc., North Chicago, Illinois
(9) Drug Discovery Science and Technology, AbbVie Inc., North Chicago, Illinois, United States
(10) FTMS Laboratory for Human Health Research, Department of Chemistry, North Carolina State University; Molecular Education, Technology, and Research Innovation Center (METRIC), North Carolina State University, Raleigh, North Carolina, United States

Mass spectrometry imaging (MSI) is an effective analysis method for determining the spatial distribution of molecules within biological tissues. Infrared matrix-assisted laser desorption ionization mass spectrometry (IR-MALDESI-MS) is a versatile MSI platform that has been utilized to image diverse biological samples without tedious sample preparation. To improve IR-MALDESI-MSI analyses, user-friendliness of the source, and overall versatility of the platform, the key components of the source have been upgraded. Here, we present the next generation (NextGen) IR-MALDESI source which features a vertically mounted IR laser, planar translation stage with computer-controlled z-axis sample stage, external source component adjusters, redesigned source-to-mass spectrometer interface plate, and reduced plastic use. Also, the NextGen IR-MALDESI source can be easily interfaced with numerous Orbitrap mass spectrometers to accommodate a greater number of research labs. In this work, we present the characterization of the NextGen IR-MALDESI source to establish it as an improved platform for MSI and direct analysis applications.

Contact: Kevan Knizner, ktknizne@ncsu.edu

P01.11 A Universal and High-Throughput Proteomics Sample Preparation Platform

Christopher A. LeClair (1), Andrew P. Burns (1), Ya-Qin Zhang (1), Tuan Xu (1), Zhengxi Wei (1), Qin Yao (2), Yuhong Fang (1), Valeriu Cebotaru (2), Menghang Xia (3), Matthew D. Hall (1), Ruili Huang (3), Anton Simeonov (1), Dingyin Tao (1)
(1) NCATS/NIH, Rockville, Maryland, United States
(2) University of Maryland School of Medicine, Baltimore, Maryland, United States
(3) NCATS/NIH, Rockville, Maryland

Proteomics has become an essential tool to address critical biological questions through the identification and characterization of total protein content within a biological system. The discovery of distinct protein biomarkers for specific disease targets can aid in the development of therapeutic treatments. Major advances have been made to improve the sensitivity of mass analyzers, spectral quality, and the speed of data processing enabling more comprehensive proteome discovery and quantitation. While focus has recently begun shifting toward robust proteomic sample preparation efforts, high-throughput proteomics sample preparation capabilities has been lacking. This led us to develop a highly automated

universal 384-well plate sample preparation platform with high reproducibility and adaptability for the extraction of proteins from cells within a culture plate. Digestion efficiency was excellent in comparison to a commercial digest peptide standard with minimal sample loss while sample preparation throughput was improved by 20- to 40-fold (the entire process from plated cells to clean peptides is complete in ~300 min). Analysis of six human cell types including two primary cell samples, identified and quantified ~4,000 proteins for each sample in a single HPLC-MS/MS injection with only 100 – 10,000 cells, thus demonstrating universality of the platform. A selected protein was further quantified with a developed HPLC-MRM method for HeLa digests utilizing a spiked in heavy-labelled internal standard peptide. Excellent linearity was achieved across different cell numbers indicating a potential for target protein quantitation in clinical research. These reported advancements in high-throughput proteomics sample preparation brings us a step closer to realizing a fully automated proteomics screening platform thus enabling an efficient large-scale characterization of diverse biological signatures across a wide array of conditions and parameters.

Contact: Christopher LeClair, leclairc@mail.nih.gov

P01.12 Real-Time Database Search Multiplex Quantitative Proteomics: Instrument Method and Data Processing

Yang Liu (1), Frank Berg (2), William D. Barshop (1), Jesse D. Canterbury (1), David M. Horn (1), David Bergen (1), Romain Huguet (1), Rosa Viner (3)
(1) Thermo Fisher Scientific, San Jose, California, United States
(2) Thermo Fisher Scientific, Bremen, Germany
(3) Thermo Fisher Scientific, San Jose, California

Introduction Real-time search (RTS) on the Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer has enabled the capability to triggered SPS-MS3 scans only upon confident peptide spectrum matches (PSMs) identified from MS2 spectra via online database search using Comet. Here we introduce RTS TMT quantitation workflow using the latest version of instrument control software (ICSW) and the implementation of Comet in Thermo Scientific™ Proteome Discoverer™ Software (PD) 3.0 to provide the best alignment between online and post-acquisition data analysis. Method For RTS experiments and its extended features, 500ng Thermo Scientific™ Pierce™ TMT11plex Yeast Digest Standard or TMT pro 18 plex Yeast Digest Standard (prototype) was analyzed by an Orbitrap Eclipse Tribrid mass spectrometer (ICSW 3.5). MS^2 spectra were searched against a yeast proteome database during acquisition using the Comet search algorithm (2019.01 rev.1). The data were analyzed with PD 3.0, using both Comet and SequestHT search engine. Result RTS TMT SPS MS^3 strategy was able to quantify 40-50% more protein groups and 30-40% more peptide groups, compared the SPS-MS^3 method. The close-out feature can further boost the number of quantified protein groups without sacrificing quantification accuracy. The online FDR estimation can improve the percentage of quantified protein groups. In both SequestHT and Comet search in PD, parameters were matched to the RTS Comet settings. 98% of PSMs that were confirmed by RTS Comet search were confidently identified in the post-acquisition analysis using PD Comet algorithm. By combining both SequestHT and Comet in PD processing workflow, the number of identification and quantified IDs can be improved by 10-15%. Conclusion The RTS-MS3 workflow using the latest ICSW, together with post-acquisition data analysis optimization using Comet node in PD 3.0, largely improved the performance of the multiplex quantification proteomics.

Contact: Yang Liu, yang.liu6@thermofisher.com

P01.13 Benchmark of Micro-flow Chromatograph for Robust Proteomics Analysis

Yang Liu (1), Amirmansoor Hakimi (1), David M Horn (1), Runsheng Zheng (2), Oleksand Boychenko (2), Daniel Lopez-Ferrer (1)
(1) Thermo Fisher Scientific, San Jose, California, United States
(2) Thermo Fisher Scientific, Germering, Germany

Introduction Liquid chromatography-mass spectrometry (LC-MS) has been a powerful analytical tool in protein identification and quantification. In the past few decades, nano-flow LC-MS has been the primary approach due to its high sensitivity. However, challenges always come from the needs of high throughput, reproducibility and robustness. Here we present a micro-flow LC-MS workflow using a robust setup with Thermo Scientific™ NG micro-flow UHPLC System coupled to Thermo Scientific™ Orbitrap Exploris™ platforms. Methods Thermo Scientific™ Pierce™ HeLa protein digest standard and Thermo Scientific™ Pierce™ TMT-11plex Yeast digest standard were analyzed on the Orbitrap Exploris 240 MS for LFQ analysis. Digested peripheral blood mononuclear cells (PBMCs) from a variety of animal species (human, mouse, etc) were analyzed to demonstrate the robustness over 100 injections. Thermo Scientific™ High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Pro Duo interface was installed to provide gas phase fractionation. Data was analyzed on Thermo Scientific™ Proteome Discoverer™ 3.0 software using MSPepSearch and CHIMERYS in parallel. Result We were able to identify ~3400 protein groups and ~27500 peptide groups from 1ug of HeLa digest, ~3800 protein groups and ~34000 peptide groups from 5ug of HeLa digest within 30min gradient. The micro-flow LC-MS system showed excellent reproducibility of protein group IDs (<3% coefficient of variation, CV) and protein group abundance (median CV ><11%). In multiplex quantitation 90% of identified proteins and peptides were successfully quantified. Digested PBMC peptides were fractioned by 6 compensation voltages using FAIMS Pro Duo interface and were acquired in triplicate on the Orbitrap Exploris 480MS. GPF provides deep proteome profiling without the need for off-line RPLC fractionation, which reduced the overall experiment time. Conclusion This micro-flow LC-MS setup has been demonstrated to be highly reproducible and robust without sacrificing performance for both discovery and quantitation.>

Contact: Yang Liu, yang.liu6@thermofisher.com

P01.14 DIA Analysis in Discovery Identifications and Label-free Quantification on Orbitrap Exploris 240 Mass Spectrometer

Yang Liu (1), Khatereh Motamedchaboki (1), Sega Ndiaye (2), Aman Makaju (1), Amarjeet K. Flora (3), Ryan Bomgarden (3), Daniel Lopez-Ferrer (1)
(1) Thermo Fisher Scientific, San Jose, California, United States
(2) Thermo Fisher Scientific, Courtaboeuf, France
(3) Thermo Fisher Scientific, Rockford, Illinois, United States

Introduction Label-free quantitation (LFQ) is the simplest way to explore the proteome. Data-independent acquisition (DIA) mass spectrometry (MS) can rapidly identify and reproducibly quantify all ions without the bias from peak intensity, providing deep proteome analysis for large-scale profiling and quantitative studies. In this work, we present a MS-only workflow that combines gas-phase fractionation (GPF) and DIA acquisition, saving significant experiment time while maintaining high reproducibility Methods Different amount of digested E.coli peptides (in ratio 0:1:2:4:8) were added to a fixed amount of HeLa digest peptides. Mixed peptides were loaded on 25cm Aurora Series emitter column (25 cm*75 µm ID, 1.6 µm, C_{18}), separated by a 90 min LC gradientat in direct injection mode on an EASY-nLC 1200 system before being injected onto the Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer. Empirically corrected spectral library was built by matching

the gas-phase fractionations (GPF) DIA chromatogram spectral library to the predicted spectral library of HeLa and E. coli tryptic peptides[1]. Acquired DIA data were analyzed by Spectronaut™ 14.0. Results Over 9,000 proteins were identified from ~1 µg of the mixed proteome sample when matching to the empirical spectral library. Over 75% of the proteins were quantified in each of 15 LC-MS runs. In addition, the Orbitrap Exploris 240 mass spectrometer achieved over 6 orders of magnitude in protein dynamic range. Furthermore, we also proved that matching with the library built at a higher sample load would improve the identification coverage and confidence of identifications. Therefore, once built at ultimate setup, the spectral library can be used for universal DIA discovery of the same sample. Conclusions Orbitrap Exploris 240 MS had excellent performance in DIA analysis in discovery and quantitation. References 1. Searle, B.C., Swearingen, K.E., Barnes, C.A. et al. Nat Commun 11, 1548 (2020).

Contact: Yang Liu, yang.liu6@thermofisher.com

P01.15 Semi-automated and high-throughput homogenization technique for in-depth analysis of various tissue proteomes

Kristine Bissell (1), Berit Mang (1), Katrin Hartinger (1), Sebastian H. Johansson (1), Jasmin Johansson (1), Nils A. Kulak (1)
(1) PreOmics GmbH, Planegg, Germany

Purpose: Efficient protein extraction is a crucial and challenging step in tissue sample preparation for LC-MS analysis. While conventional methods are faced with various challenges - such as low-throughput and time-consuming workflows - our novel tissue lysis workflow on the semi-automated BeatBox platform enables efficient protein extraction for 96 samples in parallel in as little as 10 minutes. Coupled to the robust iST technology that is fully automatable on common liquid handling platforms, we combined in-depth sample preparation for large-scale applications with minimal hands-on time. Methods The described workflow is based on 96 well sample processing on the BeatBox platform and applicable for a plethora of tissue types ranging from soft brain to rigid heart muscle samples. Utilizing innovative bead-based technologies, a defined energy input is applied to each sample facilitating highly efficient protein extraction. For complete nanoLC-MS sample preparation, we combined the BeatBox-based tissue lysis with the iST workflow. Improved proteomic depth could be achieved by a 3-step peptide fractionation. Results The BeatBox platform speeds up and simplifies the tissue preparation for proteomic experiments. From intact tissue sample to finished data acquisition (DDA on timsTOF Pro, Bruker) in less than 4 hours, we identified ~2300 proteins from mouse heart muscle, ~3000 proteins from mouse liver tissue and >4000 proteins from mouse lung tissue. In comparison to the traditional iST workflow, we gained up to 13 % more protein identifications. At the same time, excellent technical variability (e.g. for mouse liver tissue: median CV of 9.2 %) was obtained. Furthermore, a combination of the BeatBox-iST workflow with tip-based peptide fractionation increased the protein identifications by over 40 % compared to unfractionated samples. Conclusion: The innovative BeatBox-based workflow will set a new standard in large-scale tissue sample preparation by enabling ultra-fast and highly efficient protein extraction in a high-throughput manner.

Contact: Berit Mang, mang@preomics.com

P01.16 Multi-omics sample preparation workflow for DNA and proteins from a single starting sample using ProMTag

Amber Lucas (1), Stephanie Biedka (1), Jonathan Minden (1)
(1) Impact Proteomics, Pittsburgh, Pennsylvania, United States

Extracellular vesicles (EVs) are complex, cell-derived nanoparticles generated by all cell types. EVs are composed of lipid bilayer membranes and their associated membrane proteins, nucleic acids, and luminal proteins. The mechanism by which Gram-positive bacteria shed EVs is still unknown. EVs from the Gram-positive human pathogen S. pneumoniae, which is a major cause of otitis medi and pneumonia, are of particular interest because of how they EVs modulate the host immune response. To uncover possible mechanisms for EV production and shedding in S. pneumoniae, we have performed a comparative proteomics analysis of EV membrane proteins versus whole-cell membrane proteins.

Membrane proteins were enriched from intact S. pneumoniae cells or their EVs using a ProMTag labeling and capture workflow. ProMTag is a bifunctional protein tag where one moiety of the tag is able to form a reversible, covalent link to primary amines on proteins. The other moiety is methyltetrazine, which can form an irreversible, covalent bond with trans-Cyclooctene (TCO) on the surface of beads to capture ProMTagged proteins for cleanup and elution. Using this workflow plasma membrane proteins can be tagged, captured, washed to remove non-plasma membrane proteins, and then eluted in their original, unmodified state.

In this study, intact cells and EVs from S. pneumoniae cultures were separated and the extracellular domains of membrane proteins in these two fractions were labeled with ProMTag. The membrane proteins were then enriched, washed, and eluted using the ProMTag workflow. These membrane protein populations were then analyzed using mass spectrometry to identify what proteins were present in EV membranes and how that compared to the composition of whole cell membranes and whole EVs. Analysis revealed membrane proteins that are concentrated or absent in EV membranes relative to bulk plasma membrane from whole cells, indicating a selective process for EV formation in S. pneumoniae. With this information, we present a new model for EV formation and shedding in S. pneumoniae.

Contact: Amber Lucas, amber.lucas@impactproteomics.com

P01.17 Robust Identification of Low Abundant Proteins Utilizing a Precipitation Workflow on the ProTrap XG

Victoria A Miller (1), Sara Lahsaee Little (1), Jessica L Nickerson (2), Alan A Doucette (2)
(1) Proteoform Scientific Inc, Halifax, Nova Scotia, Canada
(2) Dalhousie University, Halifax, Nova Scotia, Canada

In biomarker discovery, conclusions are dependent on the quality of the data, which is contingent on the integrity of sample preparation. Efficient recovery of all proteins, regardless of abundance, is necessary to accurately represent the proteome. The ProTrap XG is a filter-based sample preparation cartridge that simplifies the traditional precipitation workflow. Sample preparation conditions for the ProTrap XG were optimized to deliver highly efficient and robust identification of low abundance proteins from a complex system. The effect of 2% SDS on the precipitation and recovery of 3 standard proteins, spanning 4 orders of magnitude: Beta-galactosidase (50 µg), cytochrome c (50 ng) and enolase-1 (5ng) were combined in the presence of 2% SDS. 50 ug of the mixture was subjected to acetone precipitation followed by resolubilization and trypsin digestion in the ProTrap XG. A control (no SDS) was also digested in solution. Bottom-up analysis on a Thermo Orbitrap Fusion Lumos with easy nLC identified all 3 proteins with equivalent sequence coverage in the SDS-based workflow relative to the control. Moreover, the SDS workflow yielded higher peptide abundance for Beta-galactosidase (4x) and cytochrome c (2x) vs. the control. Bovine liver was extracted in 1% SDS and 50 µg protein was processed in the ProTrap XG. Bottom-up MS yielded on average ~8000 proteins and ~30,000 peptides. Several low abundance proteins,

such as transcription factors (e.g. NFκB) were identified. Further data analysis is ongoing. A detergent-based workflow for proteome processing in the ProTrap XG is therefore shown to improve the MS identification of low abundance proteins in biological samples.

Contact: Victoria Miller, vmiller@proteoform.com

P01.18 Establishing Quality Control Procedures for Sample Preparation of High-Throughput Plasma Proteomics

Nekesa Oliver (1), Min Ji Choi (1), Albert Arul (1), Renã Robinson (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

Clinical biomarker discovery and verification that uses mass spectrometry (MS) –based proteomics must have workflows that are dependable and reliable. However, bottom-up plasma proteomic workflows can suffer from reproducibility issues in large cohorts (N>100) due to losses that occur with numerous sample preparation steps such as protein depletion, tryptic digestion, chemical labeling, sample cleanup and peptide separation. Automation of sample preparation workflows with a robotic liquid handler can alleviate these losses, manage the workflow, and increase reproducibility. Designing quality control (QC) measures is key to ensuring reliable MS data is obtained. This study of 800 plasma samples assesses robust quality control decision checkpoints for an automated proteomic preparation workflow using a robotic liquid handler. TMTpro isotopic labeling is used to increase multiplexing capabilities while an equimolar sample pool labeled with TMTzero is prepared as an instrument QC for liquid chromatography (LC)-MS/MS analysis. Across 1600 sample MARS-14 depletions (800 in duplicate), QC samples gave 2.91 and 0.68 CV% for unbound and bound peak retention times, respectively. Plasma sample concentrations, determined from bicinchoninic acid assay (BCA), varied less than 10% across this set of samples while albumin standards had less than 5% variation. Plasma digests randomly selected along with QCs for shotgun proteomic analysis resulted in protein identifications with less than 5% CV. Here, we will present preliminary data of QC performance checks of sample preparation aspects of the automated workflow and provide recommendations for large-scale studies. Overall, defining quality control decision points for a high-throughput and automated workflow is highly valuable for robust plasma proteomics in large-scale proteomics.

Contact: Nekesa Oliver, nekesa.c.oliver@vanderbilt.edu

P01.19 Establishing Quality Control Procedures for Large-Scale Discovery Based Proteomics Analysis of Human Plasma

Khiry L. Patterson (1), Albert B. Arul (1), Min Choi (1), Nekesa C. Oliver (1), Shania Hansen (2), Angela L. Jefferson (2), Timothy J. Hohman (2), Logan C. Dumitrescu (3), Renã A. S. Robinson (1)
(1) Vanderbilt University, Nashville, Tennessee, United States
(2) Vanderbilt University Medical Center, Nashville, Tennessee, United States
(3) Vanderbilt University Medical Center, Nashville, Tennessee

Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) platforms for large-scale clinical proteomics enable analysis of complex proteomes, where thousands of proteins can be quantified in biological samples. Datasets generated from such analyses are then used to evaluate significant biological findings relevant for disease, assuming that changes associated with disease are real and not a reflection of experimental variability. Experimental variability can occur at any step in the proteomics workflow (e.g., sample preparation, fractionation, and LC-MS/MS analysis); therefore, stringent

quality control (QC) protocols are needed to ensure optimal instrument performance, especially when the number of samples exceeds 100. While use of robotic platforms and sophisticated instruments can reduce variability during sample preparation and fractionation, daily consideration of LC-MS/MS metrics can inform optimal, “real-time” chromatographic and instrument performance needs. The goal of this study is to evaluate and determine what appropriate LC-MS/MS metrics should be considered on a daily basis when performing large-scale plasma proteomics analyses. Using a proteomic dataset of ~330 human plasma samples, we performed two independent QC protocols: 1) internal QC checks within tandem mass tag batches and 2) a daily QC check before sample injections. The observed coefficient of variation (CV %) for protein and peptide identifications was <5 and ~8 CV %, respectively, and similar for peptide spectral matches. Other metrics such as peptide retention times, signal intensities of precursors, and intensities of TMT reporter ions were also used to evaluate LC and instrument daily performance. This poster will present performance analysis of QCs in this plasma proteomics dataset, considerations for improving LC-MS/MS instrument performance, and recommendations for daily QC monitoring in large-scale plasma proteomics experiments.>

Contact: Khiry Patterson, khiry.l.patterson@vanderbilt.edu

P01.20 High throughput single-shot proteomics on the timsTOF Pro 2 with DIA-PASEF and DDA-PASEF acquisition mode

Francesco Pingitore (1), Verena Tellstroem (2), Kristina Marx (2), Dirk Wunderlich (2), Nagarjuna Nagaraj (2), Gary Kruppa (3)
(1) Bruker, San Jose, California, United States
(2) Bruker Daltonics GmbH & Co KG, Bremen, Germany
(3) Bruker s.r.o., Brno, Czech Republic

Introduction Successful proteomics research requires the combination of rapid and in depth coverage along with quantitative accuracy with high sample throughput in applications like clinical and personalized medicine proteomics. TimsTOF Pro 2 enables sensitive and accurate proteomic analysis from a variety of samples including cell lysates, tissues, pulldowns and body fluids in a robust fashion. Here we show the performance of the instrument to quantify cell line proteomes in different gradient lengths and discuss the best usage of the instrument for a versatile use in biological experiments. Methods Digested HeLa peptides (Pierce), digested K562 peptides (Promega) and inhouse prepared digests from HEK cell lines were used for benchmark experiments. Peptides were directly loaded on a Aurora-25 cm column using nanoElute coupled to a timsTOF Pro 2 instrument via a Captive Spray ionization source. Data were acquired using DDA- and DIA-PASEF methods with different parameters tested for best performance. Data were directly streamed to PaSER box for all DDA data unless otherwise specified and processed offline in MaxQuant. For DIA, data were processed in Spectronaut (v15). Data were filtered at a FDR of 1% for peptide and protein groups. Preliminary Data Preliminary analyses were performed with 200 ng of HeLa or K562 peptides and measured on 60-minute gradient. These runs typically resulted in identifications in the range of 6000 protein groups (PG). Using inhouse digested HEK peptides with an improved sensitive sample preparation protocol, we could quantify in the range of 7000 PG using DDA methods. With DIA analysis, HEK peptides resulted in about 8000 protein groups and more than 70000 unique peptide sequences. Measuring 20 ng of peptides in relatively shorter gradients resulted in more than 3500 PG. Novel Aspect Rapid and sensitive quantification of about 7000 protein groups in single-shot injections.

Contact: Francesco Pingitore, Francesco.Pingitore@Bruker.com

P01.21 Analyzing protein fluorosequencing data, a new technology for single molecule proteomics

Matthew Smith (1), Edward Marcotte (2)
(1) Oden Institute, The University of Texas at Austin, Austin, Texas, United States
(2) Department of Molecular Biosciences, The University of Texas at Austin (COI: E.M.M. is a co-founder, shareholder, and SAB member of single molecule protein sequencing company, Erisyon Inc.), Austin, Texas, United States

Purpose Tools for protein identification and quantification lag DNA and RNA sequencing techniques in sensitivity and throughput, pushing proteomics researchers to pursue new high-performance approaches. To address these issues, our group invented fluorosequencing, a single molecule protein sequencing technology that incorporates features of nucleic acid sequencing for proteomics. In fluorosequencing, proteins are proteolytically digested into peptides, and specific amino acids are labeled with fluorescent dyes. Labeled peptides are immobilized in a flow-cell where, using Edman degradation chemistry, they are sequenced in parallel while being imaged by single molecule microscopy. Fluorosequencing produces sequencing reads from many individual molecules simultaneously, with the expected elevation in noise and errors that must be addressed in subsequent computational analysis. Methods By modeling sequencing errors (failed chemistry, labeling, dye bleaching, etc) on large synthetic datasets, we considered four machine learning approaches to accurately assign fluorosequencing reads to parent peptides: (1) the k-Nearest-Neighbor method, (2) a Random Forest classifier, (3) Bayesian classification based on Hidden Markov Models of the chemical processes used in sequencing, and (4), a hybrid approach combining the k-Nearest-Neighbor method with the Hidden Markov Models. Results For small sets of proteins, RF, HMM, and kNN+HMM classifiers significantly outperformed kNN for classification precision and recall, with the HMM based classifier giving the best results. However, at human proteome scale, involving hundreds of thousands of possible peptides, RF and HMM models become intractable. We found that kNN+HMM offers a good compromise, scaling to the full proteome and significantly outperforming kNN. Conclusions A hybrid kNN+HMM machine learning strategy successfully assigns fluorosequences to their parent peptides at good precision and recall, while also scaling to the full human proteome and modeling known sources and rates of fluorosequencing errors in a human interpretable form.

Contact: Matthew Smith, mbsmith93@utexas.edu

P01.22 Universal Hydrophilic affinity SPE based peptide clean-up workflow for parallel sample processing

Stoyan Stoychev (1), Ireshyn Govender (1), Previn Naicker (1), Isak Gerber (1), Justin Jordaan (1)
(1) ReSynBio, Pretoria, South Africa

Several methods for proteome sample preparation exist that deal with upfront protein capture, clean-up and digestion. These include classical affinity-based precipitation, in-solution digestion, FASP and more recent on-membrane and on-bead precipitation workflows. However, on peptide level C18-based desalting is still the go-to-method for peptide clean-up. While this approach is efficient, it suffers from well-documented drawbacks such as the inability to deal with/eliminate contaminants including detergents and polymers as well as the requirement to remove (or at the very least reduce) the high concentration of the organic solvent utilised for peptide elution. Here we present a workflow that can be coupled for universal clean-up prior to LC-MS analysis, using semi-porous hydrophilic affinity material (ReSyn HILIC) in a 96 well plate format for parallel sample processing in a high-throughput manner. The performance was compared to traditional C18 SPE using colorimetric peptide quantification to estimate the recoveries. LCMS-analysis was further applied to assess potential bias in peptide capture and release, as well as the overall proteome coverage. The material allows for efficient capture of

peptides under high organic conditions and the efficiency to remove salts, polymers and detergents from peptide samples will be showcased.

Contact: Stoyan Stoychev, sstoychev@resynbio.com

P01.23 EPITOME ANALYSIS OF THE HUMAN PLASMA PROTEOME DELIVERS DIAGNOSTIC AND THERAPEUTIC CANDIDATES.

Laszlo Takacs (1), Jozsef Lazar (1), Peter Antal-Szalmas (2), Istvan Kurucz (1), Barbara Uzonyi (3), John Lamont (4), Janos Fekete (5), Ilona Tornyi (6), Alexandra Pap (7), Alexandra Matola (8), Mihaly Jozsi (3)
(1) Biosystems Immunolab Ltd, Debrecen, Hungary
(2) Univesity of Debrecen, Debrecen, Hungary
(3) Department of Immunology, Eötvös Loránd University, Budapest, Hungary
(4) 4Randox Laboratories Ltd, Crumlin, United Kingdom
(5) Adware Research Kft, Balatonfured, Hungary
(6) Biosystems Immunolab Ltd, Debrecen
(7) Department of Immunology Eotvos university, Budapest, Hungary
(8) Department of Immunology, Eötvös Loránd University, Budapest

Current proteomics technologies focus solely on the quantitative determination of individual proteins, practically no efforts have been dedicated to the development of proteomics systems tools to simultaneously monitor proteome variability, accessibility and abundance. Here, we present a robust and analytically validated mAb proteomics technology for profiling epitopes of the plasma proteome. To this end we developed mAb libraries, Quantiplasma™ against natural, accessible and immunogenic plasma protein epitopes. We show that biomarker epitope panels represent a rich and thus far unexplored source of protein biomarkers with diagnostic potential. We developed, validated and registered Pulmolisa™ as a tool to aid differential diagnostic of lung cancer. Additionally, we tested 156 complement component (C1q, C3, C4, C5, C6, C7, C8, C9, C4b-binding protein (C4BP) or factor H (FH)) epitope specific mAbs in functional assays and found 42 with biological activity (inhibition or potentiation). In summary, we validate a novel and expandable proteomics technology that goes beyond the current limitations of MS tools, apparently to the level of epitopes. The number of protein coding genes is far exceeded by the number of protein variants that become accessible via epitome analysis. The high ratio of in-vitro functional activity of the epitope specific mAbs support strongly that epitome profiling, done with BSI's natural protein immunization approach, monitors biologically relevant differences, with direct applicability as diagnostic tools and as therapeutic candidates.

Contact: Laszlo Takacs, laszlo.takacs@biosys-ilab.com

P01.24 Giving CE-MS a New Life in Proteomics: Integrated Solid-Phase Extraction and Detergent Tolerance

J. Will Thompson (1), J. Scott Mellors (2), Joshua P. Gurrette (3), Erin A. Redman (1), Christopher D. Brown (4)
(1) 908 Devices Inc., Boston, Massachusetts, United States
(2) 908 Devices Inc., Boston, North Carolina, United States
(3) 908 Devices Inc, Boston, Massachusetts, United States
(4) 908 Devices, Boston, Massachusetts, United States

Microchip CE-MS is an inherently low-volume separation which achieves high separation efficiency. It is capable of producing small 'peak volumes', meaning peptides may typically elute in about 1 to 10 nanoliters, resulting in the potential for a very sensitive separation. Additionally, peptide CE-MS separation can be performed from samples containing detergents and other carriers, which could be a distinct advantage for low-concentration samples. However, CE-MS traditionally suffers from poor concentration-sensitivity because of the small sample injections, which are on the order of 100-1000 times smaller than typical volumes for sample loading in capillary liquid chromatography. To address this microchip CE limitation, we have integrated a small solid phase extraction bed in the sample channel of the microfluidic device, allowing a large sample volume (up to 2 uL) to be loaded directly into the separation channel prior to CE-MS separation, a 400-fold volumetric loading improvement. The SPE-CE-MS method including a 0.5 mm bed packed with 5 um C18 particles demonstrates a capacity in excess of 100 ng peptide directly into the CE-MS channel. Using HeLa lysates and serum as model systems, we perform proteome separations in 11 minutes with sample concentrations below 0.2 ng/uL on a modified ZipChip CE coupled to an Exploris 240 MS. Additionally, CE-MS separations are detergent tolerant, so the SPE-CE-MS procedure can be performed with detergents and other carriers present in the sample. The ability to utilize carriers could be a major advantage for proteomics sample preparation of single cells, but has been generally avoided because of the inherent problems with reversed-phase LC intolerance to detergents. We observe up to 100-fold recovery improvements when samples are prepared and stored with nonionic surfactants present in the sample. Data will be presented on the relative merits and of various detergents, chaotropes, and small molecule carriers in the SPE-CE-MS separation.

Contact: J. Will Thompson, wthompson@908devices.com

P01.25 Quality Control of a DIA-MS High-Throughput Proteomics Workflow for CSF Samples from Parkinson's Disease Patients

Kirstin E. Washington (1), Rakhi Pandey (1), Danica-Mae Manalo (1), Alejandro Rivas (1), Connor Phebus (1), Matthew Ayres (1), Koen Raedschelders (1), Niveda Sundararaman (2), Vidya Venkatraman (1), Jennifer E. Van Eyk (2)
(1) Cedars Sinai Medical Center, Los Angeles, California, United States
(2) Cedars Sinai Medical Center, Los Angeles, California

The deployment of automated sample preparation using 96-well plate format, efficient LC-MS workflows and quality control measure is helpful for consistent large cohort proteomics analysis. As a part of the Accelerating Medicines Partnership in Parkinson's Disease (AMP-PD), 2284 human cerebrospinal fluid (CSF) samples from Parkinson's Disease Patients along with 624 Quality Control samples have been analyzed over 39 weeks. 4 digestion control reference (DCR) samples were included along with 75 patient samples per 96-well plate which were processed on the i7 automated workstation for protein denaturation (TFE), reduction (DTT), alkylation (IAA), and digestion (1:10 protein:trypsin; 4hrs; 42°C). A total of 39 x 96-well plates were processed and analyzed on a EVOSEP One system coupled to a Orbitrap Exploris 480 MS (Thermo) using DIA-MS (21 Da windows). Furthermore, a weekly MS-system suitability was carried out using a commercial Hela peptide digest reference standard and MS-technical control reference (TCR) samples consisting of a pooled pre-digested Parkinson's Disease CSF sample run 9 times per 96 plate. In total, 156 DCR, 117 HeLa digests (3 times per week) and 351 TCR were run. For efficiency, we automated QC reports that provided total # peptides and proteins with <20% and ><40% CV% for each 96-well plate. On average, the DCR and TCR samples quantified >13000 peptides corresponding to a total of 450 456 proteins with >70% of proteins having intraday CV% <40 across all 39 plates, while TCR samples quantified >600 peptides and 392 proteins, with 70% of proteins having CV% <40%. Variance analysis (PCA) of the DCRs and TCRs

across the 39 plates of PC1 contributing 12.45 and 11.31% and PC2 contributing 6.01 and 5.93%, respectively. Batch correction was undertaken using Random Forest further reducing this technical variance.>

Contact: Kirstin Washington, kirstin.washington@cshs.org

P01.26 Incorporation of glycoproteome detection into large scale unbiased proteomics studies utilizing nanoparticles

Bruce Wilcox (1), Kavya Swaminathan (1), John Blume (1), Jared Deyarmin (1), Preston Williams (1), Chinmay Belthangady (1), Manway Liu (1), Mi Yang (2), Philip Ma (1)
(1) Prognomiq, San Mateo, California, United States
(2) Prognomiq, San Mateo, California

Robust identification, characterization, and quantitation of glycosylated post-translational modifications (PTMs) is of key interest in disease pathogenesis as glycoproteins play vital roles in maintaining cellular health and glycoprotein dysregulation is directly associated with numerous diseases, including cancer. Historically, identification, characterization, and quantitation of the circulating glycoproteome from biofluids has posed challenges due to wide range of abundances, stability, heterogeneity, and data analysis limitations. While enrichment strategies yield deeper investigation of glycoproteins and glycoforms, experimental reproducibility and biases still pose considerable challenges. By leveraging Seer's Proteograph™ nanoparticle technology that enables robust and reproducible measurements across the plasma proteome, unique proteoforms are specifically captured in corona formation and allow for comprehensive assessment of the circulating glycoproteome. Furthermore, using Proteograph technology unlocks the possibility to monitor ratio abundances of glycosylated vs. native forms of proteins, which unveils a new dimension in glycoproteomics for the development of novel biomarkers. Using the five-nanoparticle panel workflow, we report the detection of distibuitions of glycoproteins across all five nanoparticles (up to 1% per nanoparticle) in a high-throughput cancer biomarker study of 212 subjects (96 non-cancer, 116 cancer). In this study, we identified a median 1,592 protein groups per subject, a total of 5,099 protein groups across all samples (process samples included) while maintaining median CV's ranging 20-25% across all 5 nanoparticles. With this technology, we detect a higher composition of glycoproteins relative to unenriched sample preparation strategies for circulating proteins in literature. We also describe a new approach to investigate glycoprotein abundance, glycoprotein/glycoform identification and characterization across a large cancer cohort study while overcoming the need for separate upfront sample enrichment. Our results highlight the application of Proteograph nanoparticle technology for deep, unbiased proteomics biomarker studies including the simultaneous enrichment of glycoproteins to further advance the detection of cancer biomarkers.

Contact: Bruce Wilcox, bruce.wilcox@prognomiq.com

P01.27 "BCA-No-More": Seamless, High-Throughput Protein Quantification Directly on S-Trap Plates

John Wilson (1), Aleisha Benjamin (1), Heather Eastwood (2), Darryl Pappin (3), John Laycock (4)
(1) ProtiFi LLC, Farmingdale, New York, United States
(2) Tecan, Baldwin, California, United States
(3) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States
(4) Tecan, Baldwin Park, California, United States

Purpose Recent advances in the throughput of analytical proteomics, now often requiring only minutes per sample for identification and quantification, necessitate concomitant advances in bottom-up sample preparation workflows. Specifically, the simplification and ideally elimination of any extraneous steps in a sample preparation workflow increases both throughput and robustness. Post lysis, one of the most standard steps preceding sample preparation is a protein concentration assay, often via BCA. While compatible with detergents such as SDS, all assays, including BCA, are not instant, consume some of your sample, and are subject to interference and, if performed in a 96-well plate, edge effects. Here, we demonstrate the new concept of direct quantification of cleaned, surface-bound protein on S-Trap 96-well plates using intrinsic tryptophan fluorescence. Methods S-Trap 96-well plates were made and used as per standard protocols. Samples of varied hydrophobicity were bound onto plates and washed as per standard protocols; standard BSA curves were also loaded. Tryptophan fluorescence in both a wet and dry state was measured with an excitation between 269 and 280 nm and an emission of 325 to 475 nm using a Tecan Spark plate reader in top-read emission mode. Protein concentrations determined via tryptophan fluorescence were compared to BCA for limit of detection, reproducibility, and dynamic range. Results Fluorescence was measured using a Tecan Spark plate reader for proteins already bound and washed on an S-Trap 96-well plate and provided with digestion buffer. Fluorescence tracked with protein load at 277 nm excitation and 350 nm emission in a roughly linear fashion up to 100ug load. The direct-determination method afforded protein quantification from 1ug-100ug in a significantly reduced time compared to BCA assays, without any loss of sample. Conclusion Direct determination of protein concentration with intrinsic clean-up removing the need for protein assays.

Contact: John Wilson, john@protifi.com

P01.28 Sample Preparation to Match Analytical Advances: 384-well S-Trap Plate

John Wilson (1), Brett Phinney (2), Michael Krawitzky (3), Darryl Pappin (4), Benjamin Orsburn (5)
(1) ProtiFi LLC, Farmingdale, New York, United States
(2) UC Davis, Davis, California, United States
(3) Bruker Scientific LLC, Billerica, Massachusetts, United States
(4) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States
(5) Johns Hopkins University School of Medicine, Baltimore, Maryland, United States

Purpose Recent advances in analytical proteomics throughput, necessitate concomitant progress in bottom-up sample preparation workflows. With the ability to handle extremely diverse samples without the need for protocol modification, the S-Trap sample preparation system has found widespread adoption. To date, S-Traps have been available as both spin columns of varying capacities and in 96-well plate format. To keep pace with ever-increasing throughput and single-cell analyses, we developed the new S-Trap 384-well plate, suited for protein loads from single cells and sub µg quantities to 100 µg. Methods 384-well S-Traps were manufactured to match the performance of S-Trap micros. Replicate sample preparations were performed on 384-well plates; S-Trap micro columns, with 96-well plates used as a baseline. The standard steps of lysis, reduction, alkylation, denaturation, binding, washing and digestion were performed. Well-to-well and plate-to-plate variation were compared via analysis by LC-MS. Results 384-well S-Trap plates performed consistently with the results of spin columns and 96-well plates. Well-to-well and plate-to-plate variation was equivalent to the reproducibility of replicate technical injections. With 1 hr digestion at 47 °C, an average sample processing speed of less than 10 sec/sample could be attained, a speed compatible with current rates of detection and quantification. Initial experiments in single-cell analyses indicate that the 384-well plate affords more accurate representation of the underlying biological states

by quenching biochemical reactions through direct dispensing of cells into 5% SDS. Conclusion Robust sample preparation suited for scales from single cell to 100 µg at an average of < 10 sec/sample.

Contact: John Wilson, john@protifi.com

P01.29 S-Trap Turbo: From Samples to Analysis in Record Time

John Wilson (1), Sandra Wilson (1), Darryl Pappin (2), Alexandre Zougman (3)
(1) ProtiFi LLC, Farmingdale, New York, United States
(2) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States
(3) University of Leeds, Leeds, United Kingdom

Purpose Recent advances in proteomics data acquisition throughput necessitate concomitant advances in bottom-up sample preparation. The elimination of any and all extraneous elements in a workflow increases both throughput and robustness. One of the more tedious steps in proteomics sample processing is post-elution sample dry-down. We present the new S-Trap turbo: yielding minimal elution volumes of highly concentrated peptides suited for immediate analysis. Methods New snap-cap S-Trap turbo micro columns were constructed via plastic injection molding. The columns incorporated newly developed, compressed polymeric traps derivatized with novel surface modifications. The standard S-Trap protocol steps of lysis, reduction and alkylation, denaturation, binding, washing and tryptic digestion were performed both for standard and turbo S-Traps. Samples were analyzed by LC-MS. Sample yield was compared and quantified using BCA and/or fluorescent assays. Sample quality was compared by peptide and protein identification rate and reproducibility of quantifications. Results Turbo traps were compared to traditional S-Traps using three samples of highly varied hydrophobicity between < 1 µg to 100 µg of protein was processed as described above. S-Trap turbo elutions as low as 5 µL were found to be reproducible and similar or better than standard S-Trap digestions as judged by completeness of digestion, peptide yield and identifications. S-Trap turbo elutions could be immediately loaded onto an autosampler. No significant loss of hydrophobic peptides between the standard S-Trap protocol and S-Trap turbo was observed. S-Trap turbo allows proteomics researchers to go from complex sample to ready-to-inject peptides in record time, with minimal steps and equipment. Conclusion A highly robust sample preparation system, which negates speedvacing and increases throughput.

Contact: John Wilson, john@protifi.com

P02: Aging and Neurodegenerative Diseases

P02.01 Deep and quantitative succinylation profiling from dietary treatment in liver

Conor Mullens (3), Joanna Bons (1), Jacob Rose (1), Kevin Peasley (2), Takuto Chiba (2), Francesco Pingitore (4), Christopher Adams (4), Eric Goetzman (5), Birgit Shilling (1)
(1) Buck Institute for Research on Aging, Novato, California, United States
(2) University of Pittsburgh, Pittsburgh, Pennsylvania, United States
(3) Bruker Datonic, San Jose, California, United States
(4) Bruker Daltonics, San Jose, California, United States
(5) University of Pittsburgh, Pittsburgh, Pennsylvania

Abstract: Protein succinylation is an unexplored PTM and thought to have dramatic consequences on protein structure and thereby function given the large size of the succinyl moiety (100 Da) and amino acid site selectivity (K) changing charge from +1 to -1. In this study, we combine succinylated peptide-level enrichment with timsTOF technology (DDA and DIA) to both identify sites of lysine succinylation and measure changes in succinylation upon dietary treatment in mouse liver tissues. Methods: Control and dietary-treated wild-type mouse livers (N=8) were homogenized, digested and enriched for succinyl modifications using the anti-succinyl (Succ-K) affinity motif from Cell Signaling Technologies. Peptides were run in both a DDA and dia-PASEF schema(s) on the timsTOF Pro using variable CE's and mobility windows in a 90 min. gradient and short 45 min. gradients. Data was analyzed by PaSER (DDA) and Spectronaut v15 in directDIA mode (dia-PASEF). Preliminary Data: The DDA search results produced in real time by PaSER identified more than 6,300 peptides of which greater than 3,400 were succinylated as identified in a single shot analysis. Within biological replicates dia-PASEF identified on average 8,200 peptide groups of which on average 4,600 were succinylated. Comparing control versus treated groups, more than 4200 succinylated peptides were shown to be statistically significant ($q<0.05$, $FC>0.58$) in either up or down regulation, where approximately 3100 were upregulated upon dietary treatment. In previous studies it has been shown that the proteome upon dietary supplement remains largely unchanged. These data suggest the complete succinylome re-wiring upon dietary feed and point toward Sirtuins as a novel drug target. Novel Aspect: High sensitivity profiling of succinyl PTMs using trapped ion mobility

Contact: Chris Adams, Christopher.Adams@bruker.com

P02.02 Utilizing an Iodine Unfolding Assay to Analyze the Effects of Aging on Protein Stability

Tyler Bateman (1), Chad Hyer (1), Lavender Lin (1), Michael Zackrison (1), John Price (1)
(1) Brigham Young University, Provo, Utah, United States

As humans age, diseases like Alzheimer's, cancer, and type II diabetes become more prevalent in the population. This is due in part to the loss of proteostasis, or the ability to regulate the concentration and quality of protein in the body. The proteostasis of an organism is dictated by the quality of its proteins. One metric of protein quality is the thermodynamic stability of the folded protein structure. Overly stable proteins can form aggregates or linger in a biological system for too long, while unstable proteins might not be retained long enough to perform their specialized tasks. Understanding protein stability can help researchers to combat age-related diseases by revealing novel mechanism of these ailments. We have developed a proteome-wide folding stability assay, wherein denaturation curves and an iodine reporter are used to quantify protein stability at multiple residues within the protein sequence. In this study, we use the unfolding assay to examine the

protein stability of mouse liver proteins during dietary interventions which modify their aging rate and lifespan. For example, caloric restriction of mice has been shown to extend their life span in previous studies. This extension is thought to result from improved proteostasis in mice with restricted diets. We can see how changes in proteostasis reflect on protein stability by comparing the stability of proteins from mice with different diets. This will provide unique insights into the effects of aging on an organism's proteome and overall human health.

Contact: Tyler Bateman, tboneb497@gmail.com

P02.03 Monitoring regionally-regulated control of protein turnover and concentration across the brain proteome using targeted Microsampling and AutoPOTS

Rebecca S. Burlett (1), Leena M. Patil (1), Holden Kelly (1), John C. Price (1)
(1) Brigham Young University, Provo, Utah, United States

Imbalances within the brain can lead to neurodegenerative diseases like Alzheimer's Disease (AD). The risk of developing AD has been linked to the isoforms of Apolipoprotein E (ApoE) that the body produces. However, it is unknown how ApoE (a lipid transport protein) affects brain proteostasis. Thus, it is important to understand the connection to AD. To understand this model mice were given deuterated water, over a timeline, prior to harvesting tissue. Utilizing a cryostat, we cut 20 micron slices of brain tissue that have been labeled metabolically with deuterium to allow analysis of the Hippocampus, Entorhinal, and Visual regions. Slices were then treated with a Nissl Stain and identified using Allen Brain Atlas. Next, Laser Capture Microdissection (LCM) was used to collect target specific 200x200x20 micron samples. AutoPOTS methods were used to prepare samples for mass spectrometry. Trypsinized samples were analyzed using LC-MS to measure protein abundance and turnover. Readings for each of the regions were compared against each other to identify patterns of regulation within protein expression. We have reproducibly identified ~2500 proteins with this targeted microsampling technique, almost as much as a normal LC-MS/MS run. Within the ~2500 proteins we have been able to identify, further analysis has provided us with turnover rates and concentrations that were put through analysis software, i.e. DeuteRater or Metaboanalyst. The outputs given by DeuteRater and Metaboanalyst were then run through python scripts that grouped the proteins into specific pathways provided by DAVID and original data. Those identified pathways have allowed us to pinpoint changes seen between the regions being compared. Partnering protein turnover rates with protein concentration we are able to investigate which pathways are being regulated by synthesis or by degradation. We are working towards a better understanding of how the brain's regions differ and are affected by AD overall.

Contact: Rebecca Burlett, burlett.rebecca@gmail.com

P02.04 The Role of Epigenetics-Mediated Neuroplasticity in Alzheimer's Disease as a Means of Understanding the Impacts of Apolipoprotein E Variants.

Noah Earls (1), Benjamin Jones (2), Katherine Brown (1), Kimberly Wagstaff (1), Ethan Smith (1), Noah Moran (1), Nathan Zuñiga (1), John C. Price (3)
(1) Brigham Young University, Provo, Utah, United States
(2) Brigham Young University, Utah, Utah, United States
(3) Brigham Young University, Provo, Utah

One of the primary modes of thought behind the causes of Alzheimer's Disease (AD), is the role of Apolipoprotein E (ApoE) variants. Several studies have highlighted the protective qualities of the ApoE 2 variant against several neurodegenerative disorders such as AD, Huntington's Disease, and Parkinson's Disease. Conversely, the ApoE 4 isoform contributes to the risk for these same neurological disorders, while ApoE 3 serves as a "neutral" predictor for AD. Additionally, AD has long been associated with the effects of neuroplastic changes found in patients with AD and how they relate to memory loss and dementia. Research has provided evidence that epigenetic factors (i.e., DNA Methylation, Histone Modifications, etc.) have a significant contribution to the neuroplasticity of AD. While ApoE has long been associated with these epigenetic variations, it remains to be seen what effect ApoE has on neuroplasticity-related proteins and how these changes demonstrate the predisposition for Alzheimer's Disease associated with ApoE 4. Our research aims to discover the connections between these ideas by analyzing the many factors that result in the neuroplastic changes common with AD. Using a transgenic mouse model on an LCMS system to understand the human proteome, we have analyzed both brain membrane and brain cytosolic proteins of these mice in search of significant changes between ApoE 2, ApoE 3, and ApoE 4. Our preliminary data shows several proteins strongly associated with neuroplasticity, such as Synaptopodin, that are significantly changed in concentration in ApoE 2 as compared to ApoE 3, and in ApoE4 as compared to ApoE 3. We expect this trend to also be expressed when we analyze the turnover rates. Here we show that we have been able to use LCMS data to detect significant changes in Neuroplasticity related proteins in both ApoE 2 and ApoE 4.

Contact: Noah Earls, noaheearls@gmail.com

P02.05 SIRT7 as a Novel Regulator of Cellular Senescence

Michael Gilbert (1), Lindsay Pino (1), Benjamin Garcia (2), Shelley Berger (1)
(1) University of Pennsylvania, Philadelphia, Pennsylvania, United States
(2) Washington University in St. Louis, St. Louis, Missouri, United States

Aging is a highly complex process with distinct cellular hallmarks that include epigenetic alterations, mitochondrial dysfunction, nutrient sensing, and cellular senescence. While these processes are distinct, they are also highly connected, as the perturbation of one of these hallmarks can influence another. As such, elucidating the areas of intersection between hallmarks could help define the mechanistic underpinnings of aging. Thus, the overarching goal of this proposal is to identify novel crosstalk mechanisms between metabolism, epigenetics, senescence, and other aging hallmarks to uncover novel targets that can promote healthy aging. With regards to cellular senescence, the process where a cell limits its ability to proliferate, it is a critical component of aging as senescent cells are accumulated in aged tissue and the selective removal of these cells can increase lifespan and ameliorate age-related diseases. Senescent cells undergo multiple cellular and molecular changes which include altered chromatin landscape, metabolic reprogramming, and the secretion of inflammatory cytokines and growth factors (SASP). We discovered that the least studied sirtuin, SIRT7, which has histone and non-histone substates, is targeted for degradation upon oncogene-induced senescence (OIS) and this precedes the upregulation of senescence associated proteins like p16 and Lamin B1. Rescue of SIRT7 loss during OIS by exogenous SIRT7 expression delayed senescence and importantly, diminished expression of SASP. These data suggest SIRT7 is a key regulator during senescence. Through immunoprecipitation-mass spectrometry (IP-MS), we determined E3 ubiquitin ligases that interact with SIRT7 upon senescence induction. These studies reveal a novel function of an epigenetic factor that is critical in aging and senescence and will allow deeper investigation into the crosstalk of epigenetics and metabolism.

Contact: Michael Gilbert, michael.gilbert@pennmedicine.upenn.edu

P02.06 Dynamics of Huntingtin Protein Interactions in the Striatum Identifies Candidate Modifiers of Huntington's Disease

Todd Greco (1), Christopher Secker (2), Eduardo Silva Ramos (2), Joel Federspiel (1), Jeh-Ping Liu (3), Alma Perez (4), Ismael Al-Ramahi (4), Jeffrey Cantle (5), Jeffrey Carroll (6), Juan Botas (7), Scott Zeitlin (3), Erich Wanker (8), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States
(2) Max Delbrück Centre for Molecular Medicine, Berlin, Germany
(3) University of Virginia School of Medicine, Charlottesville, Virginia, United States
(4) Baylor College of Medicine, Houston, Texas, United States
(5) Western Washington University, Bellingham, Washington
(6) Western Washington University, Bellingham, Washington, United States
(7) Baylor College of Medicine, Houston, Texas
(8) Max Delbrück Centre for Molecular Medicine, Berlin, United States

Introduction: Huntington's disease (HD) is an autosomal dominant fatal neurodegenerative disorder with one causative gene, huntingtin (HTT). Yet, the pathobiology is complex, suggesting that cellular factors influence disease progression. Here, we define HTT protein-protein interactions (PPIs) that are altered in abundance and stability due to the mutant protein with expanded polyglutamine in the mouse striatum, a brain region with selective HD vulnerability. By combining protein interactions investigations in mouse brain with validation in human cells and functional assays in Drosophila, we discover mutant HTT-dependent interactions whose genes modify disease pathogenesis. Method: The striatum of 2- and 10-month-old mice expressing 3XFLAG-Htt with a knock of human exon 1 containing either normal (20Q) or expanded (140Q) polyglutamine tracts were used for label-free and isotope-labeled affinity purification-mass spectrometry experiments to determine the specificity and relative stability of the interacting protein, except here adapted for tissues. Polyglutamine-regulated PPIs were validated in a human HD cell model using bioluminescence-based two-hybrid assays and Drosophila gene homologs were tested by a motor performance assay in an HD fly model expressing fragment ($HTT^{NT231[Q128]}$) or full length (HTT^{Q200}) mHTT. Results: Using metabolically labeled tissues and immunoaffinity purification-mass spectrometry, we found that polyglutamine-dependent modulation of HTT PPI abundances and relative stability starts at an early stage of pathogenesis in a 140Q HD mouse model. We identified direct and indirect PPIs that are also genetic disease modifiers using in-cell two-hybrid and behavioral assays in HD human cell and Drosophila models, respectively. Validated, disease-relevant mHTT-dependent interactions encompassed mediators of synaptic neurotransmission (SNAREs and glutamate receptors) and lysosomal acidification (V-ATPase). Conclusion: Using complementary AP-MS approaches, in-cell two hybrid assays, and fly genetic disease modifier assays, our results provide a resource for understanding mutant HTT-dependent dysfunction in cortico-striatal cellular networks, partly through impairment in synaptic communication and the endosomal-lysosomal system.

Contact: Todd Greco, tgreco@princeton.edu

P02.07 A proteomic network approach to define proteins associated with cognitive resilience in Alzheimer's disease

Cheyenne Hurst (1), Derian Pugh (2), Duc Duong (3), David Bennett (4), Jeremy Herskowitz (2), Nicholas Seyfried (1)
(1) Emory University, Atlanta, Georgia, United States
(2) University of Alabama Birmingham, Birmingham, Alabama, United States

(3) Emory Universoty, Atlanta, Georgia, United States
(4) Rush University Medical Center, Chicago, Illinois, United States

The core pathologies of Alzheimer's disease (AD) include extracellular plaques of aggregated amyloid-β (Aβ) and intracellular accumulation of tau neurofibrillary tangles. There is increasing recognition of a subpopulation of "cognitive-resilient" individuals that live to advanced age with intact cognitive function despite the presence of plaque and tangle pathology. Thus, identifying molecular mechanisms in brain associated with cognitive resilience may reveal novel therapeutic pathways and targets for AD. Here, we performed isobaric tandem mass tag (TMT) labeling coupled with off-line fractionation and liquid chromatography with tandem mass spectrometry (LC-MS/MS) on post-mortem brain tissue samples from the Religious Order Study and Memory and Aging Project longitudinal cohorts across two brain regions (Broadman area 6 and 37). This included pathology-free controls (n=24), cognitively normal "resilient" individuals with AD pathology (n=53) and symptomatic AD (n=34) cases. Following database search, ~8000 high abundance proteins were quantified. Consensus Weighted Gene Correlation Network Analysis (cWGCNA) of both brain regions generated 38 modules, groups of highly co-expressed proteins, which were correlated to clinical and pathological phenotypes and assessed for cell-type specificity and gene ontologies. Modules positively correlated to core AD pathology were significantly increased in both resilient and AD cases compared to controls. Conversely, modules enriched for synaptic and other vulnerable proteins, including VGF, NPTX2, and NRN1, were significantly decreased in AD, but not resilient cases. To further validate our systems-level data and explore potential neuroprotective mechanisms, we treated rat primary neuronal cultures with recombinant NRN1 protein with or without toxic Aβ 1-42 oligomer insult followed by dendritic spine morphometric analysis and in-depth proteomics. Morphometrics revealed NRN1 treatment in the presence of Aβ oligomers rescued overt dendritic spine loss observed with Aβ insult alone. Together, our findings support network nomination of resilience-associated proteins, including NRN1 which protects against amyloid-induced neurotoxicity that may in-turn promote cognitive resilience in AD.

Contact: Cheyenne Hurst, cheyenne.d.henckel@emory.edu

P02.08 Extracellular Matrix Proteins Differentially Expressed Between APOE Isoforms

Benjamin Jones (1), Noah Moran (1), Ethan Smith (2), Noah Earls (1), Kimberly Wagstaff (1), Katherine Brown (1), Anisha Daley (1), Nathan Zuniga (3), John Price (1)
(1) Brigham Young University, Provo, Utah, United States
(2) Brigham Young Uninversity, Provo, Utah, United States
(3) Brigham Young University, Provo, Utah

Apolipoprotein E (APOE) isoform expression is a significant risk factor for Alzheimer's Disease (AD). The three most common APOE alleles include: APOE3, the baseline variant, APOE 2, which bestows increased resistance against AD development, and APOE4, which increases the risk of the development of AD. The biochemical mechanism behind this disparity in AD risk remains unknown. To analyze the differences between polymorphisms, we homogenized brain samples from transgenic mice with each genotype; these homogenized samples were centrifuged to separate cytosolic proteins from membrane-bound proteins. Additional washes to the membrane pellet remove cytosolic contaminants that would otherwise limit our ability to see low abundance proteins on the membrane. Membrane samples were processed via state-of-the-art liquid chromatography and mass spectrometry (LC-MS), and this data was used to identify and quantify proteins in the membrane fractions.This data was further processed by comparing the concentrations of the proteins between polymorphisms to discover which proteins were upregulated or downregulated between polymorphisms. The regulation of

membrane-bound proteins in brain tissues may offer a deeper understanding of the signaling and structural differences between polymorphisms. Proteins in these tissues include receptors involved in cell signaling. Preliminary data shows a large number of proteins with significantly changed metabolism between APOE2 and APOE3. Collagen proteins and components of the extracellular matrix were some of the most significantly changed. Comparing both concentration and turnover rate changes gives us insight into the changes in protein metabolism. We will discuss our collagen and extracellular matrix protein observations in the context of the development and spread of amyloid beta aggregates, which are a leading biochemical marker for AD.

Contact: Benjamin Jones, benjaminjones1212@gmail.com

P02.10 Changes in the Proteome of Calorie Restricted Mice vs Ad Lipitum Mice

Noah Moran (1), Nathan Zuniga (1), John Price (1)
(1) Brigham Young University, Provo, Utah, United States

Human health is determined by the proper balance and function of proteins, which is called proteostasis. When proteostasis is interrupted, disease and death occur. Therefore, minimizing disease and maximizing life are top priorities in the scientific community. Some literature has shown that calorie restriction (CR) in mice prolongs lifespan relative to mice on an ad lipitum (AL) diet. This difference suggests that there are changes in the proteome. Those changes have not been fully revealed, so we did a follow up experiment. We put mice on two main diets (AL and CR) and five sub diets of protein to carbohydrate (PC) ratios (0.1, 0.2, 0.4, 0.8, 1.6). Liver tissue samples were homogenized and proteins were extracted, cleaned, and digested by trypsin. After analysis on LC/MS, data was put though PEAKS software. Our preliminary data suggest that mice on CR respond more quickly to dietary change than AL mice. This suggest changes in the proteome that lead to those fluctuations. We used LC/MS to quantify the differences in protein concentrations between AL and CR mice. This knowledge will identify changes in metabolic pathways which support longer lifespan due to nutrient type and total calories. We will present this data and compare against efforts to increase human lifespan.

Contact: Noah Moran, noah.g.moran@gmail.com

P02.11 Multi-omics analysis of the exosome senescence-associated secretory phenotype to study aging biology

Sandip Kumar Patel (1), Jacob Rose (1), Rebecca Beres (2), Joanna Bons (1), Nathan Basisty (3), Roland Bruderer (4), Lukas Reiter (5), Erin Baker (6), Judith Campisi (1), Birgit Schilling (1)
(1) Buck Institute for Research on Aging, Novato, California, United States
(2) North Carolina State University, Raleigh, North Carolina, United States
(3) National Institute on Aging, Baltimore, Maryland, United States
(4) Biognosys AG, Schlieren, California, United States
(5) Biognosys AG, Schlieren, Switzerland
(6) North Carolina State University, Raleigh, North Carolina

Cellular senescence induced by endogenous and exogenous stresses is a common aging hallmark. Senescent cells can drive aging phenotypes by paracrine signaling of the senescence-associated secretory phenotype (SASP), which can impair tissue regeneration and drive chronic age-associated diseases. Exosomes, a component of the SASP, are small, lipid-bilayer enclosed, cell-derived nanoparticles that play a vital role in cell-to-cell signaling by delivering bioactive

cargo. Senescence-associated exosomes, largely unexplored, could serve as aging biomarkers. We developed robust and efficient plasma exosome isolation and analytical workflows using genotoxic- and mitochondrial dysfunction-induced senescent IMR-90 human fibroblasts. For model development, 30 plasma samples from young and old humans and 20 conditioned media samples from senescent IMR-90 cells were analyzed using proteomics, lipidomics, and miRNA sequencing. We generated a deep plasma exosome-specific spectral library comprising ~2,300 exosome proteins by combining data-dependent acquisitions (DDA) and direct Data-Independent Acquisitions (directDIA) spectral libraries. Our data demonstrate that the directDIA spectral library enables efficient, time-saving proteomic workflows. Partial least squares-discriminant analysis (PLS-DA) of DIA-data could differentiate the plasma donors into young and old groups with no overlap. We identified 1,310 plasma exosome proteins with two unique peptides, of which 144 show differential abundance comparing old and young donors. In IMR-90 cells, 1852 exosome proteins with 2 unique peptides were identified, of which ~1000 changed with senescence. Interestingly, 42 significantly altered proteins are common in human plasma and the exosome SASP. 331 miRNAs were detected in human plasma exosomes, of which 88 and 17 miRNAs were unique to old and young individuals, respectively. Exosome lipidomics on a timsTOF ion mobility MS platform identified >300 lipid species and 38 were differentially present in young and plasma. We will validate our model with larger cohorts and integrate proteomics lipidomics, and miRNA data to discover molecular aging markers.

Contact: Sandip Kumar Patel, SPatel@buckinstitute.org

P02.12 Deep and Unbiased Plasma Protein Profiling of Alzheimer's and Mild Cognitive Impairment Subjects with a Novel Multi-nanoparticle Approach

A Siddiqui (1), RB Benz (1), M Goldberg (1), JC Cuevas (1), W Manning (1), X Zhao (1), TL Platt (1), M Ko (2), EM Elgierari (1), M Figa (2), H Xia (1), H Guturu (1), M Zamanighomi (1), J Zhang (1), S Batzoglou (1), OC Farokhzad (1)
(1) Seer, Redwood City, California, United States
(2) Seer, Redwood City, California

Blood plasma is a rich source of protein biomarkers for early detection of diseases, but its large dynamic range of protein concentrations necessitates complex workflow and trade-offs between throughput, scalability, coverage, and precision. Here we use a deep and quantitative proteome profiling platform, Proteograph™ Product Suite, which leverages multiple nanoparticles, engineered with distinct physicochemical properties to provide broad coverage of the plasma proteome at scale. In this study, we aim to identify protein biomarkers for Alzheimer's disease (AD) from blood with this untargeted plasma protein profiling approach, for AD and Mild Cognitive Impairment (MCI) condition. Previously, plasma samples from 200 subjects comprising 50 AD, 50 MCI, and 100 controls were profiled using Proteograpgh plasma protein profiling platform[1]. Using 5 injections per sample, proteins were quantified by data-independent acquisition (DIA) liquid-chromatography mass-spectrometry (LC-MS). Normalized peptide intensities were used to develop models for class discrimination. Across the samples, 2,391 plasma proteins were detected, with 2,085 in at least 25%. 25,593 peptides were detected, with 15,661 in at least 25% of the samples. Univariate analysis identified 441 and 526 proteins that were significantly different in AD or MCI versus control, respectively. Random-forest classification with ten rounds of 10-fold cross-validation yielded ROC AUCs that were at least 0.90 for AD and MCI. These analyses identified a combination of known and potential new candidate plasma protein markers, confirming the Proteograph platform's ability to generate profiling data in a deep, broad, and rapid fashion, enabling large-scale studies to detect novel insights with clinically relevant potential. We are in process of conducting a second study of 200 subjects comprising 100 AD and 100 controls and will present the results from this second cohort at the meeting. References. 1. Blume JE et al. Nature Communications. 2020;11(1):3662–1

Contact: Asim Siddiqui, asiddiqui@seer.bio

P02.13 The Liver's Role in ApoE-Mediated Alzheimer's Disease Risk

Ethan G Smith (1), Nathan Zuniga (1), Kimberly Wagstaff (1), Katie Brown (1), Noah Moran (1), John C Price (1)
(1) BYU, Provo, Utah, United States

Alzheimer's Disease (AD) carries finality to the declining quality of life in many aging members of the population. While several genetic risk factors have been characterized, the mechanisms by which they are associated with AD are not well understood. Pathological patterns have been observed in AD progression, but we do not understand them well enough to intervene medically. Using proteomics methods to better characterize the biochemistry of AD risk across different tissues and organs is a very promising way to identify the mechanisms of risk and potential interventions. The genetic risk factor that we study involves three variants of the lipid transport protein ApoE. This is among the strongest genetic correlates of AD risk, and it does much of its work between the liver and the brain. Both organs produce ApoE, and the liver produces many lipids for energy and structure in the brain. As AD risk increases, we and others find this relationship dysregulated, with changes in both organs. The clinical manifestations are far more obvious in the brain, but MS proteomic analysis also gives insight into the mechanistic underpinnings that are less obvious in the liver. Here I will present our results investigating proteomic perturbations in mice with transgenic insertions of ApoE variant 2, 3, or 4 as our AD risk models. We analyzed both membrane and cytosolic protein populations from homogenized liver and brain tissue, and we compared differential impacts of the ApoE variants on the two organs. As we compare these organs, we can better understand the full-body contributions to the risk for AD pathology.

Contact: Ethan Smith, tkdethan@gmail.com

P02.14 An Untargeted Metabolomic Investigation of Amyotrophic Lateral Sclerosis Decedent Brain Tissue by IR-MALDESI Mass Spectrometry Imaging

Alexandria L. Sohn (1), Lingyan Ping (2), Jonathan D. Glass (2), Nicholas T. Seyfried (2), David C. Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States
(2) Emory University School of Medicine, Atlanta, Georgia, United States

Amyotrophic lateral sclerosis (ALS) is a fatal, idiopathic neurodegenerative disease characterized by progressive motor neuron degradation with limited treatment options. This disease state is classified based on presence or absence of family history, respectively differentiating familial ALS (fALS) and sporadic ALS (sALS), and further specified with regard to genetic mutations if applicable (e.g. C9orf72). Because metabolites provide insight into dysregulation of upstream biological pathways, an untargeted metabolomic study of brain tissue could identify biomarkers and exogenous compounds related to ALS in hopes to optimize diagnosis protocol and extend the life expectancy of ALS patients. Infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) mass spectrometry imaging (MSI) is a powerful approach leveraged in this study to analyze the spatial distribution of analytes across a tissue with respect to their abundance without the necessity of an organic matrix. Post-mortem human brain tissue was imaged at 10-micron thickness from three subgroup classifications: non-neurological controls, sALS decedents, and sALS decedents positive for the C9orf72 repeat expansion mutation. Selection of sALS and C9orf72 positive cases for analysis respectively was due to the significantly higher rate of incidence of sALS (~90%) relative to fALS, and the C9orf72 mutation relative to other genetic mutations associated with ALS (e.g. SOD1, TARDBP, FUS). Data collected was annotated by METASPACE and further analyzed utilizing MSiReader and

Excel. Additionally, a normalization approach was developed and employed for the comparison of tissues within and across diagnosis subgroups to reduce analytical variability between biological samples for a more accurate representation of biological variability in a MSI workflow. Ultimately, over 100 annotated metabolites were common between all samples and their respective differences in abundance between groups are under further investigation for biological differences between disease classes. Furthermore, MS/MS follow-on studies will be pursued for metabolite structure validation to enhance our understanding of ALS-related metabolites.

Contact: Alexandria Sohn, asohn2@ncsu.edu

P02.15 Investigating Racial Disparities in Alzheimer's Disease using Quantitative Brain Proteomics

Jasmin B. Tindal (1), Kaitlyn E. Stepler (2), Shania Hansen (3), Timothy J. Hohman (3), Lisa L. Barnes (4), David A. Bennett (4), Renã A.S. Robinson (1)
(1) Vanderbilt University, Nashville, Tennessee, United States
(2) Vanderbilt Univerisity, Nashville, Tennessee, United States
(3) Vanderbilt Memory and Alzheimer's Center, Nashville, Tennessee, United States
(4) Rush Alzheimer's Disease Center, Chicago, Illinois, United States

Compared to the non-Hispanic White population, the African American/Black population is disproportionately affected by Alzheimer's disease (AD) with a 2-3 times higher incidence rate. Understanding the molecular basis of this disparity has been challenged by a lack of samples from African American/Black participants in basic science studies. Previously, our laboratory determined many proteomics changes in AD that are universal across racial/ethnic backgrounds. However, we also observed proteins that had significant differences between cognitively unimpaired and AD participants that were associated with racial and ethnic background of the participants. Here we applied a proteomic workflow to postmortem tissue from the inferior parietal lobule (IPL) and prefrontal cortex (PFC) brain regions across cognitively unimpaired and AD adults from the Religious Orders Study and Rush Memory and Aging Project. The individuals in the cohort self-identified as African American/Black or non-Hispanic White. We employed multiplexed proteomics with Tandem Mass Tags (TMT), reversed-phase fractionation, and nanoflow LC-MS/MS analysis using Q-Exactive HF and Orbitrap Lumos mass spectrometers. Overall, >9000 proteins were identified across the two regions and were subjected to regression analysis. Preliminary data of the IPL suggest >75 proteins had changes in AD that were significantly impacted by race. This presentation will highlight differentially expressed brain proteins in AD, including those associated with racial backgrounds, and implications for disease understanding.

Contact: Jasmin Tindal, jasmin.b.tindal@vanderbilt.edu

P02.16 Mapping the 3 D Proteome using Surface Accessibility in Animal Models of Disease

John Yates (1), Ahrum Son (1), Casimir Bamberger (1), Sandra Pankow (1)
(1) The Scripps Research Institute, Lajolla, California, United States

Methods to determine the surface accessibility of amino acids on proteins can provide information about conformational changes to proteins because of mutation, drug treatment, or other perturbations. These methods such as FPOP, SPROX or LiP are collectively described as protein footprinting. We describe an approach using formaldehyde to label accessible lysine residues in animals using perfusion of formaldehyde into the animal which is then converted to dimethyl tags using

cyanoborohydride. Proteins are labeled both in situ and in vivo. As has been used in quantitative proteomics, different stable isotope elements can be used in formaldehyde and cyanoborohydride to create different weight tags and thus a secondary labeling after proteolytic digestion of proteins using different weight reagents labels lysine residues that were not. After removal of tissue it is homogenized, solubilized and digested and then labeled with a different weight combination of formaldehyde and cyanoborohydride. Quantitative differences in surface accessibility can be observed to identify changes from a control or healthy state. This method is being used to look at protein misfolding in connection with Alzheimer's disease in a mouse model.

Contact: John Yates, jyates@scripps.edu

P02.17 Increased Sequence Coverage of Low-abundance Biomarkers in Human Serum to Quantify Protein Folding Stability

Michael Zackrison (1), Lavender Lin (1), JC Price (1), Tyler Bateman (1), Chad Hyer (1)
(1) Brigham Young University, Provo, Utah, United States

The structural integrity of proteins is foundational for their ability to carry out their biological function. The protein folding stability (PFS) can be quantified by the amount of energy needed to unfold the structure. The alteration of PFS can result in aggregation and a decrease in proteostasis. With the novel iodinated protein folding stability assay developed by our lab, we can take a snapshot of localized PFS of the proteins across the proteome and compare the stability between various experimental conditions. Cardiac amyloidosis is a disease caused by amyloid aggregation of misfolded serum protein transthyretin (TTR) in the heart. Unfortunately, it usually can't be diagnosed until it is too late. TTR functions as a tetramer. When the tetramer can't be stabilized, it dissociates into beta sheet-stabilized monomers that then aggregate as amyloids. Our goal is to measure the PFS of TTR under different states of disease and intervention and develop a metric to aid early diagnosis. However, TTR is a low-abundance protein in human serum, and most of its signals are covered by the most abundant serum protein: serum albumin. Consequently, it is challenging for us to repeatedly measure the PFS of the same site on a TTR protein across different treatment conditions, affecting our ability to have a reliable measurement. Here we present our efforts on increasing the sequence/PFS site coverage of TTR while also retaining the effect from its interactions using offline HPLC fractionation and serum albumin depletion after iodinating our samples. The methods used here not only promote the PFS measurement of TTR, but also can be applied to other low-abundance biomarkers for further understanding the mechanism of disease.

Contact: Michael Zackrison, zackpac08@gmail.com

P03: Biomarkers and Targeted MS Assays

P03.01 Urine Proteomics for Noninvasive Monitoring Health and Disease in Preterm Neonates: Urinary Biomarkers for Bronchopulmonary Dysplasia

Saima Ahmed (1), Oludare A. Odumade (1), Patrick van Zalm (1), Kinga K. Smolen (1), Kimino Fujimura (1), Jan Muntel (1), Melissa S. Rotunno (1), Abigail B. Winston (2), Judith Steen (1), Richard B. Parad (2), Linda J. Van Marter (3), Stella Kourembanas (1), Hanno Steen (1)
(1) Boston Children's Hospital and Harvard Medical School, Boston, Massachusetts, United States
(2) Brigham & Women's Hospital and Harvard Medical School, Boston, Massachusetts
(3) Brigham & Women's Hospital and Harvard Medical School, Boston, Massachusetts, United States

Introduction: Current techniques to diagnose and/or monitor critically ill neonates with bronchopulmonary dysplasia (BPD) require repeated invasive sampling of body fluids, which by itself interferes with the wellbeing of these extremely frail patients. We tested our hypothesis that it is feasible to use non-invasively collected urine samples for a urine proteomics study with extremely low gestational age newborns (ELGANS) at risk for bronchopulmonary dysplasia to identify proteins, which have previously been associated with BPD albeit never in urine. Methods: Our urine proteomics pipeline requires as little as 50 microliters of urine and is thus fully compatible with sample amounts collected from ELGANs in Neonatal Intensive Care Units (NICUs). We performed a proof-of-concept urine proteomics study validating proteins, which had previously been identified in invasively collected samples such as blood, broncho-alveolar lavage and/or tracheal aspirates. We used urine collected within 72 hours of birth from ELGANS (gestational age (26+1.2) weeks) who had been admitted to a NICU; half of whom eventually developed BPD (n=21), while the other half remained BPD free, i.e., serving as controls (n=21). Results: Our sample-sparing urine proteomics pipeline clearly identified several BPD-associated changes in the urine proteome recapitulating well-described changes in the blood, broncho-alveolar lavage and/or tracheal aspirate proteomes. In addition, several novel urinary biomarker candidates were discovered that allowed to identify the ELGANs at risk for BPD. Interestingly, sixteen of the identified urinary proteins are known targets of FDA-approved drugs pointing towards potential novel therapeutic strategies. Conclusion: In addition to validating numerous proteins, previously found in invasively collected blood, tracheal aspirate, and broncho-alveolar lavage, that have been implicated in BPD pathophysiology, urine proteomics also suggested novel potential therapeutic targets. Ease of access to urine could allow for repeated non-invasive monitoring of disease progression and impact of therapeutic intervention in future studies.

Contact: Saima Ahmed, saima.ahmed@childrens.harvard.edu

P03.02 Reducing Sample Input Requirements for Parallel Reaction Monitoring of Affinity-Enriched Trace Biomarkers

Nathaniel Axtell (1), Jacob J. Kennedy (2), Rachel Lundeen (2), Andikan Nwosu (1), Thy Truong (1), Jeffrey R. Whiteaker (2), Xiaofeng Xie (1), Lei Zhao (3), Amanda G. Paulovich (2), Ryan T. Kelly (4)
(1) BYU Department of Chemistry and Biochemistry, Provo, Utah, United States
(2) Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States
(3) Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington
(4) BYU Department of Chemistry and Biochemistry, Provo, Utah

Studies of protein pathways often reveal key proteins whose expression levels can be excellent indicators of normal, pathogenic, and response states of biological systems. Such proteins make ideal biomarkers for medical diagnostics, as well as targets for therapeutics. Potential protein biomarkers are routinely quantified in laboratory settings using mass spectrometry (MS) methods such as multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM). A significant challenge in translating these measurements to clinical application is overcoming sensitivity hurdles in analyzing limited clinical sample material. Immunoaffinity enrichment paired with MRM/PRM analysis has been demonstrated to increase measurement sensitivity for some biomarker candidates; however, this can still require hundreds of micrograms of protein input to yield quantitative measurements. Here we report on improvements to the standard immuno-PRM workflow that reduce required sample input requirements by more than an order of magnitude. By employing a MS-compatible surfactant during sample preparation, we simplified clean-up steps, preventing protein losses due to nonspecific surface adsorption, and maximizing target proteotypic peptide recovery. Also, our ultra-low flow nanoLC setup increased ionization efficiency significantly, further reducing sample loss. These changes resulted in a cumulative increase in immuno-PRM assay sensitivity over current methods. We demonstrated the applicability of our immuno-PRM workflow through detection and quantitation of DNA damage response proteins FANCD2 and RAD50 expressed in lymphoblast cells. We validated our results through comparison with a previous study utilizing a conventional immunoaffinity-enriched MRM method.

Contact: Nathaniel Axtell, axtelln@byu.edu

P03.03 Zeno MS/MS powers quantitative and qualitative workflows for proteomics using the ZenoTOF 7600 system

Alexandra Antonoplis (1)
(1) SCIEX, Redwood City, California, United States

Proteomic workflows cover a wide range depending on project goals, from fully untargeted data dependent acquisition (DDA) approaches for protein identification, comprehensive data independent acquisition (DIA) strategies for large scale quantification, and fully targeted quantitative assays for the highest specificity/sensitivity (MRM). Novel Zeno trap functionality can greatly improve duty cycle in orthogonal pulsing region of a QTOF system, providing large gains in MS/MS sensitivity. This work investigates the impact of increased MS/MS sensitivity on protein/peptide identification and targeted peptide quantitation. Microflow chromatography was performed on a Phenomenex Kinetex column at 5 µL/min using a range of gradients (5, 10, 20, and 45 min). Digested K562 cell lysate was used for DDA testing and the PQ500 kit (Biognosys, 804 heavy peptides) was used for peptide quantification. Using the ZenoTOF 7600 system, all experiments were performed with and without the Zeno trap activated. Data was processed using OneOmics suite, SCIEX OS software and Skyline. To test the impact on peptide MS/MS sensitivity, a targeted peptide quant assay was built using Skyline for the PQ500 kit (804 peptides). Using the final MRM^{HR} assay with Zeno trap off vs. Zeno trap on, the average gain in MS/MS fragment peak area was ~5.6 fold. Median peak area CV was 6.1% for 10 replicate injections and median LLOQ was found to be 193 amol on column. Next, Zeno IDA acquisition parameters were optimized for 4 different microflow gradient lengths using a design of experiments (DOE) approach. Using optimized settings, comparison data for Zeno trap on and off was generated for all gradients. Specific gains with Zeno on increase with gradient length, with gains in protein identifications of more than 35% for longer gradients. Sample loading curves were also generated at all gradient lengths to fully explore acquisition space, and results will be discussed.

Contact: christie hunter, christie.hunter@sciex.com

P03.04 Developing and Validating Targeted Mass Spectrometry Assays for Pan-Herpesvirus Viral Protein Detection and Monitoring of Infection Progression

Michelle Kennedy (1), Joel Federspiel (1), Cora Betsinger (1), Matthew Tyl (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

INTRODUCTION: The presence and abundance of viral proteins within host cells signify the cellular stages of viral infections. Viral proteins are either brought into cells by infectious particles or expressed at specific steps of the replication cycle. However, methods that can comprehensively detect and quantify these proteins are limited, particularly for viruses like herpesviruses that boast large protein coding capacities. Importantly, the development of such methods would provide a comprehensive portrait of viral replication and allow for the screening of small molecules and other cellular perturbations with potential therapeutic and clinical applications. METHODS: By integrating mass spectrometry and molecular virology, we have designed and experimentally validated a set of targeted proteomics assays for monitoring human viruses representing the three Herpesviridae subfamilies—herpes simplex virus type 1, human cytomegalovirus, and Kaposi's sarcoma-associated herpesvirus. Our assays target 200+ unique peptides covering 50-80% of the predicted proteomes for each of these viruses and span proteins that are representative of different temporal protein expression classes as well as virion components. RESULTS: During wild-type virus infections, we first demonstrated that the breadth of proteins monitored by our methods captures the temporal cascades of the replication cycles of these viruses. Additionally, we illustrated that our assays can detect viral proteins at clinically relevant levels of infection. We next showed that these assays can be used to quantify the effects of long-established and recently-discovered antiviral agents, and further captures their precise temporal regulation of specific viral proteins. Finally, we demonstrated their broad utility for monitoring different viral strains, including laboratory and clinical isolates. CONCLUSIONS: Altogether, our assays provide a reproducible framework for monitoring the progression of herpesvirus infections and are broadly applicable across a variety of model systems and contexts, including drug screening, detecting infections in clinical settings, and genetic manipulations of virus or host factors.

Contact: Michelle Kennedy, mak4@princeton.edu

P03.05 Beyond Concentration: Quantifying changes in protein fold stability of human serum proteins due to drug binding using a novel iodination assay

Hsien-Jung L. Lin (1), Isabella James (1), Chad Hyer (1), Connor Haderlie (1), Michael Zackrison (1), Tyler Bateman (1), Anisha Daley (1), Monica Berg (1), Stella Park (1), Nate Zuniga (1), John Price (1)
(1) Brigham Young University, Provo, Utah, United States

Currently, most biomarker discoveries focus on concentration differences. However, concentration is not sufficient to assess the impact of the observed changes. There is a need to quantify the quality of the proteins in the proteome to assist diseases diagnosis. Our lab has developed a novel assay that uses protein folding stability (PFS) as a metric to describe protein quality in the proteome. Our assay measures the classical idea of denature midpoint ($C_{1/2}$). A tissue homogenate is aliquoted into a GdmCl gradient to denature each protein of the proteome in a controlled manner. Iodine is then added to covalently label surface-exposed amino acids (Y, H, C, M, W). The modified peptides served as reporters of solvent accessibility. The different amino acids we targeted compared to other PFS methods allow us to track the stability of a protein across its sequence, which provides another perspective to understand protein structure and protein-protein interaction. Here we present the application of the assay on human serum albumin (HSA) from undepleted human serum. HSA is the most abundant protein in serum. It accounts for 60% of the serum proteome, which often presents a challenge

for serum proteomics. In most of the applications, HSA, and other abundant proteins, are removed for meaningful biomarkers analysis. However, studies have shown that the stability of HSA is altered in disease states. The results show how the PFS at different regions of HSA, and the PFS of other serum proteins, respond to drug interventions. It also shines a light on the potential of using HSA as a diagnosis tool. Further, our folding stability assay in combination with other proteome scale metrics like concentration, turnover, or PTM can promote in vivo assessments of proteome regulation in humans, and aid in biomarker discoveries.

Contact: Hsien-Jung Lin, llhj13@chem.byu.edu

P03.06 Patient-specific autoantigen sample preparation and analysis using ProMTag

Jonathan Minden (1), Dana Ascherman (2), Amber Lucas (3)
(1) Impact Proteomics, Pittsburgh, United States
(2) University of Pittsburgh, Pittsburgh, Pennsylvania, United States
(3) Impact Proteomics, Pittsburgh, Pennsylvania, United States

Autoimmune diseases affect >20 million people in the US today. Currently, disease-specific autoantibodies are thought to be the best biomarkers for diagnosis. Conventional immunoprecipitation methods have been used to identify autoantigens from the most common autoimmune diseases. However, these diseases account for only 6.5 million of the 20 million patients suffering from autoimmune diseases, leaving many without diagnoses until irreversible damage occurs. The remaining 13.5 million patients have >70 autoimmune disorders without well characterized autoantibodies. The state-of-the-art diagnostic test of these remaining diseases relies on gel electrophoresis of immunoprecipitated radiolabeled proteins, which cannot be identified by MS due to safety issues and the overwhelming presence of immunoglobulins. We have created an immunoprecipitation method that uses serum from patients with any autoimmune disorder to identify patient-specific autoantigen proteins. This method uses a reversible click chemistry tag, called ProMTag. One end of the ProMTag forms a reversible, covalent bond with protein by coupling to lysines and amino termini. The other end of the ProMTag can form an irreversible, covalent bond with a solid bead support via a click chemistry, methyltetrazine-TCO, pairing. In this study, the proteins of cell lysates that contain potential autoantigens were labeled with ProMTag. The ProMTagged-proteins were exposed to patient antibodies bound to Protein A beads, thus capturing the ProMTagged autoantigens. All proteins were released from the Protein A beads, including ProMTagged-autoantigens and untagged-antibodies. The ProMTagged-autoantigens were subsequently coupled to TCO beads, and the untagged-antibodies were washed away. The linkage between the ProMTag and autoantigens was then reversed, yielding autoantigen proteins with greatly reduced antibody contamination ready for MS analysis. MS analysis successfully identified autoantigens from serums of patients with rheumatoid arthritis. This autoimmune biomarker discovery method can accelerate sample testing for known autoantigens and facilitating rapid discovery of novel autoantigens for both diagnostic and predictive biomarkers.

Contact: Jonathan Minden, minden@cmu.edu

P03.07 Quantitative Proteomic Analysis of the Senescence-Associated Secretory Phenotype in Primary Human Endothelial Cells to Identify Senescence Biomarkers

Francesco Neri (1), Pei-Hsun Wu (2), Akos Gerencser (1), Sandip Patel (3), Jacob Rose (1), Shuyuan Zheng (2), Jude Phillip (2), Pierre-Yves Desprez (4), Denis Wirtz (2), Judith Campisi (5), Birgit Schilling (1)
(1) Buck Institute for Research on Aging, Novato, California, United States

(2) Department of Chemical and Biomolecular Engineering, Johns Hopkins Physical Sciences Oncology Center, Johns Hopkins Institute for Nanobiotechnology (INBT), Johns Hopkins University, Baltimore, Maryland, United States
(3) Buck Institute for Research on Aging, Novato, Novato, California, United States
(4) Buck Institute for Research on Aging, Novato, CA, USA, Novato, Maryland
(5) Buck Institute for Research on Aging, Novato, California

Age is the primary risk factor for major human pathologies, such as cardiovascular disorders, cancer, and neurodegenerative diseases. Senescent cells and their senescence-associated secretory phenotype (SASP) drive aging and promote such age-related pathologies. Identifying robust and easily-accessible biomarkers for the quantification of senescence burden in humans is essential for translating therapeutic interventions that target senescent cells. Endothelial cells line blood vessels and can secrete biomolecules directly into circulation. Thus, we hypothesized that studying the SASP of senescent endothelial cells holds great promise for the identification of plasma biomarkers of senescence burden. To do this, we optimized a cell culture model of primary human senescent endothelial cells by treating lung microvascular endothelial cells – a cell type that comprises 50% of the capillary surface of our entire body – with the genotoxic chemotherapeutic drug doxorubicin. Our protocol induces senescence robustly, as shown by high senescence-associated β-galactosidase staining, halted proliferation, and a senescent gene signature. We are now optimizing a workflow to study the SASP of senescent endothelial cells by data-independent acquisition (DIA) mass spectrometry. This approach will enable us to study the senescent secretome in a comprehensive and unbiased fashion and identify potential plasma biomarker candidates for human senescence burden. We are also investigating the senescent cell morphology in order to eventually identify senescent cells in tissues based on their morphological traits. Thus, we started quantifying morphological features of senescent endothelial cells by combining immunofluorescence staining of known senescence markers and cytoskeleton dyes using an artificial intelligence-based algorithm that enables high content image analysis. Our preliminary data shows that senescent cell populations show a strikingly altered morphology and are highly heterogeneous compared to control populations. Thus, our translational approach utilizes multi-omics technologies for potential senescence-derived biomarker identification.

Contact: Francesco Neri, fneri@buckinstitute.org

P03.08 SureQuant targeted mass spectrometry standards and assay panel for quantitative analysis of phosphorylated proteins from multiple signaling pathways

Bhavin Patel (1), Penny Jensen (2), Amirmansoor Hakimi (3), Sebastien Gallien (4), Aaron Gajadhar (3), Ana Martinez Del Val (5), Jesper Olsen (5), Andreas Huhmer (6), Daniel Lopez-Ferrer (3), Ryan Bomgarden (7), Kay Opperman (1), John Rogers (1)
(1) Thermo Fisher Scientific, Rockford, Illinois, United States
(2) Thermo Fisher Scientific, Rockford, United States
(3) Thermo Fisher Scientific, San Jose, California, United States
(4) Thermo Fisher Scientific, Paris, France
(5) University of Copenhagen, Copenhagen, Denmark
(6) Thermo Fisher Scientific, San Jose, California
(7) Thermo Fisher Scientific, Rockford, Illinois

Introduction There is broad interest in quantifying dynamic protein phosphorylation states in cellular signaling pathways under different conditions. We have combined EasyPep™ technology, phosphopeptide enrichment, validated multipathway

AQUA™ heavy-labeled phosphopeptide standards, and SureQuant™ targeted mass spectrometry to quantitate changes in phosphorylated protein abundance across multiple stimulated cell lines. Specific phosphopeptide standards were chosen representing phosphosites from several different pathways. This novel workflow enables targeted quantitation of biologically relevant phosphorylation sites with high accuracy, precision, and specificity. Methods Multiple cell lines were grown with different stimulation conditions before in-solution digestion using EasyPep Maxi MS sample prep kit. One milligram of each digest spiked with phosphopeptides standard was subjected to analysis using the Thermo Scientific™ Pierce™ Fe-NTA phosphopeptide enrichment kit. Discovery and targeted LC-MS/MS analysis were performed using Thermo Scientific Dionex nanoLC™ system or Thermo Scientific Vanquish Neo UHPLC system coupled to Thermo Scientific™ Q Exactive™ HF Hybrid Quadrupole-Orbitrap or Thermo Scientific™ Orbitrap Exploris™ 480 or Orbitrap Eclipse™ Tribrid™ Mass Spectrometers. To ensure optimal measurement for targets, a novel SureQuant targeted acquisition method was performed where real-time heavy peptides detection triggered high-sensitivity measurement of endogenous targets. Data analysis was performed with Proteome Discoverer and Skyline software. Results Our optimized workflow includes EasyPep MS sample prep kits and Fe-NTA enrichment with 131 AQUA heavy-isotope phosphopeptide standards to monitor multi-pathway signaling. Two targeted MS methods (PRM and SureQuant) were compared to assess the relative performance for quantitation of the desired endogenous peptides. More than 100 endogenous phosphopeptides from multiple stimulated cancer cell lines and all 131 heavy phosphopeptides were quantitated with high sensitivity and reproducibility. SureQuant method allowed quantitation of endogenous phosphopeptides at 10X lower levels than PRM. Conclusion SureQuant multipathway phosphopeptide standard with novel SureQuant MS analysis allows reproducible, routine, and simultaneous quantitation of functionally relevant phosphorylation sites.

Contact: Bhavin Patel, bhavinkumar.patel@thermofisher.com

P03.09 Compared dia-PASEF and prm-PASEF approaches for the absolute quantification of 500 human plasma proteins in colon cancer plasma samples.

Gunnar Dittmar (1), Antoine Lesur (1), Francois Bernardin (1), Elisabeth Letellier (2), Gary Kruppa (3), Pierre-Olivier Schmit (4)
(1) LIH-Luxembourg Institute of Health, Strassen, Luxembourg
(2) UNiversity of Luxembourg, Luxembourg, Luxembourg
(3) Bruker Daltonik, Bremen, Germany
(4) Bruker France S.A., Wissembourg, France

Introduction Targeted proteomics methods are a traditional choice for absolute protein quantitation of proteins in biofluids. We recently evaluated the prm-PASEF acquisition strategy which, compared to traditional prm approaches, allows to further increase the number of addressable targets and the method's selectivity without compromising the sensitivity. In parallel, the even higher multiplexing potential of dia-PASEF approaches (now regularly used in discovery studies) also triggered our interest. We are now applying both approaches to the absolute quantitation of 500 blood proteins in colon cancer plasma samples. Methods The plasma sample cohort consisted in 10 patients affected by a colon cancer (adeno carcinoma) and 10 controls. Plasma samples were depleted with a Mars 14 depletion column (Agilent), digested with a trypsin protease and spiked the PQ500 (Biognosys) synthetic peptides mixture. Samples and controls were separated by nano-HPLC (nanoElute, Bruker Daltonics) on a 25cm pulled emitter column (IonOpticks) using a 100 min gradient. Peptides were analyzed on a timsTOF Pro2 instrument (Bruker Daltonics) operated in prm-PASEF and dia-PASEF modes. Data processing has been done with Spectronaut (Biognosys), MaxQuant and Skyline-daily. Results We evaluated the quantification performance of the prm-PASEF in depleted plasma samples by monitoring 1564 precursor-ions

corresponding to 782 peptides from 565 proteins while using a 2 min retention time window. Some peptides could be quantified down to 7 amol with prm-PASEF and 20 amol with dia-PASEF . The median relative standard deviation of the signal of the peptides was of 3%. 98% of the 574 quantified peptide pairs could be quantified from the prm-PASEF experiment, while 96% could be quantified using dia-PASEF. The results obtained with both approaches were highly correlated. Conclusions Both approaches have been successfully applied to the analysis of colon cancer plasma samples and allowed to spot regulated proteins that had formerly been spotted from cancer tissue analysis.

Contact: Pierre-Olivier SCHMIT, pierre-olivier.schmit@bruker.com

---

P03.10 Discovery of Senescence Biomarker Candidates from Monocyte Proteomes and Secretomes

DIMITRIOS TSITSIPATIS (1), ANJANA RAM (1), MARTINA ROSSI (1), MYRIAM GOROSPE (1), NATHAN BASISTY (1)
(1) National Institute on Aging, Baltimore, Maryland, United States

Introduction: Cellular senescence is a complex stress response characterized by permanent cell-cycle arrest and the activation of a pro-inflammatory senescence-associated secretory phenotype (SASP). Accumulation of senescent cells and SASP are drivers of aging and many pathological processes including chronic inflammation and cancer. Biomarkers of senescent cell burden may be clinically useful for treatment of age-related pathologies in humans. Here we perform comprehensive and quantitative proteomic profiling of intracellular and secreted proteomes of senescent monocytes to identify biomarker candidates of monocyte senescence burden in human peripheral blood. Methods: To model the full range of physiological oxygen exposure experienced by monocytes in circulation, we induced senescence using ionizing radiation in two monocyte cell lines, U937 and THP-1, under hypoxic or hyperoxic oxygen conditions. After 6 days, proliferating and senescent cells were placed in serum- and phenol-red-free RPMI media for a 24-hour period. Secreted media and cell pellets were collected for LC-MS/MS analysis and assessment of a panel of canonical senescence and viability markers. Analysis of all samples was performed on the Q-Exactive HF Orbitrap mass spectrometer using Data-Independent Analysis and analyzed in Spectronaut. Results: We established a model of senescence in two independent monocyte cell lines that is characterized by senescence markers of reduced proliferation (assessed by EdU incorporation), increased expression of p21, p16, DPP4, and IL6, and increased senescence-associated β-galactosidase activity. We report the first comprehensive and unbiased assessment of the senescent monocyte intracellular proteome and secretomes, with many unique features no shared with other senescent cell types. We compared the proteomes of senescent monocytes with existing biomarkers of aging in humans to identify promising peripheral biomarkers of senescence in humans. Conclusion: These results may provide clinically useful information for establishing therapeutic targets or biomarkers to aid in the translation of senescence-targeted therapies to treat chronic inflammation and age-related decline.

Contact: DIMITRIOS TSITSIPATIS, dimitrios.tsitispatis@nih.gov

---

P03.11 Monitoring sirtuin dynamics during viral infection using a targeted mass spectrometry assay

Matthew D. Tyl (1), Yana D. Miteva (1), Ileana M. Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Sirtuins are NAD^+-dependent enzymes that function as “sensors” of the cellular environment by transmitting information through the addition or removal of posttranslational modifications (PTMs). Although initially designated as Class III histone

deacetylases, recent work has demonstrated that sirtuins have a range of enzymatic activities and substrates through which they maintain cellular homeostasis. Our lab has previously uncovered that sirtuins exert both proviral and antiviral roles during infection, but the underlying mechanisms remain unknown. Sirtuins may have compensatory effect, with one sirtuin modulating the levels and functions of other sirtuins. Here, we developed a targeted mass spectrometry assay based on parallel reaction monitoring that allows simultaneous detection and quantification of all human sirtuins. We demonstrate the sensitivity of this assay for detecting these low abundance proteins and its generalizability to different cell types. We profiled changes after treatment with the antiviral compound, trans-resveratrol, which activates sirtuins. Further, we monitored sirtuin dynamics during cellular remodeling events: differentiation of monocyte-like cells into macrophages, as well as during infection with two herpesviruses, herpes simplex virus 1 (HSV1) and human cytomegalovirus (HCMV). Finally, we used this assay in our investigation of the nuclear sirtuin 6 (SIRT6) during infection. Using SIRT6 knockdown, CRISPR-mediated knockout, and overexpression, we assessed the effect of SIRT6 on HCMV replication in primary human fibroblasts (MRC5). We further investigated whether perturbing SIRT6 levels drives compensatory effects by other sirtuins. Lastly, our quantitative sirtuin measurements were integrated with investigations of SIRT6 protein interactions and localization during the progression of HCMV infection.

Contact: Matthew Tyl, mtyl@princeton.edu

P04: Cancer Proteomics

P04.01 Proteomic Analysis Reveals that Musashi-2 Promotes Signaling Pathways Associated with Cell Proliferation and Metastasis in Neuroblastoma Cells

Pritha Bagchi (1), Adeiye Pilgrim (1), Selma Cuya-Smith (2), Dongdong Chen (1), Nicholas Seyfried (1), Robert Schnepp (1)
(1) Emory University, Atlanta, Georgia, United States
(2) Kennesaw State University, Kennesaw, Georgia, United States

Neuroblastoma (NBL) is the most common extracranial solid tumor of childhood. Many patients present with widespread metastasis and may have secondary malignancies and other toxicities in adulthood even after extensive multimodal treatment. Thus, there is a dire need for druggable targets, which minimize toxicity and maximize efficacy in the treatment of NBL. RNA-binding protein Musashi-2 (MSI2) has been shown to be a mediator of numerous cancer-related hallmarks, but its role in NBL remains unexplored. Based on our preliminary data from cell-based assays, MSI2 appears to be an oncogenic driver in NBL as well. Hence, we explored the MSI2-influenced proteome to understand the mechanism and to identify potential NBL-specific therapeutic targets. We genetically reduced MSI2 expression in two NBL cell lines, SK-N-BE(2)C and Kelly, using short hairpin RNA (shRNA) clones with two unique targets within the MSI2 coding sequence. Subsequently, we performed unbiased proteomics analysis using LC-MS/MS. The label-free quantification algorithm in MaxQuant was employed for protein identification and quantification; statistical analysis was performed by Perseus and DAVID. The proteomics data verify successful knockdown of MSI2 protein in both cell lines. Gene Ontology (GO) analysis reveals that MSI2 knockdown affected different pathways in a cell-line specific manner. In SK-N-BE(2)C cells, processes related to cell cycle progression and cell growth are downregulated (e.g., DNA replication and telomere maintenance via recombination), whereas mechanisms that induce apoptosis are upregulated (e.g., glycosphingolipid metabolic process and lysosome). In Kelly cells, downregulated proteins show overrepresentation for “nervous system development” and “postsynaptic density”, indicating attenuation of metastasis, while upregulated proteins are related to “mitochondrial matrix” and “metabolic pathways”. In summary, NBL cells with genetically reduced MSI2 demonstrate cell cycle arrest, increased apoptosis, and altered metabolism. Thus, MSI2 and its downstream proteins could potentially be targeted to develop cancer therapeutics, especially to inhibit tumor growth and metastasis in neuroblastoma.

Contact: Pritha Bagchi, pritha.bagchi@emory.edu

P04.02 Robust, Reliable, Reproducible Sample Preparation Enabling Confident LC-MS Analysis of Proteins in FFPE Tissues

DEBADEEP BHATTACHARYYA (1), Eugenio Daviso (1)
(1) Covaris, Woburn, Massachusetts, United States

Formalin-fixation and paraffin-embedding (FFPE) of tissue samples are routinely used in clinical histopathology for preserving tissue architecture. However, paraffin removal is needed for high quality LC-MS data. Typically Xylol is used but has health risks and compromises reproducibility. Better sample preparation removes unwanted matrix components while ensuring reliability, reproducibility and robustness. Covaris' Adaptive Focused Acoustics® (AFA®) Technology can enable a modified protein aggregation capture (PAC) protocol with remarkable robustness, reliability, and reproducibility. FFPE tissue scrolls can be incubated in a Covaris plate post-treatment with “Covaris Tissue Lysis Buffer”. The plates were processed with the Covaris LE220-Plus Focused-ultrasonicator for deparaffinization and DNA

shearing. PAC protocol enabled extraction from FFPE lysate and Sera Mag Speed beads induced on-bead protein aggregation followed by digestion Trypsin/LysC before conducting LC-MS analysis. This work demonstrates a non-toxic, scalable, and robust method for LC-MS based analysis of FFPE scrolls with human adenoma samples that were processed in AFA compatible 96-well plate. Heating and AFA based sonication ensured reversal of crosslinks and paraffin removal from tissue samples. Protein purification was achieved using the PAC protocol and a tryptic 'on-bead' digestion resulted in a clean peptide solution for LC-MS analysis. A comparison of the xylol-free workflow with that of the common xylol deparaffinization resulted in highly similar peptide, protein identification rates, and protein abundance ranges. On average, 3,900 proteins and 20,842 peptides per single run measurement were quantified. A comparison of this method with the xylol-based protocol showed that crosslink reversal was equally efficient in both protocols. The developed protocol not only offers ability to characterize and quantify a large number of proteins from complex tissue matrices, but also can support low-medium-high throughput capabilities as needed by the organization.

Contact: DEBADEEP BHATTACHARYYA, debadeep@gmail.com

P04.03 Deep Plasma Proteomics at Scale Enabling Proteogenomic Analyses in a Lung Cancer (NSCLC) Study

Harendra Guturu (1), Mahdi Zamanighomi (1), Jian Wang (1), Amir Alavi (1), Tristan Brown (1), Daniel Hornburg (1), Moaraj Hasan (1), Shadi Ferdosi (1), Khatereh Motamedchaboki (1), Margaret Donovan (1), Theodore Platt (1), Ryan Benz (1), Asim Siddiqui (1), Serafim Batzoglou (1)
(1) Seer, Inc., Redwood City, California, United States

Our ~20,000 genes encode over one million protein variants, given alternative splice forms, allelic variation, and protein modification. Though large-scale genomics studies have expanded our understanding of biology, similarly, scaled deep and untargeted proteomics studies of biofluids have remained impractical due to complex workflows. To address this need, we have previously described Proteograph(TM), a novel platform that leverages the protein-corona interactions of nanoparticles for deep and untargeted proteomic sampling at scale. Using Proteograph(TM) in a non-small cell lung cancer (NSCLC) cohort, we previously conducted a deep interrogation of plasma from 141 subjects: 61 early-stage NSCLC subjects and 80 non-cancer controls. We identified 2,499 plasma proteins, with 1,992 present in ≥ 25% of the samples. Leveraging this data, we created a biomarker classifier distinguishing NSCLC from controls with area under the receiver operating characteristic curve of 0.91 [1]. In this study, we now re-analyze the data with the more sensitive DIA-NN software to enhance protein depth while preserving the accuracy of the classifier. In addition, to show the added value of our platform in combination with genomic data, we integrate previously sequenced exome data with our proteomic data to build a multi-modal proteogenomic deep learning classifier. Our results outline proteogenomic workflows for robust biomarker discovery and cohort subtyping. The Proteograph(TM) platform interrogates the plasma proteome at previously impractical combinations of scale, depth and coverage, and enables the development of improved classification models and the study of proteogenomics. [1]Blume et al. Nat. Comm. (2020)

Contact: Harendra Guturu, hguturu@seer.bio

P04.04 Deep, Rapid and Unbiased Plasma Proteomics with Differential Analysis of Proteoforms Enabling Proteogenomic Studies in a NSCLC Lung Cancer Study

Yingxiang Huang (1), Margaret Donovan (1), John Blume (1), Jian Wang (1), Marwin Ko (1), Ryan Benz (1), Theodore Platt

(1), Juan Cuevas (1), Serafim Batzoglou (1), Asim Siddiqui (1), Omid Farokhzad (1)
(1) Seer Inc., Redwood City, California, United States

Comprehensive assessment of the human proteome remains elusive due to multiple forms of a protein, each of which can serve distinct functions, arising from alternative splicing, allelic variation, and protein modifications. Characterization of the variable protein forms, or proteoforms, will expand our understanding of the molecular mechanisms underlying disease, however identification of these variable forms requires unbiased protein coverage at sufficient scale. Scalable, deep, and unbiased proteomics studies have been impractical due to cumbersome and lengthy workflows required for complex samples, like blood plasma. Here, we demonstrate the power of Proteograph in a proof-of-concept proteogenomic analysis of 80 healthy controls and 61 early-stage non-small-cell lung cancer (NSCLC) samples to dissect differences between protein isoforms arising from alternative gene splicing, as well as the identification of novel peptides arising from allelic variation. Processing 141 plasma samples with Proteograph yielded 22,993 peptides corresponding to 2,569 protein groups. Using peptides with significant abundance differences ($p < 0.05$; Benjamini-Hochberg corrected), we extracted proteins comprised of peptides where at least one peptide had significantly higher plasma abundance, and another significantly lower plasma abundance in controls vs. cancer, resulting in a set of four proteins. For three of these proteins (BMP1, C4A, C1R), the abundance variation is possibly explained by underlying protein isoforms. To identify protein variants, we performed exome sequencing on 29 individuals from the NSCLC study, created personalized mass spectrometry search libraries for each individual, and identified 464 protein variants. In conclusions, Proteograph can generate unbiased and deep plasma proteome profiles that enable identification of protein variants and peptides present in plasma, at a scale sufficient to enable population-scale proteomic studies.

Contact: Yingxiang Huang, yhuang@seer.bio

P04.06 Low-Input Proteomic Phenotyping of a Metabolically Defined Putative Mammary Epithelial Stem Cell Population

Matthew Waas (1), Pirashaanthy Tharmapalan (1), Rama Khokha (1), Thomas Kislinger (1)
(1) Princess Margaret Cancer Centre, Toronto, Ontario, Canada

Breast cancer, a heterogenous disease of multiple origins, is the most common cancer in women. Studies of the various subtypes of breast cancer have revealed similarities between the molecular profile of these subtypes to distinct populations of mammary gland epithelial cells. This evidence supports the hypothesis that the different epithelial cell populations serve as the cell-of-origin for the corresponding subtypes of cancer. Our hypothesis is that defining functional and molecular identities of distinct stem cell populations within the mammary gland epithelium will reveal potential markers and molecular vulnerabilities. Recently, we identified a metabolically defined subpopulation of mammary epithelial basal cells which harbor an increased capacity for colony formation representing a novel putative progenitor/stem population. We have adapted an innovative droplet-digestion sample preparation strategy relying on commercially available consumables and, in combination with systematic optimization of sample preparation, instrument acquisition, and data analysis methods, we reproducibly obtain >3,000 proteins on a Thermo Fusion mass spectrometer from 500 sorted cells. We have performed a repeatability study to characterize the observed variability inherent to sample preparation and data acquisition. The median correlation between cell type replicates was 0.96 and 0.91 within and between days, respectively. Finally, we have applied this platform to profile the proteomic phenotype of the metabolically defined putative epithelial stem/progenitor cells from six individual mice (~2000 cells/mouse). We identify over 3500 proteins per sample with high repeatability. Comparisons to

control populations reveal 9 potential cell surface markers and 5 potential druggable targets. Future studies will compare the proteomic phenotype to stem cell populations sorted with other established markers and will investigate the functional capacity of this population. We anticipate the results of this research will harmonize our understanding of mammary stem cell biology and support breast cancer prevention, prognosis, and treatment efforts.

Contact: Matthew Waas, waasmatthew@gmail.com

P04.07 Characterizing the Proteomic Response of Models of High-Grade Serous Tubo-Ovarian Cancer to Taxane Treatment

John Muroski (1), Oscar Pundel (2), Krystal Lum (1), Benjamin Neel (2), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States
(2) New York University, New York, New York, United States

Ovarian cancer is the deadliest form of gynecological cancer, with high-grade serous tubo-ovarian cancer (HGSC) presenting itself as the most common form. Advancements in treating HGSC have shown less progress in overall survival, due in part to the high heterogeneity of cancer cells. Recent advances in organoid generation have allowed the development of genetic models that represent different subclasses of HGSCs. Major subtypes include a homologous-recombination (HR)-deficient model ($Trp53^{-/-}$;$Brca1^{-/-}$;Myc^{OE}) and a model of a poorly characterized subtype ($Trp53^{-/-}$;$Pten^{-/-}$;$Nf1^{-/-}$). Previous work to elucidate features of this model unexpectedly revealed a significantly greater susceptibility to treatment with a commonly used anti-cancer drug, Paclitaxel, than compared to other subclass models. Here, we used proteomics to analyze 2D and 3D cell culture model systems of HGSC in order to better understand drug hypersensitivity and cell autonomous/non-autonomous effects on HGSC evolution. Alterations in the proteomes of cell and organoid cultures, with and without Paclitaxel treatment, were analyzed using quantitative mass-spectrometry. Over 700 proteins significantly altered in abundance and major alterations in the HR-deficient model indicate an increase in proteins related to cytoskeletal organization, extracellular components, and immune signaling. Protein network analysis highlights enriched processes such as mitosis and immune signaling. In contrast, proteins enriched in the $Trp53^{-/-}$;$Pten^{-/-}$;$Nf1^{-/-}$ model are involved in oxidation reduction and lipid metabolism. These findings indicate that alterations in cytoskeletal compositions do correlate with differences in Paclitaxel sensitivity. To further define the system-wide changes occurring in drug-treated HGSC models, thermal proximity coaggregation analysis (TPCA) is used to investigate activated or inhibited cellular pathways in these model systems.

Contact: John Muroski, jm3100@princeton.edu

P04.08 Proteogenomics reveal over-expression of HLA-I in cancer

Ying Wang (1), David Fenyo (1)
(1) New York University, New York, New York, United States

Accurate quantification of HLA class I gene expression is important to understand the interplay of tumor microenvironment of antitumor cytotoxic T cell activity. Because HLA-I sequences are highly variable, common RNAseq and mass spectrometry based quantification workflows using standard genome and protein sequence references do not provide HLA-I allele specific quantifications. This study improved the quantification processes by using personalized HLA-I nucleotide and protein reference sequences based on the subjects' HLA-I genotypes, and surveyed tumor and adjacent normal samples

from patients from nine cancer types. DDA-MS data was validated to be sufficient to estimate HLA-A protein expression at allele level. B15 allele group was over-represented in LUAD patient with allele frequency of 76% more than the expected value. HLA-I proteins were observed to be at significantly higher levels in tumors compared to adjacent normal tissues by 41% - 63% of head and neck, endometrial, and kidney tumors. These higher protein levels were mostly driven by increased level of HLA-I transcripts from both alleles. Most immune cell types are universally enriched in HLA-I high tumors while endothelial and neuron cells showed divergent relationships with HLA-I. Pathway analysis revealed that tumor senescence and autophagy activity was likely to influence the level of HLA-I proteins in glioblastoma. Most proteins correlated to HLA-I protein levels are directly involved in HLA-I function in immune response and cell death, while co-expression of glycosylation genes with HLA-I are exclusively observed at protein level.

Contact: Ying Wang, Ying.Wang2@nyulangone.org

P05: Cardiovascular Proteomics

P05.01 CellSurfer platform reveals the first experimentally derived surfaceome map of primary human cardiac cells
Linda Berg Luecke (1), Matthew Waas (2), Jack Littrell (3), Rebekah Gundry (3)
(1) Medical College of Wisconsin, Milwaukee, Wisconsin, United States
(2) Princess Margaret Cancer Centre, Toronto, Canada
(3) University of Nebraska Medical Center, Omaha, Nebraska, United States

In the heart, cell surface glycoproteins in cardiomyocytes and cardiac fibroblasts are essential for sustaining normal organ function and play critical roles in cardiac development, disease, and drug uptake. However, the lack of a detailed cell type- or chamber-resolved view of the cell surface proteome of the adult human heart currently limits discovery of new targets for precision drug delivery and the development of practical approaches for studying how different cell types contribute to the development of cardiac disease. CellSurfer, a new analytical platform, was applied to cardiac cells isolated from human hearts. Briefly, cell surface N-glycoproteins on ~1 million cells were labeled, digested, selectively enriched using streptavidin magnetic beads, cleaned using SP2, and analyzed by MS. Sample preparation was automated using liquid handling robotics. MS data were analyzed using Proteome Discoverer, Spectronaut, MSstats, and R. Results were curated and annotated using Veneer. Integrating CellSurfer with an optimized strategy for isolating intact primary cardiomyocytes and fibroblasts from human donor heart tissue resulted in the generation of the first chamber-, cell-type-, and patient-specific map of the cell surface N-glycoproteome in the adult heart. Overall, >1100 cell surface N-glycoproteins were detected, including 48 and 21 putative selective markers for cardiomyocytes and cardiac fibroblast, respectively. Novel monoclonal antibodies generated for one cardiomyocyte marker uniquely localize to cardiomyocytes within human heart tissue sections and stem cell derivatives, suggesting its value for cell-type specific targeting and immunophenotyping. Comparisons of explanted cardiac fibroblasts within the first three passages reveals previously undescribed remodeling of the surfaceome, justifying caution when using cultured cells. These data represent the first major step towards a comprehensive, donor, cell-type, subtype, and chamber-resolved reference map of cell surface phenotypes in the adult human heart and reveal new targets for immunophenotyping, drug delivery, and benchmarking explanted cells and stem cell derivatives.

Contact: Linda Berg Luecke, lbergluecke@mcw.edu

P05.02 Collagen-derived Matricryptin Modulates MMP-9 Substrate Preference in Left Ventricles
Sirin Nazan Cakir (1), Lisandra de Castro Brás (1)
(1) East Carolina University, Greenville, North Carolina, United States

Purpose: Matrix metalloproteinases (MMPs) are an endogenous family of enzymes that contribute to matrix remodeling in several diseases, including cardiomyopathies. After cardiac ischemia, MMP induction and activation may generate a predominantly proteolytic environment, leading to degradation of extracellular matrix (ECM). Matricryptins are biologically active peptides, generated from ECM proteolysis, able to regulate cell function and survival. We recently identified a collagen-derived matricryptin (p1159), gradually forms post-myocardial infarction (MI) and is a substrate of MMP-9. Mice treated with p1159 post-MI show both improved function and reduced dilation of the left ventricle (LV). To evaluate whether p1159 presence modulates MMP-9 substrate preference in the myocardium, we used N-terminal amine isotopic labeling of substrates (TAILS) proteomics to compare and quantify the LV peptidome. Methods: iTRAQ-TAILS multiplex quantitative proteomics was applied to naïve LV tissues (male, C57Bl/6 mice, 6-8 months old, n=6/group) to assess effects of p1159 in

MMP-9 substrate preference. One mg of protein was used for each sample ± p1159 ± MMP-9; samples were processed according to the previously published protocol (1). The samples were run on a Thermo Scientific Q-Exactive Orbitrap, then analyzed using Proteome Discoverer 2.5™. Results: We identified 837 MMP-9 substrates in naïve mouse LV. Of these, 49 were ECM molecules and 47 were plasma membranes. For the first time, cofilin-1 and nischarin were identified as potential MMP-9 cardiac substrates. The presence of p1159 changed the relative abundance of MMP-9 cleavage products. Conclusion: This study both identified novel MMP-9 cardiac substates and determined p1159 presence changes MMP-9 substrate preference. Future studies will identify which substrates may be responsible for the peptide beneficial effects on LV remodeling. Literature cited: (1) Kleifeld, Oded et al. “Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates.” Nature protocols vol. 6,10 1578-611. 22 Sep. 2011, doi:10.1038/nprot.2011.382

Contact: Sirin Nazan Cakir, cakirs18@students.ecu.edu

P05.03 Proteomics to understand cholesterol efflux capacity of High density Lipoprotein (HDL) particles from human plasma

Zeeshan Hamid (1), Anamika Gangwar (2), Kip Zimmerman (1), Sneha Deodhar (2), Kit Wallis (1), Anand Rohatgi (2), Michael Olivier (1)
(1) Wake Forest University School of Medicine, Winston Salem, North Carolina, United States
(2) University of Texas Southwestern Medical Center, Dallas, Texas, United States

Cardiovascular disease (CVD) remains the leading cause of death worldwide and atherosclerosis has been shown to be a major underlying factor. HDL is known to exert various anti-atherosclerotic functions with its strong ability to efflux cholesterol from periphery and transport it to the liver for excretion. Cholesterol efflux capacity (CEC) of HDL has been extensively studied in various animal models, but has not been studied in detail in humans. Here, we identified HDL plasma proteins with high throughput mass spectrometry and associated protein abundance with measured CEC. A bottom up proteomics workflow was followed involving protein precipitation from HDL particles, followed by reduction, alkylation and trypsin digestion. Resulting peptides were cleaned up using C18 desalting tips, and data was acquired on a Q-Exactive mass spectrometer with an easyspray nano-LC. Raw data analysis was done with MaxQuant using H. Sapiens uniprot protein database. We obtained quantitative data for 540 proteins across all samples. A series of linear mixed models were computed on the log-transformed and normalized abundance of each distinct protein. Protein abundance was the outcome in our model and CEC value was the predictor. The model included age as a covariate as well as the individual identifier as a random effect to account for repeated measures. A Benjamini-Hochberg FDR was computed on the raw p-values to adjust for multiple comparisons, and proteins meeting an FDR < 0.05 were retained for further downstream analysis. Of 540 quantified proteins, 47 proteins were either positively or negatively associated with the CEC values. Using DAVID enrichment analysis of these 47 proteins, we observed strong enrichment of proteins involved in lipoprotein metabolic process, high density lipoprotein particle remodeling and reverse cholesterol transport (APOA2, APOC2, APOC3, APOE and LCAT). Overall, this study has highlighted some significant associations of some proteins involved in CEC of HDL particles.

Contact: Zeeshan Hamid, zhamid@wakehealth.edu

P05.04 Cardiac proteomics reveals sex chromosome-dependent differences between males and females that arise prior to gonad formation

Josiah Hutton (1), Wei Shi (2), Xinlei Sheng (1), Kerry Dorr (2), James Emerson (2), Haley Davies (2), Tia Andrade (2), Lauren Wasson (3), Todd Greco (1), Yutaka Hashimoto (4), Joel Federspiel (1), Zachary Robbe (2), Xuqi Chen (5), Arthur Arnold (5), Frank Conlon (2), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States
(2) UNC-Chapel Hill, Chapel Hill, North Carolina, United States
(3) UNC-Chapel Hill, Chapel Hill, North Carolina
(4) Princeton University, Princeton, New Jersey
(5) University of California, Los Angeles, Los Angeles, California, United States

Sex disparities in cardiac homeostasis and heart disease are well documented, with differences attributed to actions of sex hormones. However, studies have indicated sex chromosomes act outside of the gonads to function without mediation by gonadal hormones. Here, we performed transcriptional and proteomics profiling to define the differences between male and female mouse hearts. We first applied this workflow to the C57BL/6J mouse strain, where we observed that there were fewer protein expression differences between male and female hearts compared to the number of differentially expressed mRNA transcripts. We next performed quantitative proteomic analysis of eight different mouse strains and determined that proteins involved in metabolic processes were significantly enriched in male mouse hearts, while proteins involved in nuclear membrane structure were significantly enriched in female mouse hearts. We then utilized the Four Core Genotype (FCG) mouse strain to demonstrate, contrary to current dogma, cardiac sex disparities are controlled not only by sex hormones but also through a sex-chromosome mechanism. Using Turner syndrome (XO) and Klinefelter (XXY) mouse models, we determined that the sex-chromosome pathway is established by X-linked gene dosage. Additionally, we demonstrated that cardiac sex disparities occur at the earliest stages of heart formation, a period before gonad formation. Throughout these analyses, alpha-1B-glycoprotein (A1BG) was consistently observed to be significantly increased in female hearts compared to male hearts. We utilized CRISPR/Cas9 to generate cardiac conditional loss-of-function alleles of A1BG, and we demonstrated that that the loss of A1BG in the heart leads to cardiac defects in females, where we specifically observed defects in the outer left ventricle wall function and thickness, but these cardiac defects were not observed in A1BG-KO males. These results provide a powerful resource for studying sex based cardiac disease states, and we demonstrated the critical role of increased A1BG expression in female heart function.

Contact: Josiah Hutton, josiahh@princeton.edu

P05.05 Integrating a high-throughput proteomic workflow for in vitro screening of cell-based assays

Saeed Seyedmohammad (1), Alejandro Rivas (1), Simion Kreimer (1), Qin Fu (1), Jennifer Van Eyk (1)
(1) Cedars Sinai Medical Institute, Los Angeles, California, United States

Introduction: We have developed a robust and reproducible high-throughput (HTP) workflow for large-scale protein assay. This bottom-up proteomics workflow directs automated processing of cells grown in 96-well plates for LC-MS analysis. The HTP system leverages the COVARIS LE220 Plus sonication system for cell lysis and solubilization in tandem with the Beckman i7 liquid handling workstation for sample preparation. Methods: AC16 human cardiomyocytes were grown to different cell densities in 96-well plates. Sonication parameters were optimized for maximal performance, and suitability of different solubilization buffers were determined on COVARIS 96 AFA-Tube TPX 150 Plates. The i7 automation workstation were employed in conjunction with S-trap digestion for protein denaturation, Cys reduction, alkylation, and trypsin digestion

(1:20 µg ratio of trypsin/protein) in a humidity regulated on-deck incubator (2 hours, 47°C). Data analysis involved BCA measurements and LC-MS acquisition on Thermo Orbitrap Lumos Micro system coupled to capLC column at 60-minute gradient. Results: Average protein concentration of 1 µg/µl (n=4, 1.7 STD) were obtained from approximately 100,000 cells sonicated in each well of a 96-well COVARIS plate. Optimal cell lysis was observed at the representative AFA settings condition (200 PIP, 50 DF and 1,000 CPB) in 5 % SDS using the V-shaped AFA-Tube TPX 150 plate (PN-520291, COVARIS). A specifically designed cell-based prototype plate (PN-520310, COVARIS) was also optimized to achieve DNA sheering and full cell lysis (half a fold increase with higher reproducibility and less miss-cleavages), by incorporating an initial high power-burst cycle prior to standard AFA sonication. Conclusions: By employing the COVARIS sonication system, simultaneous cell lysis of multiple samples was achieved. The developed HTP pipeline can be applied to both label free and stable isotope-labeling as well as investigating important key regulators in cell-based assays, including post-translational modifications. Furthermore, drug targets and cellular perturbations could be rapidly investigated and studied.

Contact: Saeed Seyedmohammad, saeed.seyedmohammad@cshs.org

P06: Cellular Signaling and Systems Biology

P06.01 Heterogeneity of PTEN and PTEN-alpha expression in prostate cancer

Alakesh Bera (1), Meera Srivastava (2), Madhan Subramanian (1), Eric Russ (1), Surya Radhakrishnan (1), Adam Landa (1), Nahbuma Gana (1), Lukas Bubendorf (2)
(1) Uniformed Services University of the Health Sciences, Bethesda, Maryland, United States
(2) University Hospital Basel, Basel, Swaziland

Prostate cancer (PrCa) is among the most common cancers and remains the second leading cause of cancer-associated death in men worldwide. Several genetic alterations are known to be responsible for the progression, aggressiveness, chemo-resistance, and heterogeneity of prostate cancer. For instance, phosphatase and tensin homolog (PTEN) is an essential tumor suppressor gene which regulates the PI3K-AKT pathway and whose mutation and deletion are among the most common genomic aberrations in prostate cancer. Moreover, recent studies have indicated that the longer isoform, termed PTEN-Long (PTEN-Ln) or PTEN-alpha, also plays a critical role in normal cellular functions such as proliferation and apoptotic pathways. This is significant as our findings from tissue micro-array and cellular assays, which compared the protein expression of the standard PTEN (now termed PTEN-st) and PTEN-Ln, have indicated a robust correlation between loss of PTEN expression and the balance between PTEN-st and PTEN-Ln protein expression; showing adverse pathological features with aggressive Gleason grades of prostate cancers. In other words, in patients with prostate cancer, the ratio of PTEN-st to PTEN-Ln expression drastically decreases. We also found a novel Alu element insertion in the PTEN gene of a prostate cancer cell line, similar to the reported Alu element insertions at a PTEN hotspot in Cowden syndrome. In this current study, we have determined the molecular basis of PTEN-standard and PTEN-Long expression variability and their genetic aberrations, including the Alu insertion mechanism, which leads to the loss of functional PTEN activity. Furthermore, efforts have also applied to determine the signaling cascade which is involved in regulating this genetic aberration and how to inhibit that signaling pathway. Finally, our current findings could help in the development of both clinical diagnostic capabilities and in predicting patient response to targeted therapy in prostate cancer.

Contact: Alakesh Bera, alakesh.bera.ctr@usuhs.edu

P06.02 Spatial proteomics for atlasing healthy human kidney

Danielle Gutierrez (1), Melissa Farrow (2), Jamie Allen (1), Carrie Romer (1), Angela Kruse (1), Elizabeth Neumann (1), Tina Tsui (1), Mark deCaestecker (3), Richard Caprioli (1), Jeffrey Spraggins (4)
(1) Vanderbilt University, Nashville, Tennessee, United States
(2) Vanderbilt University Medical Center, Nashville, Tennessee, United States
(3) Vanderbilt University Medical Center, Nasville, Tennessee
(4) Vanderbilt University, Nashville, Tennessee

The National Institutes of Health Human BioMolecular Atlas Program was established to drive the large-scale mapping of healthy human tissue from the level of single cells to whole organs. As part of this effort, we assessed an initial set of tumor resections comprising histologically normal tissue from 35 patients. The body mass index (BMI) of these patients ranged from normal (20-25) to severe obesity (>40), and additional conditions, such as hypertension and diabetes, existed within the cohort. While it is well established that obesity can negatively impact kidney health, the molecular mechanisms responsible have yet to be comprehensively elucidated. Additionally, the effect of confounding comorbidities can further

obscure these pathomechanisms. We applied bulk proteomics and systems biology workflows to interrogate molecular changes associated with BMI, comorbidities, and sample composition (i.e., percent cortex versus medulla). The initial data reveal major trends in changes to oxidative phosphorylation and immune response associated with BMI. Oxidative stress can trigger inflammation, leading to a negative feedback loop and furthering tissue injury. These data will be interrogated further to understand molecular changes with increasing levels obesity and other comorbidities. Furthermore, we will implement spatial proteomics and imaging mass spectrometry to resolve a spatially specific understanding of the molecular effects of these conditions on kidney tissue. Spatial omics will enable the association of specific molecular mechanisms with tissue functional regions, such as glomeruli. These approaches will be applied to additional patient samples in order to map the proteomic landscape of healthy kidney in relation to BMI, comorbidities, and tissue functional regions.

Contact: Danielle Gutierrez, d.gutierrez@vanderbilt.edu

P06.03 Proteomic characterization of the virus microenvironment and proximal immune signaling

Bokai Song (1), Xinlei Sheng (1), Peter Metzger (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Efficient virus dissemination relies on the release of infectious viral particles from the infected cells and attachment to the neighboring cells. Viruses have acquired mechanisms to alter the proteome of infected cells to promote infection progression. Upon infection, the infected cells secrete various chemokines and cytokines to signal the immune system to prevent the spread of infection, while viral proteins are also secreted to dampen the host response. However, it remains largely unknown how cells either in close or distant proximity to the infected cell respond to the infection. No study has yet attempted to investigate cellular proteomes of the virus microenvironment. Here, we integrated fluorescence-activated cell sorting (FACS), quantitative mass spectrometry, microscopy, and molecular virology assays to provide the first view of cellular proteome alterations within the virus microenvironment. We generated fibroblasts secreting a lipid-permeable mCherry tag, which we infected with GFP-tagged human cytomegalovirus (HCMV). Using a strategy that involved the co-culture of these cells with wild type cells, we distinguished and enriched infected, neighboring, and distal cells. By defining the differential proteomes of these cell populations, we discovered that multiple interferon-inducible proteins are down-regulated in infected and neighboring cells compared to the distal cells, suggesting a viral strategy to promote the spread of infection. We corroborated these results by confocal microscopy analyses. Next, we performed superinfection assays with both HCMV and herpes simplex virus type 1 (HSV-1), observing increased susceptibility of neighboring cells to superinfection compared to distal cells. Furthermore, neighboring cells also display alterations in cell cycle progression when compared to distal cells. Overall, our study is the first proteomic characterization of the virus microenvironment, revealing global alterations in immunity-related pathways. These findings provide insights into how virus infection reshapes the surrounding environment to facilitate progression of infection.

Contact: Bokai Song, bokais@princeton.edu

P06.04 Defining Toll-like Receptor Signaling Networks in Cellular Space and Time Using Quantitative Mass Spectrometry

Trisha Tucholski (1), Aleksandra Nita-Lazar (1)
(1) Functional Cellular Networks Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland, United States

Pattern recognition receptor (PRR) signaling initiates host innate immune response upon the recognition of pathogen-associated molecular patterns (PAMPs). Toll-like receptors (TLRs) are an important class of evolutionarily-conserved PRRs localized to the surface and endosomes of professional innate immune cells. Protein expression data alone may not sufficiently describe regulation of early innate immune response. Quantitative experimental data detailing the subcellular location and movement of signaling network components in the cell as innate immune response progresses are important for understanding how hosts initially respond to infection. Herein, we use subcellular fractionation combined with quantitative mass spectrometry (MS)-based proteomics to define the spatial proteome of human monocyte-derived macrophages, an important professional innate immune cell type. Preliminarily, we have demonstrated reproducible enrichment of organelle protein markers using the spatial proteomics workflow. We have also identified important proteins involved in TLR signaling, including components of the myddosome (e.g., MyD88, IRAK1, IRAK4). We show that the proteins TAB1, TAB2, TAK1, which associate during TLR signaling, are co-enriched in subcellular fractions. Furthermore, we aim to characterize early TLR signaling by incorporating spatial proteomic analysis of macrophages following time-course stimulation of TLR4, and important sensor of lipopolysaccharides derived from Gram negative pathogens. We will monitor protein subcellular location at early timepoints post-stimulation (15, 30, 45, and 60 min) to capture early protein translocation events. Global spatiotemporal proteomic profiles as well as targeted absolute quantitative data will be collected. The quantitative data will eventually be used as input for computational modeling of innate immune signaling networks in the spatial and temporal dimensions. The resulting quantitative spatiotemporal proteomic data and computational models of innate immune response can ultimately contribute to improved understanding of how proteins cooperate in space and time to enact cellular processes.

Contact: Trisha Tucholski, trisha.tucholski@nih.gov

P07: Computational Proteomics and Data Science

P07.01 Characterizing and Correcting Batch Effects in Large Scale Proteomics Data.

Amir Alavi (1), Harendra Guturu (1), Mahdi Zamanighomi (1), Tristan Brown (1), Jian Wang (1), Sam Cutler (1), Biao Li (1), Khatereh Motamedchaboki (2), Asim Sidiqqui (1), Serafim Batzoglou (1)
(1) Seer, Inc., Redwood City, California, United States

Recent advances in liquid chromatography mass spectrometry (LCMS)-based proteomics have enabled the efficient profiling of thousands of proteins from single LCMS runs. The ability to run untargeted, high throughput LCMS experiments has opened the door to large-scale cohort studies for biomarker and drug target discovery. When conducting large-scale cohort studies, technical confounding can be introduced as samples are run across different MS instruments, LC columns, dates, and geographic locations. In order to integrate these samples across datasets for joint analyses, one needs to both diagnose this batch effect and apply methods to correct for it. In order to benchmark batch effect characterization and batch effect correction methods for protein abundance data, we collect a batch-diverse dataset using Seer's Proteograph™ Product Suite. Our dataset includes over 800 LCMS runs across two biosamples, multiple nanoparticles, three MS instruments, and eight LC columns. We evaluate the presence of a batch effect using multiple batch effect diagnosis methods, including Principal Components Analysis-based approaches, local-neighborhood diversity measures, machine learning classifier-based methods, and a statistical hypothesis testing framework. We show that technical covariates such as mass spectrometry machine and liquid chromatography column have a significant contribution to the variability observed in protein abundance in this data. To address the observed batch effect, we survey the space of batch effect correction methods for protein abundance data. These include traditional methods often used in proteomics and genomics such as ComBat, dimensionality reduction methods, and nearest neighbors-based approaches. In addition, we apply a novel deep learning-based batch correction method along with published methods, including metric learning-based methods and adversarial approaches. We find that our proposed deep learning-based approach can remove technical variation while preserving biological variation better than prior methods and conclude with a discussion on the advantages and disadvantages of these different techniques.

Contact: Amir Alavi, aalavi@seer.bio

P07.02 Deep Learning enables a Joint Embedding of Peptides and MS/MS Spectra allowing for Exploratory Analyses via UMAP

Tom Altenburg (1), Maroua Filali (1), Jennifer I. Daniel Onwuchekwa (1), Thilo Muth (2), Bernhard Y. Renard (1)
(1) Hasso Plattner Institute, Potsdam, Germany
(2) Federal Institute for Materials Research and Testing (BAM), Berlin, Germany

MS-based proteomics, in contrast to other omics types, typically requires additional efforts in terms of peptide identification before the actual data exploration, or even biological interpretation can take place. Here, we present a deep learning approach that jointly embeds MS/MS spectra and peptides into the same vector space such that embeddings can be compared easily via euclidean distances. As a result, manifold learning or clustering approaches can be directly applied to the embeddings of MS/MS spectra. Note, our approach differs from existing spectral embedding approaches because our joint embedding yields meaningful spectrum embeddings. In particular, the manifold of spectrum embeddings is tied to its

peptide counterpart. We share a collection of UMAP-visualizations of our embeddings that render useful for sample exploration, quality control, and annotation of peptide properties.

Contact: Tom Altenburg, tom.altenburg@hpi.de

P07.03 Subtyping the Aging Brain Based on an Unbiased Brain Proteome Network Clustering Approach

E. Kathleen Carter (1), Eric B. Dammer (2), Duc M. Duong (2), David A. Bennett (3), James J. Lah (2), Allan I. Levey (4), Nicholas T. Seyfried (1)
(1) Emory University School of Medicine, Department of Biochemistry, Atlanta, Georgia, United States
(2) Emory University, Atlanta, Georgia, United States
(3) Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, United States
(4) Emory University School of Medicine, Department of Neurology, Atlanta, Georgia, United States

Aging is the largest risk factor for Alzheimer's disease (AD)--a progressive neurological disease characterized by memory loss and cognitive dysfunction. Recent studies have shown that it may be possible to define subtypes of AD based on clustering of brain transcriptomic data. Here, we aimed to develop an approach to classify the brain proteome in a community-based cohort of normal aging that included AD cases using unbiased network clustering to uncover distinct proteomic subtypes. Four hundred dorsolateral prefrontal cortex tissues were analyzed from the Religious Orders Study and Memory and Aging Project (ROSMAP), a longitudinal study of aging, by tandem mass tag mass spectrometry (TMT-MS)-based quantitative proteomics. After data processing, MONET M1 was used to cluster the samples using network modularity optimization while minimizing the percentage of samples not assigned to the network. The resultant classes were then related to case metadata. Three distinct classes of brain proteome profiles were resolved. The first class was enriched with cognitively normal cases with modest cortical amyloid and tau pathological burden, like cases typically defined as controls. The second class was enriched with cases having both high amyloid and tau burden and significant global cognitive dysfunction prior to death, consistent with an AD classification. Interestingly, cases within the third class showed changes in episodic memory performance without a significant difference in amyloid and tau burden compared to the "control" class, suggesting that this group may represent an intermediate stage in disease progression, or be a phenotype distinct from AD. By integrating the results of the unbiased classification with a recent consensus AD brain proteomic network, we found that this third class had distinct patterns of protein expression compared to the AD class. These findings highlight the utility of unbiased network analysis to uncover classes and subtypes of the aging brain in postmortem tissue.

Contact: Kathleen Carter, emma.kathleen.carter@emory.edu

P07.04 Breaking privacy in large scale proteomics data

Russell Bowler (1), Andrew Hill (1), Ani Manichaikul (2), Elizabeth Litkowski (3), Farnoush Banaei-kashani (4)
(1) National Jewish Health, Denver, Colorado, United States
(2) University of Virgina, Charlottesville, Virginia, United States
(3) University of Colorado, Aurora, Colorado, United States
(4) University of Colorado at Denver, Denver, Colorado, United States

Introduction Privacy protection is a core principle of genomic research, but has not been addressed in high-throughput proteomic platforms. Methods Using 1,184 subjects from COPDGene and 1,500 subjects from JHS who had 1,305 proteins

measured by SomaScan, we identified 100 proteins which had independent single nucleotide polymorphism (SNP) quantitative trait loci (pQTL) and then calculated genotype probability functions for each protein-genotype combination (training dataset). To test privacy protection (i.e., assess the ability to identify subjects based on their proteome profile), we applied a naïve Bayes framework to match the protein profile to the genotype profile. Accuracy of the algorithm was tested using 547 independent protein profiles from 9,970 genotyped subjects in COPDGene, 258 independent protein profiles from 2,638 genotyped subjects in SPIROMICS, 1,000 independent profiles from 5,000 genotyped subjects in JHS, and 980 subjects from MESA. Results For SPIROMICS, the probabilities of a correct exact subject match (1 in 2,638) was 86%, a near subject match (top 3 in 2,638) was 93% and a close subject match (top 1%) was 96%. In COPDGene, the probabilities for a correct exact subject match (1 in 9,970) was 83%, a near subject match (top 3 in 9,970) was 87%, and a close subject match (top 1%) was 94%. Accuracy was lower in Blacks compared to other ancestries. When newer data sets with 4,776 proteins were used, exact subject matching increased to 98%. Reidentification accuracy plummeted and approached that of a random guess when we adjusted out the mean genotype effect of proteins. Conclusion This study demonstrates that proteomic data cannot be considered deidentified and suggests that large scale proteomic data used in proteomic research should be given the same privacy protections as genomic data. Alternatively, bioinformatic transformations could be applied to protect privacy (such as adjustment for genotype effect).

Contact: Russell Bowler, bowlerr@njhealth.org

P07.05 CHalf: A User-friendly, Publicly Available Computational Tool for Protein Folding Structure Quantification

Chad D. Hyer (1), Connor Haderlie (1), Monica Berg (1), Hsien-Jung Lavender Lin (1), Isabella James (1), Tyler Bateman (1), Michael Zackrison (1), S Anisha Daley (2), John C Price (1)
(1) Brigham Young University, Provo, Utah, United States
(2) Brigham Young University, Provo, Utah

Traditionally, protein structure has been described using methods such as X-ray crystallography and NMR spectroscopy. While these methods have greatly expanded our understanding of protein structure, they are inherently flawed as they do not accurately describe proteins in biological conditions and are only able to describe a small fraction of the proteome. We have developed an assay that examines in vivo protein stability through mass spectrometry and a computational tool known as CHalf that can calculate the chemical denature midpoints (a term we have dubbed CHalf values) of residues throughout proteins across the proteome and is the first user-friendly, publicly available program of its kind. This assay and CHalf can help remediate flaws in existing protein structure calculation methods by providing a quantitative measure of protein folding strength in biological conditions across the entire proteome. Using findings from this assay and CHalf, we plan to use agnostic machine learning methods to examine possible correlations between CHalf values and current literature solvent-accessible surface area, Rosetta energy, and b-factor values as well as qualitative observations in protein structure. Using these methods, we hope to be able to accurately describe protein folding quantitatively and provide insight into implications of structural motifs on folding strength and vice versa. These aims, if realized, have potential to greatly expand the framework of computational protein structure determination and extrapolation and to provide insight into the effects of protein structure motifs on folding strength and function. Applying these findings in biological conditions, we expect to be able to detect changes in folding strength across conditions to identify changes in protein folding as a result of proteomic diseases. This application shows promise in increasing our understanding of folding patterns of diseased protein biomarkers and is useful in understanding the biochemical pathways of disease and the effects of drug treatments.

Contact: Chad Hyer, hyer.chad@gmail.com

P07.06 Progress Identifying and Analyzing the Human Proteome: 2021 Metrics from the HUPO Human Proteome Project

Gilbert S. Omenn (1), Lydie Lane (2), Christopher M. Overall (3), Young-Ki Paik (4), Ileana M. Cristea (5), Fernando J. Corrales (6), Cecilia Lindskog (7), Susan Weintraub (8), Michael H.A. Roehrl (9), Siqi Liu (10), Nuno Bandeira (11), Sudhir Srivastava (12), Yu-Ju Chen (13), Ruedi Aebersold (14), Robert L. Moritz (15), Eric W. Deutsch (15)
(1) University of Michigan, Institute for Systems Biology, Ann Arbor, Michigan, United States
(2) CALIPHO Group, SIB Swiss Institute of Bioinformatics, Geneva, Switzerland
(3) University of British Columbia, Vancouver, British Columbia, Canada
(4) Yonsei Proteome Research Center and Yonsei University, Seoul, Korea, Republic of
(5) Princeton University, Princeton, New Jersey, United States
(6) Centro Nacional de Biotecnologia, Madrid, Spain
(7) Uppsala University, Uppsala, Sweden
(8) University of Texas Health San Antonio, San Antonio, Texas
(9) Memorial Sloan Kettering Cancer Center, New York, New York, United States
(10) BGI Group, Shenzhen
(11) University of California, San Diego, San Diego, California, United States
(12) National Cancer Institute, Bethesda, Maryland, United States
(13) Institute of Chemistry, Academia Sinica, Taipei, Taiwan (Greater China)
(14) ETH-Zurich and University of Zurich, Zurich, Switzerland
(15) Institute for Systems Biology, Seattle, Washington, United States

The 2021 Metrics of the HUPO Human Proteome Project (HPP) show that protein expression has now been credibly detected (neXtProt PE1 level) for 18,357 (92.8%) of the 19,778 predicted proteins coded in the human genome, a gain of 483 since 2020 from reports throughout the world reanalyzed by the HPP. Conversely, the number of neXtProt PE2, PE3, and PE4 missing proteins has been reduced by 478 to 1421. This represents remarkable progress on the proteome parts list. The utilization of proteomics in a broad array of biological and clinical studies likewise continues to expand with many important findings and effective integration with other omics platforms. We present highlights from the Immunopeptidomics, Glycoproteomics, Infectious Disease, Cardiovascular, Musculo-Skeletal, Liver, and Cancers B/D-HPP teams and from the Knowledgebase, Mass Spectrometry, Antibody Profiling, and Pathology resource pillars, as well as ethical considerations important to the clinical utilization of proteomics and protein biomarkers.

Contact: Gilbert Omenn, gomenn@umich.edu

P07.07 Deep Learning Algorithm Using Collision Cross-Section (CCS) for DIA data analysis

Robin Park (1), Patrick Garrett (2), Tharan Srikumar (3), Sven Brehmer (4), Christopher Adams (5), Ignacio Jauregui (6), Jens Decker (4), Dennis Trede (7), John Yates (2), Rohan Thakur (8)
(1) Bruker, San Diego, California, United States
(2) Scripps Research, San Diego, California, United States
(3) Bruker, Toronto, Canada
(4) Bruker, Bremen, Germany
(5) Bruker, San Jose, California, United States
(6) Bruker, Madrid, Spain

(7) Bruker, Bremen
(8) Bruker, San Jose, California

Since Venable et al. first introduced data-independent acquisition (DIA) in 2004, DIA acquisition and data analysis tools have been continuously improved, making DIA a vital technology to identify and quantify thousands of proteins with high reproducibility and deep proteomics coverage. The advantage of DIA strategy is to cover all ions in a sample by a pre-defined set of precursor isolation windows over the entire m/z range. Meanwhile, a recent breakthrough of 4D proteomics by the timsTOF Pro adds an extra dimension of separation, providing a solution to accurately clarifying a significant number of spectra with CCS (collisional cross-section) values. We developed algorithms to build features using CCS values and incorporate them into deep learning to improve peptide identification. The features include an ion mobility-based search score called TIMScore, p-values derived from the CCS distribution based on peptide candidates. To generate predicted CCS values, we used the training data from the PRIDE repository (PXD019086) and developed a model using transformer, neural network architecture. To evaluate the algorithms, we acquired the spectra on a timsTOF PRO from 400ng of digested Hela proteins using a 30-minute gradient. We ran CCS-aware deep learning algorithm together with the CCS-prediction model on the DIA data on the triplicates of the samples. We plugged the algorithm into DIA-NN open-source version and compared it to the original version. We identified 7941 protein groups, which is 20 percent more than the original version. We also identified 43,783 stripped peptides, which is 40 percent more. CCS-aware deep learning algorithm showed better classification for identifying both proteins and peptides.

Contact: Robin Park, robin.park@bruker.com

P07.08 CHIMERYS: an AI-driven leap forward in peptide identification

Aaron Robitaille (5), Martin Frejno (1), Daniel Zolg (1), Tobias Schmidt (1), Siegfried Gessulat (1), Michael Graber (1), Florian Seefried (1), Magnus Rathke-Kuhnert (1), Samia Ben Fredj (2), Kai Fritzemeier (3), Frank Berg (4), Waqas Nasir (3), Bernard Delanghe (3), Christoph Henrich (3), David Horn (5), Bernhard Kuster (6), Mathias Wilhelm (6)
(1) MSAID, Munich, Germany
(2) MSAID, Munich
(3) Thermo Fisher Scientific, Bremen, Germany
(4) Thermo Fisher Scientific, Bremen
(5) Thermo Fisher Scientific, San Jose, California, United States
(6) Technical University of Munich, Munich, Germany

Matching peptide sequences to tandem mass spectra is integral to bottom-up proteomics, where chimeric spectra are estimated to constitute >50% of data-dependent acquisition data. Some search engines allow multi-pass searches or account for multiple possible precursors, however such approaches result in an over or underutilization of measured fragment ions. This introduces errors or leaves valuable information unused, resulting in far fewer peptide identifications than could be obtained in the data. Here, we describe Proteome Discoverer 3.0 software with CHIMERYS, an intelligent search algorithm, that rethinks the analysis of tandem mass spectra. This innovative approach routinely doubles the number of peptide identifications and reaches identification rates of >80% for typical proteomic data sets. CHIMERYS uses accurate predictions of peptide fragment ion intensities and retention times provided by the deep learning framework INFERYS. CHIMERYS aims to explain as much measured intensity with as few candidate peptides as possible, resulting in the deconvolution of chimeric spectra. We validated this approach in multiple ways including entrapment searches and an in-silico chimeric spectra system. Here, analyzing a HeLa cell lysate digest with CHIMERYS identified 114k PSMs, 61k unique

peptides and 7,300 unique protein groups at 1% FDR. This is a 3.5-, 2- and 1.5-fold increase when compared to Sequest HT, respectively, resulting on average in 2.5-fold more identified peptides per protein. CHIMERYS is compatible with all Orbitrap mass spectrometers but provides more additional identifications from the increased sensitivity of recent instruments. CHIMERYS provides exceptional performance with short chromatographic gradients and high protein loads, enabling higher throughput and a deeper mining of data. CHIMERYS is the first search algorithm that embraces chimeric spectra in a highly scalable, cloud-native, AI-powered implementation.

Contact: Aaron Robitaille, aaron.robitaille@thermofisher.com

P07.09 MetaNetwork: Systems Biology Insights through Quantitative Proteomics and Weighted Correlation Network Analysis

Austin Carr (1), Brian Frey (1), Mark Scalf (1), Anthony Cesnik (1), Zach Rolfs (1), Kyndal Pike (1), Bing Yang (2), David Jarrard (3), Mark Keller (4), Lloyd Smith (5)
(1) Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States
(2) Department of Urology, University of Wisconsin-Madison, Madison, Wisconsin, United States
(3) Department of Urology, University of Wisconsin-Madison, Madison, Wisconsin
(4) Department of Biochemistry, University of Wisconsin-Madison, Madson, Wisconsin, United States
(5) Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin

Proteomics experiments can enable deep insight into biological differences between sample groups. Differential expression analysis is commonly used to identify the subset of proteins with the most statistically significant and interesting fold-changes. By contrast, weighted correlation network analysis (WCNA), uses all measured gene product abundances in an unsupervised clustering and network analysis with a proven capacity for revealing important networks and pathways. However, the use of WCNA in proteomics has been limited due to difficulties in implementing WCNA workflows. Consequently, we developed MetaNetwork, which includes a customizable workflow, visualization of outputs, and a built-in functional annotation enrichment tool. We use MetaNetwork to identify changes in the biological pathways that may enable the differentiation of indolent from aggressive prostate tumors. Six aggressive and six indolent radical prostatectomy specimens sectioned into tumor or histologically “normal” tissue. Two-dimensional LC/MS-MS was performed on trypsinized peptides derived from each sample. The data was searched with MetaMorpheus version 0.0.311. WCNA was performed on the protein intensity values using MetaNetwork v1.0 with default settings. From the 2441 proteins uploaded to MetaNetwork, the WCNA workflow identified ten modules. We found that the networks identified by MetaNetwork’s unsupervised clustering are highly enriched in protein-protein interactions. Next, we used module eigenprotein analysis to identify two statistically significant up-regulated and one down-regulated modules. These significant changes only occurred between the aggressive tumor samples and their paired histologically normal tissue. Using MetaNetwork’s built-in functional annotation enrichment analysis tool, we identified potential biological roles for modules with statistically significant changes: the down-regulated module is enriched in terms related to the cytoskeleton and the extra-cellular matrix, and the two up-regulated modules are enriched in terms related to mRNA translation, RNA metabolism, and protein trafficking. MetaNetwork is also capable of generating many additional layers of information to facilitate the interpretation of differential expression analysis.

Contact: Michael Shortreed, mshort@chem.wisc.edu

P07.10 Estimating absolute protein abundance from CPTAC colon cancer label-free and tandem mass tag (TMT) global proteomics data

Jixin Wang (1), Rachel D'Anna (1), Wen Yu (1), Wenyan Zhong (1)
(1) AstraZeneca, Gaithersburg, Maryland, United States

Purpose: Clinical Proteomic Tumor Analysis Consortium (CPTAC) represents the largest proteogenomic data and is a valuable resource for enhancing our understanding of cancer biology. Proteomics data from CPTAC was primarily generated using TMT multiplex approach, which was designed to provide relative protein quantification between samples. We aim to develop a computational approach to estimate absolute protein abundance from CPTAC TMT global proteomics data. Methods: Label-free proteomics raw intensity data (100 tumors) of colon cancer from CPTAC were processed using MaxQuant. TMT-10 plex mzML files (97 tumor & 100 NAT) were downloaded from CPTAC Data Portal and processed through FragPipe to generate ratio and abundance (ion abundance) based protein quantification. Label-free raw intensity and TMT abundance data were analyzed by total protein approach (TPA, two methods) and intensity based absolute quantification (iBAQ) to derive estimated absolute protein abundance. Results: Using CPTAC colon cancer dataset, we found that estimated absolute protein abundance derived from label-free proteomics data using TPA (two methods) and iBAQ are highly correlated (TPA1 vs TPA2: r = 0.90; TPA1 vs iBAQ: r = 0.85; and TPA2 vs iBAQ: r = 0.95). Further, our analysis revealed high concordance between TMT ratio data generated by FragPipe and published value (Vasaikar et al. 2019) (r = 0.86). Lastly, we showed that estimated absolute protein abundance derived from TMT and label-free using TPA displayed relatively good correlation (TPA1: r = 0.72; TPA2: r = 0.76). Conclusions: We observed relatively good correlation of estimated absolute protein abundance derived from both label-free and TMT global proteomics data of colon cancer in CPTAC, suggesting TPA and iBAQ could be applied to TMT global proteomics data to derive estimated absolute protein abundance. Reference: Vasaikar et al. 2019, Cell, 177, 1035-1049

Contact: Jixin Wang, jixin.wang@astrazeneca.com

P07.11 Identifying Drivers of Human Therapeutic Antibody Production in Chinese Hamster Ovary Cells via Proteomics and Bioinformatics

Junmin Wang (1), Raghothama Chaerkady (1), Lina Chakrabarti (1), Shao Huan Samuel Weng (1), Jie Zhu (1), Diane Hatton (1), Lisa Cazares (1), Sonja Hess (2)
(1) AstraZeneca, Gaithersburg, Maryland, United States
(2) AstraZeneca, Gaithersburg, Maryland

The Chinese hamster ovary (CHO) cell is the primary workhorse for human therapeutic antibody production. Chakrabarti et al. (2019) showed that mitochondrial membrane potential (MMP) identifies CHO cells with high recombinant protein productivity. Nevertheless, our ability to engineer high-yielding cell lines has been limited by insufficient knowledge of proteins driving antibody production in CHO cells. Poor characterization of protein functions and limited sample sizes constitute major challenges for mass spectrometry-based studies of the Chinese hamster proteome. To identify drivers of antibody production, we performed an LC-MS/MS proteomic analysis on a Q-Exactive HF-X mass spectrometer in CHO cells with normal and high MMP. The MaxQuant protein and peptide output was used for subsequent analysis. To improve protein annotations, we mapped CHO proteins to murine homologs via pairwise global sequence alignment and a logistic regression model with sequence percentage identity as its feature. Subsequently, we repurposed CAMERA, a gene set test developed by Wu and Smyth (2012), for differential protein expression analysis. Our approach permits analysis to be carried out on an ensemble of peptides originating from the same protein, facilitating the validation of subtle protein changes. Our

alignment-based regression model enables us to map 99% of proteins in the Chinese hamster proteome to the corresponding murine homologs. To validate our CAMERA-based method, we compared its performance with LIMMA. Results suggested that our method outperformed LIMMA for proteins associated with numerous peptide identifications. Using peptide-level data increased the sample size, leading to improved statistical power. Altogether our method identified 2.8 times more proteins that were significantly changed than LIMMA. These proteins, including pantetheinase and mitochondrial glycerol-3-phosphate dehydrogenase, defined an altered metabolic landscape of high-MMP CHO cells. Our work facilitates a deeper understanding of CHO cell biology. It will enable us to engineer high-producing CHO cell lines.

Contact: Junmin Wang, junmin.wang@astrazeneca.com

P07.13 Increasing Human Leukocyte Antigen Peptide Identifications Using Fragpipe and MSBooster

Kevin Yang (1), Fengchao Yu (1), Guo Ci Teo (1), Vadim Demichev (2), Markus Ralser (2), Alexey Nesvizhskii (1)
(1) University of Michigan, Ann Arbor, Michigan, United States
(2) Charite Universitatsmedizin, Berlin, Germany

Peptide antigen presentation via the major histocompatibility complex (MHC) is a crucial process by which the immune system can detect and target bacterial and viral infections as well as cancers. The human leukocyte antigen (HLA) peptidome presents a unique challenge for database searching of LC-MS/MS data, as the number of possible sequence identities far exceeds those possible in standard enzymatic digests such as tryptic peptide databases. Previous software solutions have used deep learning (DL) models to rescore PSMs using features describing the consistency of predicted MS2 spectra and retention time with experimentally observed values. Here, we present a workflow in Fragpipe with Percolator for PSM rescoring of HLA-I peptides. Incorporation of DL features is available through MSBooster, a new addition to Fragpipe. Using reversed-phase fractionated samples from an A∗02:01 monoallelic cell line (Klaeger et al, 2021), rescoring with DL features improved peptide identifications by more than 25% compared to a baseline model without DL rescoring. Another difficulty of searching HLA data is that they fragment differently from tryptic peptides due to their non-enzymatic nature. Some HLA peptides display strong internal fragment ion series, which are not typically handled by DL models predicting b- and y-ions. We perform a comparison of spectral similarity features using either 1) a baseline model trained on tryptic peptides, 2) the Prosit HLA model, or 3) the PredFull full spectrum prediction model. This is the first application of testing full spectrum prediction for HLA peptide rescoring, rather than using just b- and y-ions.

Contact: Kevin Yang, yangkl@umich.edu

P07.14 One-stop DIA data analysis using MSFragger-DIA coupled with FragPipe

Fengchao Yu (1), Kevin Yang (1), Vadim Demichev (2), Sarah Haynes (1), Markus Ralser (2), Alexey Nesvizhskii (1)
(1) University of Michigan, Ann Arbor, Michigan, United States
(2) Charité - Universitätsmedizin Berlin, Berlin, Germany

Data-independent acquisition (DIA) is a widely used data acquisition approach in bottom-up proteomics. Most DIA peptide identification methods require extraction of precursor and fragment ions, either by detecting features or using spectral libraries, prior to database searching. Here, we propose an approach to directly search DIA spectra against a database using MSFragger-DIA, a new module in MSFragger. We can identify peptides directly from DIA data and then build a spectral library for quantification. To increase the sensitivity of peptide identification, we have also developed MSBooster, a

rescoring tool that leverages deep learning-predicted spectra. These features are now available in FragPipe, a proteomics data analysis suite, enabling straightforward library-based and library-free analyses. Furthermore, with MSFragger DDA and DIA modes, we can build a hybrid library using both DDA and DIA data to boost the number of quantified peptides. We first used six gas-phase fractionated runs and four single-shot runs from Searle et al. 2020 to demonstrate the performance of MSFragger-DIA. We showed that MSFragger-DIA coupled with FragPipe can detect 30% more peptides compared to the one reported in the original publication. To evaluate the false-discovery rate (FDR), we generated an entrapment database containing the same number of target proteins and false proteins. Both target and false proteins were used to generate decoy proteins for FDR estimation. Re-analyzing the same dataset using the entrapment database showed that the actual FDR was much lower than 1%. To test the performance of MSFragger-DIA in analyzing phosphorylation-enriched DIA data, we used two datasets from Bekker-Jensen et al. 2020 and Gao et al. 2021. We compared these results to those from other tools including DIA-Umpire, Spectronaut, and DIA-NN, and show that MSFragger-DIA has the highest sensitivity. Finally, we performed two runtime comparisons to show that MSFragger-DIA was fast in analyzing proteome and phospho-enriched datasets.

Contact: Fengchao Yu, yufe@umich.edu

P07.15 moPepGen: a fast custom database generator from multi-omics data for proteogenomics

Chenghao Zhu (1), Lydia Y. Liu (2), Paul C. Boutros (1)
(1) Department of Human Genetics, University of California, Los Angeles, Los Angeles, California, United States
(2) Department of Medical Biophysics, University of Toronto, Toronto, Canada

Purpose: Cancers are driven by genomic variants such as SNV (single nucleotide variants) and INDEL, often accompanied by many transcriptional variants. Modern mass spectrometry based proteomics is able to identify and quantify peptides and proteins comprehensively, however the variant-harboring peptides that are absent in canonical databases are largely under-studied. A common problem in custom database construction is the large number of combinations of variants to consider. Thus major proteogenomic studies often choose a strategy that only captures peptides harboring a single variant. Existing algorithms also suffer from limited sources of variation. Methods: Here we present moPepGen (multi-omics peptide database generator) that aims at accelerating proteogenomic researching by generating custom peptide database from variety of genomic and transcriptional variants. MoPepGen uses a graph-based algorithm that achieves a linear time complexity in integrating variants in contrast to the exponential complexity of exhaust searching. MoPepGen integrates variants from a variety of sources including SNV, indel, transcription fusion, alternative splicing, RNA editing, and circRNA. It is able to generate peptides harboring any combinations of variants on the single peptide. MoPepGen is also highly extensible for other types of genomic and transcriptional variants. Results: With moPepGen we created sample-specific database noncanonical peptide harboring SNV, transcriptional fusion and RNA editing events from ~1000 cancer cell lines. The run time for generating custom database is around 40 minutes per sample. We then applied a tiered library searching strategy using three independent search engines and identified tens to thousands of variant-harboring peptides across 36 cancer types. Conclusion: This study demonstrated the capacity of moPepGen in identifying noncanonical peptides in proteogenomics studies and extended the potentiality of proteomics in precision medicine and biomarker discovery.

Contact: Chenghao Zhu, ChenghaoZhu@mednet.ucla.edu

P08: Data Analysis and Visualization

P08.01 Automated workflows for DIA data using DIA-NN on the PaSER platform

Christopher Adams (1), Robin Park (1), Sven Brehmer (2), Nagarjuna Nagaraj (2), Tharan Srikumar (1)
(1) Bruker Daltonics, San Jose, California, United States
(2) Bruker Daltonics, Bremen, Germany

Abstract: Data independent acquisition has become the go to method for deep and quantitative proteomic analysis given the ability to sample large m/z windows in a reproducible and non-stochastic manner. Using a method termed dia-PASEF on a TIMS enabled Q-TOF lends additional advantages in both duty cycle and selectivity using the ion mobility space. Dia-PASEF allows for deep proteomes in short gradient times (<20 min.) and therefore 100's of LCMS runs can be generated in short times. Data analysis in a streamlined automated fashion expedites the time from experiments to discovery. DIA-NN is a novel software package that uses neural networks providing best-in-class DIA output. Here we integrate DIA-NN onto the PaSER platform for a streamlined workflow for the analysis of many samples in a short analysis time with no file transfer or data migration. Methods: A timsTOF Pro using variable CE's and mobility windows in gradients ranging from 5 min. to 90 min. were used. DIA-NN was modified to become CCS-enabled and process data in the most expedient fashion. The PaSER GUI was designed such that first-pass analysis is predefined automatically triggering quantitative analysis. Retrospectively, match-between-runs (MBR) analysis can be triggered on the whole project or subset of user defined experiments. Preliminary Data: Human, Yeast and E. coli (HYE) digested mixtures at different but known ratios with injection loads from 50ng to 600ng were run at different gradient lengths in replicate resulting in >2500 proteins at short gradients to >9000 proteins identified and quantified at longer gradients. Quantitative accuracy was shown to be <20% CV. Using the DIA-NN as integrated into PaSER creates a seamless approach to dia-PASEF analysis. Novel Aspect: Automated workflows for dia-PASEF using DIA-NN on PaSER streamlines experiments to results >

Contact: Chris Adams, Christopher.Adams@bruker.com

P08.03 A Cloud-scalable Software Suite for Large-Scale Proteogenomics Data Analysis and Visualization

Margaret Donovan (1), Harsharn Auluck (1), Arjun Vadapalli (1), Yan Berk (1), Aaron Gajadhar (1), Khatereh Motamedchaboki (1), Yuandan Lou (1), Theo Platt (1), Asim Siddiqui (1)
(1) Seer, Redwood City, California, United States

Assessment of the flow of genetic information through multi-'omic data integration can reveal the molecular consequences of genetic variation underlying human disease. Next-generation sequencing (NGS) can be used to identify genetic variants, while mass spectrometry can be used to assess the proteome. The ProteographTM Product Suite[1], leverages multiple nanoparticles with distinct physiochemical properties to enable large-scale, deep plasma proteome analyses. Integration of proteomics and genomics data requires many tools of which require complex workflows that can act as a barrier for researchers to adapt new analysis tools. Here, we present a cloud-based, data analysis software platform called Proteograph Analysis Suite (PAS) for proteogenomic data analyses through the integration of Proteograph proteomics data with NGS variant information. PAS includes an experiment data management system, analysis protocols, analysis setup wizard, and result visualizations. PAS can support both Data Independent Analysis (DIA) and Data Dependent Analysis (DDA) workflows and is compatible with variant call format (.vcf) files, enabling personalized database

searches. To assess data quality, PAS includes metrics for identified peptides and protein groups like intensity, protein sequence coverage, abundance distributions, and counts. Visualizations, including principal component analysis, hierarchical clustering, and heatmaps, allowing identification of experimental trends. To enable biological insights, differential abundance analyses results are displayed as volcano plots, protein interaction maps, and protein-set enrichment. From data to insight, PAS provides an easy-to-use and efficient suite of tools to enable proteogenomic data analysis. Here, we apply PAS by analyzing 141 Proteograph NSCLC plasma dataset[1] and performing a database search. This search was launched through the user interface requiring only three clicks, provisioned 142 servers, and completed in approximately five and half hours. Together, these results show the utility of PAS for seamless and fast proteomic data analysis. Reference: [1]Blume, J. E. et al. Nature Communications, 2020

Contact: Margaret Donovan, mdonovan@seer.bio

P08.04 User-Friendly, Versatile Proteomics Analysis with FragPipe

Sarah Haynes (1), Dmitry Avtonomov (1), Fengchao Yu (1), Guo Ci Teo (1), Felipe Leprevost (1), Daniel Geiszler (1), Daniel Polasky (1), Kevin Yang (2), Alexey Nesvizhskii (1)
(1) University of Michigan, Ann Arbor, Michigan, United States
(2) University of Michigan, Ann Arbor, Michigan

FragPipe is a complete proteomics pipeline for a variety of workflows and experiment types. It can be used on both Linux and Windows systems, where users can choose from an array of complete analysis workflows to easily load and run without the need to set options manually. Typical analyses with FragPipe involve only a few steps: load a workflow, add LC-MS/MS spectral files, and click "Run". FragPipe is a graphical interface that brings together a number of tools which can be used in complete workflows or as individual data processing steps. Fast database searching can be performed with MSFragger. Direct analysis of data-independent acquisition (DIA) data can be done with MSFragger-DIA or DIAUmpire combined with DIA-NN quantification. Crystal-C can be used for post-search of mass shift artifacts from open search. Peptide validation can be performed with either PeptideProphet or Percolator, with protein inference by ProteinProphet. Philosopher can perform isobaric labeling-based quantification and generate multi-level reports. PTM-Shepherd provides summarization and characterization of possible post-translational modifications (PTMs) found from open searches. Label-free quantification (with optional FDR-controlled match-between-runs) can be accomplished with IonQuant, and TMT or iTRAQ data can be analyzed with TMT-Integrator. Available workflows include traditional (closed) search with label-free quantification; open (mass tolerant) or mass offset searches for PTM discovery; non-specific (HLA or peptidome) searches; isobaric quantification (TMT or iTRAQ) at the gene, protein, peptide, and PTM site levels; isotopic labeling (SILAC, dimethyl, custom MS1-based labels); glycopeptide (N- and O-linked; and DIA search and quantification. FragPipe can be used with a variety of LC-MS/MS data types, including raw files directly from Orbitrap and timsTOF instruments as well as spectral files converted to the mzML format.

Contact: Sarah Haynes, hayse@umich.edu

P08.05 TIMS Viz for Mobility Offset Mass Aligned interrogation of complex samples

Tharan Srikumar (3), Philipp Strohmidel (1), Sebastian Wehner (1), Jens Decker (1), Ignacio Auregui (1), Christopher Adams (2), Sven Brehmer (1)
(1) Bruker Daltonics GmbH & Co. KG, Bremen, Germany

(2) Bruker Scientific LLC, Billerica, United States
(3) Bruker Ltd, Milton, Ontario, Canada

The PASEF® acquisition mode of the timsTOF Pro has the power to isolate co-eluting, quasi-isobaric peptides separately for fragmentation, based on differences in the peptide's ion mobility. Such an event is called Mobility Offset Mass Aligned (MOMA) and results in non-chimeric spectra, despite a quadrupoles fidelity. This is especially valuable for PTM analysis, for example to resolve positional isomers of phosphopeptides. TIMS Viz was introduced in PaSER to visualize MOMA events in complex samples. During acquisition, MS/MS spectra are streamed to a GPU-powered processing computer performing a real-time database search, called PaSER. The database search utilizes all four dimensions – retention time, CCS value, m/z and fragment spectra – to increase confidence in the identification results. TIMS Viz, a novel data visualization tool to display an interactive heatmap in the m/z ion mobility space, maps MOMA features. Herein, we show the number of MOMA groups, which are sets of at least two MOMA features, that could be identified by TIMS Viz in two different data sets with different m/z tolerance settings. Setting tolerances to 500 mDa (typical lower limits of quadrupole isolation) and a retention time window of 10 s resulted in more than 40,000 MOMA groups containing more than 90,000 spectra for both, the cell lysate sample and the phosphopeptide enriched sample. Without the power of ion mobility separation these spectra would likely be chimeric in nature. Lowering the m/z tolerance to 25 mDa (well below the tolerance of any quadrupole) still leads to more than 18,000 MOMA groups (> 40,000 spectra) for the cell lysate and more than 23,000 MOMA groups (> 52,000 spectra) for the phosphopeptide enriched sample. TIMS Viz helps user to explore their data for MOMA features and is a powerful demonstration how the TIMS dimension can improve the spectral quality for co-eluting, quasi-isobaric peptides.

Contact: Tharan Srikumar, tharan.srikumar@bruker.com

P08.06 ASSESSMENT OF DISCORDANCE BETWEEN PEPTIDE AND PROTEIN LEVEL DIFFERENTIAL EXPRESSION ANALYSIS RESULTS IN CANCER BIOMARKER DISCOVERY

Ginny Xiaohe LI (1), Alexey Nesvizhskii (1), Marcin Cieslik (1), Michael Mumphrey (1)
(1) University of Michigan, Ann Arbor, Michigan, United States

Nowadays, molecular biomarker discovery relies highly on mass spectrometry proteomics. In bottom-up proteome inference and quantitation, a single abundance measurement on proteins are routinely used in comparing between subgroups of patients. Such measurement on protein is obtained by aggregating intensities of peptides mapped. There are two major sources of inaccuracy that might weaken the reliability of protein level quantification. First, a given tryptic peptide can be derived from multiple different proteoforms. Assigning it to any of the proteoforms is a simplification of reality without full confidence. Second, the assumption underlying the aggregation that multiple peptides mapping to a gene behave in the same direction is often violated. Hence, protein level aggregation of multiple peptides may average out true biological effect. In this manuscript, we study discordant discoveries made when contrasting abundance measurements of tumor with normal adjacent tissue (NAT) samples in protein-level as opposed to peptide-level based on data in major cancer cohorts from the Clinical Proteomics Tumor Analysis Consortium (CPTAC). To start with, we investigate the discordances seen in the highly polymorphic HLA proteins. As an example, in 212 Lung Squamous Cell Carcinoma (LSCC) samples, 31 peptides are assigned to HLA-A protein (ENSP00000366002.5), among these peptides, 4 and 2 peptides are found to be significantly enriched and depleted in tumor samples compared to NAT samples respectively. We then calculate protein level abundance measurement only from peptides that are uniquely mapped to the protein, and compare with conventional protein level quantification. In addition, we provide visualization of the peptide-proteomform mapping topology together with projection of conserved protein domains and known PTMs. In sum, this work explores the complexity in bottom-up proteomics presented

by discordances in peptide and protein level differential expression analysis results and provides insight into improved peptide-proteoform assignment.

Contact: Ginny Xiaohe Li, lixiaohe@umich.edu

P08.07 Absolute Quantification (AQUA) of Vitellogenin Proteins in Striped Bass Ovary: Considering the Traditional PRM and the Nontraditional DDA Workflows

Taufika Islam Williams (1), Cara Kowalchyk (2), Leonard Collins (1), Benjamin Reading (1)
(1) NCSU, Raleigh, North Carolina, United States
(2) NC Division of Marine Fisheries, Raleigh, North Carolina, United States

Mass spectrometry has steadily but surely moved into the forefront of quantification-centered protein research. The AQUA method of Protein Cleavage Isotope Dilution Mass Spectrometry (PC-IDMS) is a proven way for quantifying proteins by using an isotope-labeled analogue of a peptide fragment of the parent protein as an internal standard. Such targeted analyses are widely performed using triple-quadrupole-MS-based Multiple Reaction Monitoring (MRM) experiments and orbitrap-based Parallel Reaction Monitoring (PRM) workflows. PRM has lately been the go-to approach for such quantification, as it is believed to offer highest instrument sensitivity in an orbitrap-based instrument, just as MRM offers lowest limits of detection (LOD) in a triple-quadrupole-MS. Enhanced sensitivity is considered to be elemental in the accurate and precise delivery of quantitative data. We performed a comparative study of DDA- and PRM-based workflows using the quantification of vitellogenins in striped bass (Morone saxatils) eggs. Striped bass are an important recreational and commercial fish species native to North Carolina, however, natural recruitment is limited in key water systems such as the Tar, Neuse, and Cape Fear Rivers. The study of vitellogenins, which are glycolipophosphoproteins critically important to the striped bass reproductive cycle, can shed some light on adaptations associated with wild type stock diminishing. Vitellogenins are precursors to egg yolk proteins (YPs), and the ratios of key vitellogenins plays an important role in egg quality throughout development and different environmental conditions. We employed both DDA and PRM experiments to quantify VTG Aa, Ab and C in striped bass ovaries. Our results clearly indicate that it may not be necessary to perform a targeted analysis such as PRM, as DDA is able to provide both absolute quantification data and discovery data in a single nanoLC-MS/MS run.

Contact: Taufika Williams, Taufika_Williams@ncsu.edu

P08.08 SimpliFi: Democratizing the Analysis and Accessibility of Multiomics Data

John Wilson (1), Jim Palmeri (1), Mark Pitman (1), Darryl Pappin (2)
(1) ProtiFi LLC, Farmingdale, New York, United States
(2) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States

Purpose Technological advances mean we are becoming increasingly efficient at producing large amounts of omics data; however, our ability to translate this into tangible meaning is falling behind. Large numbers of analytical tools tends to hinder, rather than help; the complexity only getting worse with multi-omics studies. To enable everyone to understand and use omics, we created SimpliFi, the world's first cloud-based, GPU-driven, streamlined and browser-accessible data-to-meaning engine for integrated omics analyses. SimpliFi obligates QC including batch and run order effects and results. Results are easily shared, explored or published simply by sending a URL. Methods SimpliFi models biology, using

exclusively non-parametric statistics, with biological replicates defining their own distributions; p-values and fold-changes are determined as a function of biological variation, number of samples and observations and measurement error. Optimized GPU CUDA routines generate confidence intervals via resampling for all values, including p-values. Crucially, SimpliFi does not transform data and accounts for increased data variance at low or high intensities. SimpliFi's user interface is intuitive and user-friendly. Results SimpliFi makes the following statistical assumptions: (1) Two means are indistinguishable without further knowledge of their scatter = distribution. (2) Biological replicates yield biological empirical distributions which give p-values in comparisons between groups. (3) Bootstrapping handles all distributions and yields confidence intervals for p-values and fold changes. (4) Measurement error is a function of measurement intensity and must be controlled for. (5) Condition p-values represent "the chance the observed difference between conditions could occur due only to intrinsic variability within 'same'." Alternatively, they are "FDRs of different." (6) This computational framework is applicable to all kinds of biological entities (protein, peptide, metabolite, etc.), and allows unified analyses of multi-omics data. Conclusion First cloud-omics analytics engine designed from the ground up, from fundamental first-principles of mathematics, data analysis, visualization and user design.

Contact: John Wilson, john@protifi.com

P09: Glycoproteomics and Glycomics

P09.01 Increasing Glycopeptide Identifications through Targeted Y Ion Detection

Robert Chalkley (1), Peter Baker (1)
(1) UCSF, San Francisco, California, United States

Intact glycopeptide analysis using mass spectrometry has become increasingly effective due to hybrid fragmentation protocols such as EThcD and stepped HCD. These allow the identification of thousands of glycopeptide spectra from complex mixtures. However, within these datasets there are typically spectra of other glycoforms of the same glycopeptides that are not identified due to either poorer quality fragmentation data or unanticipated modifications. Bioinformatic strategies to uncover these extra glycoforms would provide more comprehensive sample characterization. Different glycoforms of the same glycopeptides produce the same Y0/Y1/Y2 ions. We have previously adapted the software MS-Filter to automatically calculate the expected masses of these ions from list of input peptides and report all spectra in peak list files that contain these ions along with tentative glycan assignments. We have now added peptide and glycan scoring to the output, which allows reliability filtering of these results and viewing of annotated spectra for all assignments. We have evaluated the performance of this strategy for analysis of both EThcD and HCD data. Comparing to results from database searching in each case we were able to identify more glycoforms. For many of these spectra the same identification was reported by database searching but was below the 1% FDR threshold applied. A particular advantage of the strategy is the ability to report masses of modifications that do not correspond to anticipated glycan structures. This has allowed detection of unusual glycan modifications such as multiply phosphorylated glycans and the presence of salt and metal adduct spectra.

Contact: Robert Chalkley, chalkley@cgl.ucsf.edu

P09.02 Optimizing Glycopeptide Fragmentation for Characterization of Multiply-Fucosylated Glycoproteins in Head and Neck Squamous Cell Carcinoma

Kevin Chandler (1), Daniel Marrero (1), Robert Sackstein (1)
(1) Translational Glycobiology Institute, Florida International University College of Medicine, Miami, Florida, United States

Analytical workflows relying on higher energy collisional dissociation (HCD) tandem mass spectrometry have evolved into powerful tools for glycopeptide analysis, but topological characterization of multiply-fucosylated glycopeptides continues to challenge the field. Common glycopeptide fragmentation strategies prioritize peptide backbone fragmentation, and rarely consider the value of glycan topological information in glycopeptide tandem mass spectra. We sought to apply stepped collision energy HCD (sceHCD) and electron transfer dissociation with HCD supplemental activation (EThcD) to the analysis of multiply-fucosylated glycopeptides from head and neck squamous cell carcinoma lysates, to produce information-rich glycopeptide spectra with both peptide and glycan topological information and aid in the assignment of fucosylated carbohydrate epitope topology. Fucosylated glycopeptides were enriched via (a) AAL lectin, and (b) anti-Le^{y} antibody from CAL27 and HSC-3 tongue squamous cell carcinoma lysate digests, followed by nLC-MS/MS on an Orbitrap Eclipse Tribrid Mass Spectrometer with an EASY-nLC 1200 system. An Acclaim PepMap 100 trapping column and a PepMap RSLC C18 analytical column were employed for chromatographic separation, and HCD (35% collision energy) followed by oxonium ion triggered sceHCD and EThcD dependent scans were performed. Analyses were performed in triplicate. Spectra were analyzed using Byonic, and assigned spectra were manually verified. The nLC-MS/MS analyses led to the successful assignment of 212 unique high confidence glycopeptides (score ≥ 200) representing 73 glycoproteins, and 52% of the

glycopeptides were fucosylated. sceHCD and EThcD facilitated topological assignment of multiply-fucosylated carbohydrate substructures and contained higher intensity $HexNAc_1Hex_1Fuc_2$ (m/z 512.1974) and $HexNAc_1Hex_1Fuc_2$ (m/z 658.2533) oxonium ion peaks compared to HCD spectra acquired at a single collision energy. Future analyses will focus on additional optimization of sceHCD parameters to further improve the presence and intensity of topologically useful fucosylated glycopeptide fragments in glycopeptide MS2 spectra.

Contact: Kevin Chandler, kchandle@fiu.edu

P09.03 Correlation of Quantitative Analysis of Glycoproteomic Biomarkers from Dried Blood Spots and Serum.

Hector Huang (1), Liam Magee (1), Apoorva Srinivasan (1), Prasanna Ramachandran (1), Darragh Buckley (1), Kaity Moser (1), Ley Hian Low (1), Carrie Smith (2), Xin Cong (1)
(1) InterVenn Biosciences, South San Francisco, California, United States
(2) InterVenn Biosciences, South San Francisco, California

Dried blood spot (DBS) cards have been used as a biosampling method for over a century and were introduced into the clinic fifty years ago. These cards, which collect finger or heel prick blood in infants, provide an effective and affordable alternative to traditional phlebotomy in many different applications including genomics, metabolomics and proteomics. We have recently demonstrated that glycoproteomic analysis of peripheral blood proteins provides a powerful method for ovarian cancer screening. Here we present a method for extracting glycoproteins from DBSs followed by Multiple Reaction Monitoring (MRM) liquid chromatography-mass spectrometry (LC-MS) analysis to quantify glycopeptide abundances. The work presented here provides a quantitative comparison of glycopeptide abundances extracted from DBS and serum. In a pilot study we found that glycopeptides can be extracted from DBS samples in a reproducible and consistent manner, and that relative abundances of individual glycopeptides correlate well with relative abundances determined in serum with a Pearson product-moment correlation coefficient > 0.8. To validate these results, we applied our extraction protocol to DBS samples collected from 29 individuals enrolled in the Clinical Validation of the InterVenn Ovarian CAncer Liquid biopsy clinical trial (ClinicalTrials.gov NCT03837327), which examines the ability to discriminate malignant from benign pelvic tumors, based on their blood glycoproteomic signature. Roughly half of the 29 assessed individuals had benign histologies, while the other half were confirmed to have an advanced form of ovarian cancer. Principal component analysis of glycoproteomic signatures of DBS samples confirmed this distribution. Our results indicate that samples collected on DBS can be used as a source for MRM-based analysis of glycoproteomic biomarkers in peripheral blood, with sufficiently robust performance to be applied for distinguishing between benign and malignant ovarian tumors.

Contact: Hector Huang, hhuang@venn.bio

P09.04 Spatiotemporal glycosylation profile at the maternal-fetal interface

Ke Leow (1), Richard Drake (2), Mike Angelo (1)
(1) Stanford University, Palo Alto, California, United States
(2) Medical University of South Carolina, Charleston, South Carolina, United States

Aberrant glycosylation has been observed in solid tumors for decades, but little is known about the significance of altered glycosylation and its impact on immunomodulation and tumor infiltration. We hypothesize that glycosylation patterns at the maternal-fetal interface are highly reminiscent of that in tumors, and the goal of this study is to determine changes in

glycosylation across gestation and to elucidate roles of glycosylation in pregnancy and in cancer. Given that functions of glycosylation structures are often dependent on their interactions with glycan receptors expressed on cells in the local environment, spatial information is critical to comprehensively determine glycosylation changes and their roles at the maternal-fetal interface. The advent of state-of-the-art imaging technologies such as MALDI glycan imaging and multiplex IHC methods, including multiplexed ion beam imaging by time of flight (MIBI-TOF) developed by our lab, presents tools to interrogate the spatial distribution of glycan structures and glycan receptor expression at the glycomics and proteomics levels respectively. We performed MALDI glycan imaging on a tissue microarray and full tissue sections from a cohort of maternal decidua obtained from women who underwent elective terminations of otherwise healthy pregnancies at 6-20 weeks gestation. We observed temporally- and spatially- dependent changes in decidual glycan patterns, and we correlated these patterns with MIBI-TOF and multiplex IHC data to associate glycosylation and cellular profiles. Our preliminary data and past literature suggest close parallels in glycan profiles between tumor and decidua. For example, highly branched glycans were enriched in the anchoring villi where fetal cells invade into the maternal decidua. In tumors, glycan branching has been positively correlated with cancer progression and invasiveness. Overall, this is the first N-glycan imaging study in the human maternal endometrium, and we are employing computational tools to predict glycan expression and function at the subcellular level.

Contact: Ke Leow, kxleow@stanford.edu

P09.05 Selective enrichment of cell surface proteins using ProMTag to detect O-glycosylation-dependent changes

Amber Lucas (1), Jonathan Minden (1), Collin Bachert (2), Adam Linstedt (2)
(1) Impact Proteomics, Pittsburgh, Pennsylvania, United States
(2) Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

O-linked glycosylation, which is regulated by the enzyme GalNAc transferase, plays a vital role in cellular function by regulating membrane protein composition and sorting. Changes in O-linked glycosylation are very common in diseases such as cancers, diabetes, and Alzheimer's. It is important to study the effects of changes in O-link glycosylation on cell membrane composition and function to understand how these changes affect the onset, phenotype, and progression of these diseases. However, membrane proteins are notoriously challenging to analyze at the proteomic level due to their hydrophobicity and the inability to separate the membrane proteins from other proteins in the cell.

Here we introduce a new sample preparation workflow that enriches membrane proteins for mass spectrometry analysis to understand changes in protein composition of the plasma membrane as the result of changes in O-linked glycosylation. To do this, we used the bifunctional tag ProMTag to label and capture plasma membrane proteins via their extracellular domains in cells treated with GalNAc Transferase inhibitors. One end of the ProMTag forms a reversible, covalent bond with protein the other end of the ProMTag is capable of forming an irreversible, covalent bond with a solid bead support via the click chemistry pair of methyltetrazine and trans-Cyclooctene (TCO). The extracellular domains of proteins from intact wild type HEK cells, HEK cells treated with GalNAc Transferase inhibitors, and HEK cells with GalNAc Transferase mutations were ProMTagged, captured, cleaned up, and eluted. The samples were then TMT labeled and analyzed using mass spectrometry (MS). MS confirmed enrichment of plasma membrane proteins after sample preparation using this new ProMTag workflow. Protein changes in the plasma membrane as a result of GalNAc Transferase inhibition or mutation were characterized. This work establishes a method for enrichment of plasma membrane proteins and reveals how changes in O-linked glycosylation via GalNAc Transferase inhibition leads to changes in plasma membrane protein composition due to altered membrane protein stability and sorting.

Contact: Amber Lucas, amber.lucas@impactproteomics.com

P09.06 Sialic Acid Derivatization and Bioorthogonal Labeling for Multiplexed Mass Spectrometry and Fluorescence Imaging

Colin McDowell (1), Xiaowei Lu (1), Rachel Stubler (1), Luke Wisniewski (2), ChongFeng Gao (2), Peggi Angel (1), Brian Haab (2), Richard Drake (3)
(1) Medical University of South Carolina, Charleston, South Carolina, United States
(2) Van Andel Institute, Grand Rapids, Michigan, United States
(3) Medical University of South Carolina, Charleston, South Carolina

Sialic acid-modified glycans are critical regulators of cell-cell communication through their direct interaction with a host of carbohydrate binding proteins. α2,3- and α2,6-sialic acid linkage isomers have disparate biological functions and are differentially modulated by cancer and other nonmalignant diseases. Imaging mass spectrometry (IMS) is a well-established tool for the analysis of glycosylation, however the inability to resolve isomers by mass makes characterization of sialic acids challenging. Here we present a novel sialic acid derivatization approach for isomeric characterization by mass spectrometry and fluorescence imaging. In preparation for N-glycan imaging mass spectrometry, sialic acid residues on fixed cells and FFPE tissues were differentially modified with alkyne (α2,3) and amide (α2,6) moieties by a novel propargylamine-based amidation-amidation, termed AAXL. AAXL-labeled cell slides were then subsequently integrated into a previously described N-glycan IMS sample preparation workflow[1]. For fluorescence imaging, sialic acids were derivatized with azide (α2,3) and amide (α2,6) groups via a similar amine-azide-based amidation, termed AAN_3. Azide-functionalized α2,3 isomers were further reacted with a biotin-alkyne through click-chemistry, which then captured a streptavidin-Cy5 probe for fluorescent detection. The AAXL synthesis successfully differentiates α2,3- and α2,6-sialylated N-glycan isomers by imaging mass spectrometry analysis. AAXL-derivatization introduces +37.0316 m.u. (α2,3) and +27.0473 m.u. (α2,6) mass shifts which resolve the isomers by mass spectrometry. Via AAXL we detect over 100 sialylated N-glycan species from FFPE tissues and fixed cells, where AAXL revealed distinct sialic acid isomer localization. Using AAN_3 and subsequent biotin labeling of α2,3-linked isomers, we have successfully detected α2,3-sialylated glycans by fluorescence imaging. The AAN_3 fluorescent labeling strategy demonstrated disparate α2,3-sialic acid expression across FFPE tissue regions and between multiple cell lines. This dual bioorthogonal labeling system is a novel tool for multiplexed sialic acid imaging, which is further adaptable to glycomic and glycoproteomic enrichment strategies.

Contact: Colin McDowell, mcdoweco@musc.edu

P09.07 Cross-Ring Fragmentation of N-linked Glycans via Lithium-Doped ESI Using Matrix-Assisted Laser Desorption Electrospray Ionization Mass Spectrometry

Tana Palomino (1), Crystal Pace (1), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States

N-linked glycans play an important role in essential biological processes. They consist of highly complex structures due to their immense diversity in possible isomers and linkages. To characterize glycan structures with high specificity, tandem mass spectrometry is required to induce glycosidic cleavages and cross-ring fragments. While glycosidic cleavages provide sequencing and branching information, cross-ring fragmentation elucidate to specific linkage information between monosaccharides. However, cross-ring fragmentation is not as abundant as glycosidic cleavages and thus, linkage

information is difficult to obtain for N-linked glycans. Previous research has shown that lithium adducts assist in promoting cross-ring fragmentation of glycans as the metal adduct stabilizes the glycosidic bonds. In this study, we use lithium-doped ESI to promote cross-ring fragmentation of enzymatically cleaved N-linked glycans from the glycoprotein, bovine fetuin. We applied lithium doping to our novel ionization source IR-MALDESI where ions are desorbed by IR ablation and post-ionization by lithium doped ESI. Adequate concentrations of lithium ion were added to the ESI solvent and samples will be directly analyzed by IR-MALDESI. Ultimately, these methods provide crucial structural information that greatly benefit and expand the field of glycan mass spectrometry with applications towards glycan imaging of tissues.

Contact: Tana Palomino, tvpalomi@ncsu.edu

P09.08 MS-Based Glycomics and Glycoproteomics facilitating the Identification of Glycan and Glycopeptide Isomers as Reliable Alzheimer's Disease Biomarkers

Yehia Mechref (1)
(1) Texas Tech University, Lubbock, Texas, United States

Alzheimer's disease (AD) results in a progressive loss of cognitive function and dementia, affecting in the United States more than one in nine people by the time they reach 65 years of age. Strikingly, the neuropathological changes in the brain can precede the onset of symptoms by decades, providing an important diagnostic and therapeutic window. Reliable, sensitive, and specific biomarkers for early diagnosis and disease monitoring would be of paramount importance in clinical practice and for the development of specific therapeutic interventions. Glycosylation is emerging as a new frontier in biomarker development, but it is still a largely unexplored field in AD. The relevance of using glycome as a biomarker source is underlined by the key role of glycan structures in disease pathogenesis and pathology modulation, owing to their influence on a variety of essential pathobiological functions, including cell growth and differentiation, cell communication, and neurodegeneration, among others. Also, glycans represent organ-specific markers. We have developed a powerful, highly sensitive, and reproducible LC-MS/MS-based platforms to analyze glycoproteins and glycans. Due to the fact that all well-known AD-related proteins are glycosylated, AD patients brain tissue, CSF and serum would present a unique glycome/glycoproteome biosignature indicative of AD pathophysiology;which will have distinctive representation that show common and unique representation in tissue and biofluids. Specific glycans, glycan isomers and/or glycoproteins detectable in biofluids are markers associated with neurodegenerative processes in patients with AD. Appropriate markers, alone and/or in combination, can provide information critical for the diagnosis of neurodegenerative diseases such as AD. LC-MS/MS will be utilized to define the glycome and glycoproteome changes associated with the development and progression of AD. The LC-MS/MS to be used will facilitate the characterization of glycan and glycopeptide isomers, providing the first assessment of the changes in the isomeric distribution of glycan/glycopeptide in AD.

Contact: Yehia Mechref, Yehia.Mechref@ttu.edu

P09.09 Full Glycoproteomics Analysis Using FragPipe: From Peptide Identification to Glycan FDR Assessment and Quantitation

Daniel Polasky (1), Daniel Geiszler (1), Fengchao Yu (1), Yi Hsiao (1), Hui-Yin Chang (1), Guo Ci Teo (1), Alexey Nesvizhskii (1)
(1) University of Michigan, Ann Arbor, Michigan, United States

Analysis of intact glycopeptides by tandem mass spectrometry offers great potential for simultaneously identifying glycosylation sites in proteins and the glycans that occupy them at a proteome-wide scale. Fully characterizing a glycopeptide involves identifying the peptide sequence, site(s) of glycosylation, and glycan composition and/or structure, a challenging task given the heterogeneity and complexity of glycosylation. Large-scale glycoproteomics analyses also require validation of these identifications at a given false discovery rate, and it is often desirable to make quantitative comparisons between states, such as between healthy and diseased samples. Here, we present a full workflow for glycoproteomics analysis in FragPipe, a computational pipeline and user interface containing many individual tools. Our method is built around the highly sensitive and fast MSFragger glyco search, which identifies glycopeptides as the combination of a peptide sequence and a glycan mass and supports both N- and O-glycans and glycan and peptide databases of arbitrary size. MSFragger can localize glycans fragmented by collisional activation using peptide fragments retaining a single HexNAc, or intact glycans using localization-aware offset search. We have recently added glycan composition assignment in PTM-Shepherd, a post-search annotation tool, along with glycan-specific FDR control, and we demonstrate its capability to maintain the desired glycan FDR in the presence of entrapment glycans known not to be present in the sample. Finally, assigned glycopeptides can be quantified using IonQuant for label-free quantitation or TMT-Integrator for TMT-based quantitation. In addition to TMT-Integrator's summary reports at gene, protein, peptide, and modification site levels, we have added an additional summary at the glycan level, allowing the abundances of individual glycans at a given glycosite to be compared across states when appropriate, or collapsed to glycosite, peptide, protein, or gene levels to support a wide range of analyses.

Contact: Daniel Polasky, dpolasky@umich.edu

P09.10 Ancestry-Dependent Metabolic Links to Socioeconomic Stresses Revealed in Normal Breast Tissue at Risk for Breast Cancer

Denys Rujchanarong (1), Yeonhee Park (2), Danielle Scott (1), Sean Brown (3), Richard Drak (1), Marvella Ford (1), Michael Ostrowski (1), George Sandusky (4), Harikrishna Nakshatri (5), Peggi Angel (6)
(1) Medical University of South Carolina, charleston, South Carolina, United States
(2) University of Wisconsin-Madison, Madison, Wisconsin, United States
(3) Medical University of South Carolina, Charleston, United States
(4) Indiana University School of Medicine, Indianapolis, Indiana
(5) Indiana University School of Medicine, Indianapolis, Indiana, United States
(6) Medical University of South Carolina, Charleston, South Carolina

A primary difference between black women (BW) and white women (WW) diagnosed with breast cancer is not incidence, but the aggressiveness of the tumor. BW have higher mortalities with similar incidence of breast cancer compared to other race/ethnicities. Black women are diagnosed at younger ages with more advanced tumors that have double the rate of lethal, triple negative breast cancers. There is ongoing debate regarding whether the underlying cause of higher mortality relates more to healthcare inequalities or to ancestry-dependent molecular features found in normal breast tissue that facilitate the aggressive phenotype found in BW. Our hypothesis is that chronic social and economic stressors that differ by race/ethnicity result in molecular responses that create a tumor permissive tissue microenvironment in normal breast tissue. In this study, we investigate molecular pattern differences in a cohort of ancestry defined normal breast tissue from the Susan G. Komen tissue bank with a significant 5-year risk of breast cancer by Gail score. The Gail model is a well-established tool for assessing breast cancer risk based on demographic and clinical data (bcrisktool.cancer.gov). We test N-glycosylation alterations, a glucose metabolism-linked post-translational modification, against stressors marital status,

education, income, reproductive history, BMI and age. Out of 55 N-glycans profiled by glycomics mass spectrometry, a specific N-glycan appeared dependent on ancestry with high sensitivity and specificity (AUC 0.788, p-value<0.0001). When controlling for BMI, 12 N-glycans could report potential ancestry-dependent differences. Interestingly, a linear regression model with ancestry as group variable and socioeconomic covariates as predictors linked N-glycans to different socioeconomic stresses. The data suggests that metabolic patterns linked to socioeconomic stresses may contribute to breast cancer risk dependent on ancestry. This study lays the foundation for understanding the complexities linking socioeconomic stresses and molecular factors to their role in ancestry-dependent breast cancer risk in black women.

Contact: Denys Rujchanarong, rujchana@musc.edu

P09.11 A confidence score for distinguishing structural isomers on N-linked and O-linked glycopeptides with ion-mobility mass spectrometry

Zia Rahman (3), Wendy Sun (1), Xiyue Zhang (1), Hieu Tran (2), Baozhen Shan (1)
(1) Bioinformatics Solutions Inc, Waterloo, Ontario, Canada
(2) University of Waterloo, Waterloo, Ontario, Canada
(3) Bioinformatics Solutions Inc., Waterloo, Ontario, Canada

Introduction: Understanding the impact of glycosylation on protein function requires detailed knowledge of individual glycan structures and their site-specific distribution. LC-MS is a very sensitive technique that is widely used for studying site-specific protein glycosylation. Ion-mobility mass spectrometry provides additional structural information, CCS values, for distinguishing structural isomers. However, very few reports focus on the algorithm to evaluate the glycan structural confidence for glycopeptide identification with the differentiation of minor changes in glycan structure. In this work, we proposed an S-Score for a glycopeptide identification from a MS2, indicating structural confidence of a glycan among different isomers. This S-Score applies both N-linked and O-linked glycopeptide analysis with tandem mass spectra.
Methods: There are three steps to calculate S-Score. 1. Given a MS2 from a glycopeptides, a protein database and glycan database, find the candidates of glycans which match the glycan-associates fragment ions. 2. For each glycan candidate, find the peptide sequence which best fits the peptide fragment ions. A glycan score is associated with each glycan candidate, including glycan fragment ions, retention time and CCS values. 3. Evaluate all glycopeptide candidates, and associated a probability score for each candidate. The S-Score is the glycan score gap of the top two candidates. Output the glycopeptides identification with best probability score, associating the S-score as the structural confidence. Preliminary data: The algorithm was implemented into PEAKS GlycanFinder, and tested with two large-scale mass spectrometry datasets from timsTOF instrument. 19315 distinct site-specific intact N-glycopeptides from 2453 glycoproteins in mouse brain tissue were identified. The glycoprotein site-specific microheterogeneity for 6932 sites was characterized. 2685 O-glycopeptides were identified on a human serum. Among all the identification, the S-Score showed good accuracy to distinguishing structural isomers. Conclusion: A structural confidence score was proposed for the glycopeptide identification.

Contact: Baozhen Shan, bshan@bioinfor.com

P09.12 Evaluation of Serum/Plasma Sample Preparation Methods for Mass Spectrometry-Based Glycoproteomic Analysis

Gege Xu (1), Rachel Rice (1), Mingqi Liu (1), Maurice Wong (1), Diane Tu (1), Klaus Lindpaintner (1), Xin Cong (1)
(1) InterVenn Biosciences, South San Francisco, California, United States

Purpose: Efficient and reproducible sample preparation is critical to proteomic and glycoproteomic analysis of body fluids. A variety of sample preparation protocols and kits are available for proteomic analysis, all based on enzymatic digestion. However, glycopeptides are more susceptible to missed cleavages and degradation, and little is known about the performance of these methods for glycoprotein digestion. Here we evaluated different serum/plasma digestion protocols to show their applicability for targeted glycopeptide quantification. Methods: Pooled human serum and plasma were used to evaluate the digestion efficiency of glycoproteins. Commercially available products including RapiGest SF surfactant (Waters), S-Trap (Protifi), and rapid trypsin/LysC (Promega) were evaluated. Samples were digested in-solution or on-column following protocols in the respective manuals and compared with samples prepared using overnight water bath incubation and microwave-assisted digestion at optimized temperatures and durations. All experiments were performed on both serum and plasma with four replicates per condition and the digested samples were analyzed by SDS-PAGE, UHPLC-QTOF and UHPLC-QQQ to assess the digestion completeness, precision, and undesired modifications. Results: RapiGest SF surfactant, S-Trap, Rapid trypsin/LysC, and microwave-assisted digestion protocols have the advantage of completing digestion within 30 to 120 min while achieving sufficient digestion efficiency and precision for most proteomics applications. However, we found that for targeted glycopeptide analysis each of these approaches introduced certain biases towards individual classes of glycopeptides. For example, Rapid trypsin/LysC resulted in incomplete digestion of glycopeptides from certain proteins, e.g. antichymotrypsin. The RapiGest SF surfactant protocol caused degradation of multiply sialylated glycopeptides due to the high acid concentration required in the quenching step. S-Trap was helpful for reducing glycopeptide missed cleavages in plasma, but less effective for serum. Conclusion: The results of this study can serve as a reference for selecting the optimal sample preparation method for glycoproteomic analysis based on digestion efficiency, time, and cost.

Contact: Gege Xu, ggxu@venn.bio

P09.13 A high-throughput C18 clean-up workflow for glycoproteomic analysis of peripheral blood samples

Bo Zhou (1), Diane Tu (1), Khushbu Desai (1), Daniel Serie (1), Klaus Lindpaintner (1), Xin Cong (1)
(1) InterVenn Biosciences, San Francisco, California, United States

Purpose: Clean-up of serum/plasma samples prior to mass spectrometry is essential for robust performance of (glyco)proteomic analyses. Reverse-phase (RP) resins, particularly C18 matrix, are often used to accomplish this. However, the hydrophilicity of glycopeptides makes them bind to C18 resins unfavorably. Moreover, manual methods are time-consuming and labor-intensive. To address these limitations, we developed a high-throughput clean-up workflow suitable for population-scale glycoproteomic analysis. Methods: For workflow optimization, human serum samples were digested with trypsin. The resulting peptides were loaded onto AssayMAP C18 cartridges on an Agilent Bravo Platform for sample cleanup, followed by analysis on a triple quadrupole mass spectrometer in dynamic multiple reaction monitoring (dMRM) mode. 1,158 MRM transitions were used to quantify 525 glycopeptides and 151 peptides. Raw data was processed using an in-house developed data processing software. Parameters such as sample loading, sample buffer, equilibration buffer, loading volume/flow rate, and elution buffer were optimized. Subsequently, the optimized workflow was applied to analyze plasma samples from 132 melanoma patients. Results: Pooled human serum samples (n=8) were used for workflow optimization. The median coefficients of variation (CVs) were 12.7% and 8.5% for glycopeptides and peptides, respectively after C18 clean-up, and 13.1% and 10.2% prior to C18 clean-up. The median recovery of glycopeptides and peptides after

C18 clean-up were 71% and 72%, respectively. By using the optimized workflow, we were able to process 96 patient samples within 3 hours. The Pearson coefficient of the peak areas for glycopeptides measured prior to and after C18 clean-up in plasma samples of 132 melanoma patients was >0.99, indicating that no systematic bias was introduced by the C18 clean-up for glycopeptides. Conclusion: We have established a method that enables a robust, high-throughput C18 clean-up workflow with minimal manual intervention for reproducible and unbiased glycoproteomic analysis.

Contact: Bo Zhou, bzhou@venn.bio

P10: Metabolomics and Imaging Mass Spectrometry

P10.01 Revealing the Wound Healing Process with Three-Dimensional Mass Spectrometry Imaging using Infrared Matrix-Assisted Laser Desorption Ionization

Hongxia Bai (1), Yining Liu (2), Alejandra Arnedo (2), Tatiana Segura (2), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States
(2) Duke University, Durham, North Carolina, United States

Chronic wound is caused by a wound failing to proceed through a timely reparative order. Millions of people live with chronic wounds presenting an economic burden to the healthcare system in the US. Understanding lipid and metabolite migration during the wound healing process will offer insights into how skin responds to a wound and may provide guidance for wound healing solutions. 3D mass spectrometry imaging (MSI) can determine volumetric molecular distribution, with the mass spectrum and localization information collected in both XY- and Z-dimensions throughout different skin layers giving a comprehensive molecular image of the wound-healing process. Most 3D MSI approaches use MALDI or DESI and reconstruct 3D images from 2D mappings of serial sections, resulting in loss of 3D information and challenges of image registration. To overcome these limitations, we applied infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI), which is a hybrid ambient ionization technique, to sequentially image a sample by consecutive ablation events for a 3D image generation. In this study, 3D IR-MALDESI MSI is applied to a wounded skin created by biopsy punch, the molecular profile of which is then compared to the healthy skin in the same nude mouse. Depth resolution was explored with a confocal laser scanning microscope to establish the imaging parameters for the 3D experiment. A total of 810 and 1269 lipids were putatively detected in the wildtype and wounded skin respectively in a 70-layer 3D experiment. Detailed lipid profiles across different skin layers were revealed and correlated with histological images to characterize the skin structure with a lateral resolution of 50 μm and a depth resolution of 10 μm. Current work involves further statistical analysis and integration of proteomics data, with the aim for multi-omics interpretation of the wound healing process and providing potential therapeutic targets for chronic wound.

Contact: Hongxia Bai, hbai3@ncsu.edu

P10.02 Metabolomic Characterization of Pseudomonas protegens Secretions for Identification of Biopesticides Against Bacterial Panicle Blight of Rice

Samantha Balboa (1), Maia Vierengel (1), Kristen Gates (2), Clemencia Rojas (2), Leslie Hicks (1)
(1) University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States
(2) University of Arkansas, Fayetteville, Arkansas, United States

Bacterial Panicle Blight (BPB) is a rice disease caused by Burkholderia glumae that has become a global concern in reducing rice yields. Currently, there are no effective chemical treatments or resistant cultivars to control BPB. We have observed that bacterial secreted fractions from the antagonistic microorganism Pseudomonas protegens PBL3 inhibits B. glumae infections in rice plants. However, the responsible antimicrobial molecules are unknown. Genome comparison between P. protegens PBL3 and the reference strain P. protegens CHA0, a well-characterized microbial biological control agent, revealed the presence of 14 known biosynthetic gene clusters (BGCs) in P. protegens PBL3. From this genomic data, it is likely that P. protegens PBL3 secretes metabolites with activity against BPB that are not produced by P. protegens CHA0. Cultures of P. protegens PBL3 and the reference strain, P. protegens CHA0, were grown in M9 minimal media

supplemented with 1% myo-inositol. The metabolite secretome of P. protegens PBL3 and CHA0 were harvested in log phase and analyzed with a Waters Acquity H-Class HPLC coupled to a Thermo Q-Exactive HF-X. In addition, active PBL3 fractions against B. glumae were analyzed via LC-MS to confirm the presence of bioactive compounds. The resulting data were searched against the Global Natural Products Social Molecular Networking (GNPS) spectral library. For potential antimicrobial peptide identification, data were searched against the Mascot database. To validate these results, known bioactive metabolites in P. protegens CHA0 were identified. By utilizing LC-MS/MS based metabolomics, several compounds were uniquely identified in P. protegens PBL3 that could be used as biocontrol agents against BPB in rice.

Contact: Samantha Balboa, sjbalboa@live.unc.edu

P10.03 Using Covaris AFA for Fast, Reliable, and Automatable Extraction of Analytes from cells for MALDI-TOF Based Analyses

Tom O'Hare (1), Jay Abramo (1), DEBADEEP BHATTACHARYYA (1)
(1) Covaris, Woburn, Massachusetts, United States

Comprehensive proteomic and metabolomic profile can help in understanding the nature of a disease to developing therapeutics to monitoring patients and determining their therapeutic regime. Regardless of the detection technology, extraction of relevant biomolecules from cells can be laborious and complex. Typical approaches such as, bead beating, tissue homogenizers, etc. require high temperature or solvents requiring long time to stabilize. Adaptive Focused Acoustics(R) technology (AFA(R)) enables reliable extraction of biomolecules under isothermal conditions and can isolate metabolites and proteins from up to 96 different samples in ~ 2 mins. This study showcases AFA based sample preparation for rapid extraction of proteins and metabolites from a variety of microbes. Covaris truXTRAC Extraction Solvent was used for suspension of cells in a 96 AFA-TUBE TPX Plate, where lysis was performed using AFA followed by elution and MALDI-TOF analysis in positive ion linear mode with a nitrogen laser in a randomized pattern. The number of cells per well of the plate ranged from 1 x 10^7 cells for yeast to 1 x 10^5 cells for bacteria. Manual analysis of MS data using E. Coli Metabolome Database and UniProt accurately identified several proteins (50S ribosomal protein, subunit of DNA polymerase, surface membrane lipopolysaccharides) and metabolites critical for cell growth. Comparison of MS data of samples extracted from E. Coli grown in LB and M9 minimal media shows several similar signals. However, several signals in the lower mass range (~1-5 kDa) was observed only in LB media highlighting production of bacteria under stressful conditions. This study demonstrates reproducible and robust MALDI-TOF MS data after AFA enabled cell lysis of cultured cells with minimal variability, and can support high-throughput requirements. This robust method can extract sufficient sample from bacteria with a 2-second pulse or from yeast in 1 minute.

Contact: Debadeep BHATTACHARYYA, dbhattacharyya@covaris.com

P10.04 Leveraging Public Untargeted Metabolomics Data To Propagate Structurally Related Molecule Annotations to Millions of MS/MS Spectra

Wout Bittremieux (1), Sydney P. Thomas (1), Nicole Avalon (1), Mingxun Wang (1), Pieter C. Dorrestein (1)
(1) UC San Diego, La Jolla, California, United States

One of the key goals of untargeted tandem mass spectrometry (MS/MS) metabolomics is discovering biologically relevant molecules. Currently there is still a lot of discovery potential: using spectral library searching, on average only ~5% of the

data can be annotated. This means that the vast majority of data that are collected do not yield any biological insights. I will present a strategy to identify molecules that are structurally related to previously known reference molecules using repository-scale molecular networking. Based on spectral similarity, information can be propagated to neighboring MS/MS spectra in a molecular network to increase the spectrum annotation rate. We have propagated annotations from molecular networks associated with 1.2 billion MS/MS spectra from 1,335 public untargeted metabolomics datasets in various metabolomics data repositories, including GNPS/MassIVE, Metabolights, and MetabolomicsWorkbench, to create the GNPS nearest neighbor suspect spectral library. It consists of 87,916 novel reference spectra corresponding to modified molecules that are structurally related to known reference molecules. Repository-scale molecular networking to create the suspect library revealed 1,350 common modification mass differences, which provide chemical insights into the processes that molecules undergo in vivo and during mass spectrometry analysis. Using the suspect library for spectral library searching boosts the spectrum annotation rate by four-fold on average, considerably increasing the interpretation rate of untargeted metabolomics beyond the state of the art. To demonstrate the performance of the suspect library, suspect annotations enabled the discovery of 969 acylcarnitines, including significant acylcarnitine signatures for Alzheimer's disease patients, providing biomedically relevant insights into changes in energy metabolism; as well as for natural products drug discovery. The nearest neighbor suspect spectral library is freely available with an open license on GNPS for community spectral library searching, where it can be used to provide novel hypotheses for previously unexplored untargeted metabolomics data.

Contact: Wout Bittremieux, wbittremieux@health.ucsd.edu

P10.05 Glycan and Collagen Imaging Mass Spectrometry Analysis of a Broad Tumor Spectrum of Prostate Cancer Tissue Histopathologies

Jordan Hartig (1), Peggi Angel (1), Lydia Liu (2), Amanda Khoo (2), Stan Liu (3), Michelle Downs (3), Paul Boutros (4), Thomas Kislinger (5), Richard Drake (1)
(1) Medical University of South Carolina, Charleston, South Carolina, United States
(2) University of Toronto Princess Margaret Cancer Centre, Toronto, Ontario, Canada
(3) Sunnybrook Health Science Centre, Toronto, Ontario, Canada
(4) University of California, Los Angeles, Los Angeles, California, United States
(5) University of Toronto Princess Margaret Cancer Centre, Toronto, Ontario

Prostate cancer progression can be characterized at the tissue level by changes in glycosylation and the emergence of reactive stroma, which represent potential target areas for cancer biomarker discovery. Using established imaging mass spectrometry (IMS) methods designed to target N-glycans and collagen/extracellular matrix (ECM), a cohort of pathology-annotated FFPE prostate tissues linked to a multi-omic and mult-investigator consortia were evaluated. Tissues represented a broad spectrum of prostate cancer histopathologies, with the goal of identifying specific glycan and collagen biomarker candidates for integration with other omic data sets from the same cohort. analyses. Each slide was processed for N-glycan IMS by peptide N-glycosidase F (PNGase F PRIME) digestion using established protocols. Released N-glycans were detected by MALDI-QTOF IMS on a timsTOF flex mass spectrometer. Following enzyme digestion and imaging, slides were stained using established H&E staining protocols. High resolution images of the stained tissues were saved. Collagenase digestion and further MALDI-QTOF detection of ECM peptides was performed following H&E staining. Each tissue sample contained different combinations of benign tumors, fused tumors, cribriform tumors, glomeruloid bodies, poorly formed glands (PFG), and intraductal carcinomas, all with a Gleason score of at least 3, 4, or 5 When comparing H&E images to glycan profiles, it was found that the major tumor-associated glycans were in the high mannose and paucimannose

categories. Multiple branched N-glycan species with different combinations of fucose and sialic acid constituents were associated with tumor incidence and progression. Additional proteomic studies are being conducted in ongoing analyses to evaluate prevalence and identity of ECM and collagen peptide distribution in correlation with different tumor subtypes. The combination of multi-enzymatic digests with histopathology annotations represents an extensive and multi-dimensional profile of prostate cancer and a novel tissue biomarker discovery approach linked with other corresponding omic datasets.

Contact: Jordan Hartig, hartig@musc.edu

P10.06 Improved Spatial Resolution of Infrared Matrix-Assisted Laser Desorption Electrospray Ionization (IR-MALDESI) Mass Spectrometry Imaging (MSI)

Alena Joignant (1), Hellena Bai (1), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States

Mass spectrometry imaging (MSI) is a powerful method for mapping chemical information to tissue structures. Infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) is a hybrid ionization method that regularly achieves 120 microns in its current configuration, yet definition of minute structures is limited to the spatial resolution capability of the optical train. Spatial resolution of this method was previously increased with oversampling, where ablation spots are overlapped to achieve cellular and sub-cellular ablation areas[1]. A different laser and optic configuration was previously shown to decrease the available spatial resolution to 50 microns, which was nearly diffraction-limited[2]. We propose a new optical train involving a reflective, Schwarzschild-like optic that gets us closer to a diffraction-limited spot size and improves IR-MALDESI spatial resolution without the need for oversampling. By avoiding oversampling, the sample does not need to be completely ablated in order to achieve high spatial resolution, which is advantageous for the use of thicker samples. Additionally, decreasing the spot size lowers the IR laser energy output required for sufficient sample ablation, which allows for increased speed of analysis. A decrease in spot size scales the spatial resolution to that of optical imaging, allowing for more direct multi-modal imaging. The optical path and laser profile are optimized as examined on thermal paper by microscopy. Although a decrease in spot area is logistically challenging due to increased analysis time and is limited by the sensitivity of the instrument, the information received by such fine combing of tissue structure has positive implications for cellular analysis, such as imaging islet cells in pancreas tissue and similar structures. References: 1. Anal. Bioanal. Chem. 2015, 407, 2265-2271. 2. J. Am. Soc. Mass. Spectrom. 2017, 28, 2099–2107.

Contact: Alena Joignant, anjoigna@ncsu.edu

P10.07 Novel Matrix Strategies for Improved Ionization and Spatial Resolution for IR-MALDESI Mass Spectrometry Imaging

Russell Kibbe (1), Allyson Mellinger (1), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States

Infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) utilizes energy-absorbing matrices that surround the analyte and desorb molecules from the surface through infrared laser ablation for post-ionization and detection. A matrix with high infrared absorbance is important to maximize laser efficiency and analyte desorption. The deposited ice layer matrix is fundamental to IR-MALDESI and mass spectrometry imaging (MSI) because of the prevalence of water in biological systems and the absorbance of IR wavelengths in the O-H stretch. In this effort, we designed new

matrix strategies with the working hypothesis being that the newly constructed matrices will absorb more infrared energy, leading to a greater number of desorbed neutral molecules, leading to higher ion abundance, than the deposited ice layer matrix. This research was carried out by coupling the IR-MALDESI with a high resolution/accurate mass Orbitrap mass spectrometer (OE 240) to measure the ion abundance of these newly developed strategies and spot size (using Leica microscope) to demonstrate the improvement over the ice matrix.

Contact: Russell Kibbe, rkibbe@ncsu.edu

P10.08 A Novel Ion-beam Imaging Technique for Detection of Chromosomal Copy Number Alterations Associated with Progression of Breast Cancer

Rashmi Kumar (1), Sushama Varma (2), Mike Angelo (1), Robert West (1)
(1) Stanford University School of Medicine, Stanford, California, United States
(2) Stanford University, Stanford, California, United States

Purpose:Ductal Carcinoma in situ (DCIS) comprises about 25% of breast cancers detected annually. It is hard to predict which DCIS cases will progress to invasive breast cancer. Recent studies have demonstrated that certain genomic copy-number-alterations (CNAs) are associated with higher risk of DCIS progression. In this project, we aim to develop a mass-spectrometric imaging method to detect and spatially map CNAs in FFPE tissues. Using this novel technique, we plan to investigate the role of tumor genomics in DCIS progression. Methods:FFPE sections were deparaffinized and prepared for in situ hybridization (ISH) assay. Haptens were incorporated into the genomic targets. Next the tissues were incubated with metal-tagged antibodies, specific to haptens. Multiplexed Ion Beam Imaging (MIBI) technique was used for analysis. MIBI is a novel imaging technique, combining immunohistochemistry with time-of-flight (TOF) mass-spectrometry. An ion-beam gun rastered on the tissue-sections containing slides under vacuum. This freed the metal ions from specimen to be measured by TOF. The abundance of each metal was read out in pixel-by-pixel fashion which corresponded to the abundance of that antibody Results:HER2, an important chromosomal ROIs in breast cancer was chosen as the target for all the experiments. Bacterial artificial chromosomes containing chromosomal region of interest were used as DNA-probes and nick translation was carried out to introduce haptens. The probes were then allowed to hybridize overnight. After hybridization, the tissue sections were incubated with metal-tagged anti-hapten antibodies followed by imaging using MIBI. We observed significant correlation in spatial map of HER2 between immunofluorescence and MIBI images. A pair of HER2 +signal in each cell of lymph-node and placenta control tissues was observed. We observed amplified HER2 signal in HER2-positive breast tissues whereas HER2 signal was depleted in HER2-negative breast tissues. Conclusions:A new technique for spatially mapping genomic-ROIs using MIBI has been developed.

Contact: Rashmi Kumar, rashmi23@stanford.edu

P10.09 Development of a Metabolomics System Suitability Sample for MS-based Metabolomics

Tracey Schock (1), Debra Ellisor (1), Clay Davis (1)
(1) National Institute of Standards and Technology, Charleston, South Carolina, United States

Prior to the analysis of biological samples, one must first assess the suitability of the analytical instrumentation. In this regard, the metabolomics community requires an everyday system suitability standard by which to benchmark instrument performance for untargeted MS based approaches. A complex solution that mimics biological samples is needed to

determine whether an analytical run is of acceptable quality and to ensure lack of contamination prior to experimental analyses. Individual laboratories use a small number of standards (5-20) to create in-house suitability solutions for assessing measurement quality of hundreds to thousands of chemicals profiled in an untargeted study of complex samples. Over-reporting and spurious conclusions are likely rampant in the literature due to the lack of a material to evaluate measurement quality across a complex omics profile. NIST is developing a large quantity, biological extract from human liver which incorporates the complexity of a metabolome, resulting in a more encompassing system suitability sample. The design of a tissue extract as a suitability standard eliminates sample preparation variation observed with biological samples while offering simplicity of use and a chemically diverse range of analytes for analysis of metabolomics platforms. Additionally, the extract can be a tool in harmonization of instrument performance across batches and in large, multi-center studies. NIST created a candidate research grade material (RGM) 10122 Metabolomics System Suitability Sample for community evaluation of the material as a quality control material for the metabolomics community. This sample is available pre-sale for those interested. Currently we are in production for the RM 8470 Metabolomics System Suitability Sample which will be the commercially available liver extract product.

Contact: Tracey Schock, tracey.schock@nist.gov

P10.10 Spatial Metabolomics of Xenopus laevis Tadpoles by IR-MALDESI Mass Spectrometry Imaging

Whitney Stutts (1), Julia Grzymkowsi (2), Nanette Nascone-Yoder (2)
(1) Molecular Education, Technology and Research Innovation Center (METRIC) and The Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, North Carolina, United States
(2) Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, North Carolina, United States

Xenopus laevis are an important model organism for studying vertebrate development and disease. At the tadpole stage, functional organs have formed and metabolic perturbations resulting from experimental interventions can be visualized. In this work, infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) mass spectrometry imaging (MSI) was employed to spatially resolve tissue-specific molecular markers in frontal, whole organism sections from Xenopus laevis tadpoles. High-resolution accurate mass measurements were acquired with a resolving power of 240,000 (FWHM m/z 200) in positive and negative ionization modes to enhance metabolome coverage. Resulting mass spectra were converted to imzML files and uploaded to METASPACE, a web-based platform for metabolite annotation of MSI data. Over 1000 metabolites and lipids were annotated at an FDR of ≤10% after filtering for on-sample only ions. Highly localized metabolites including amino acids, fatty acyls, glycerolipids, glycerophospholipids, sphingolipids and sterol lipids were detected in distinct tissue regions such as the eyes, facial cartilage, including the branchial arches and Meckel's cartilage, intestines and stomach. In addition to metabolic profiling, this work presents the application of an optimized embedding and cryosectioning method for delicate tissues such as the Xenopus gut. The results described demonstrate the utility of IR-MALDESI MSI for studying the Xenopus metabolome and illustrate the potential for novel insights in developmental biology and toxicology studies.

Contact: Whitney Stutts, wlstutts@ncsu.edu

P10.11 Investigation of Sucrose as a Cryoprotectant for Tissue Preservation and Compatibility with IR-MALDESI Mass Spectrometry Imaging

Mary Wang (1), Alexandria Sohn (1), Kevin Erning (2), Tatiana Segura (3), Juhi Samal (3), David Muddiman (1)
(1) North Carolina State University, Raleigh, North Carolina, United States
(2) Duke University, Durham, North Carolina, Kazakhstan
(3) Duke University, Durham, North Carolina, United States

Formalin-fixed paraffin embedded (FFPE) and flash-frozen tissues are methods previously used to preserve biological tissue samples prior to analysis with infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI). FFPE tissues have been used primarily for glycan research as the tissues are more conducive to IR-MALDESI MSI when compared to fresh or flash-frozen tissues. FFPE tissues require sample preparation due to formalin fixation and paraffin wax embedding. Other methods of tissue preservation such as sucrose embedding are being investigated as an alternative to FFPE and flash frozen tissues. Sucrose embedding is a specialized version of fresh or flash-frozen tissue preparation using sucrose as a cryoprotectant for samples in order to protect the cell structure morphology during the freezing process. Sucrose embedding tissues for metabolite, lipid, and/or glycan imaging could yield an alternative method – results will compare molecular classes and their ion abundance relative to fresh frozen tissues. In this study, two female mice brains underwent separate preservation methods where one was flash frozen sample without sucrose embedding and the other embedded in sucrose embedding prior to freezing. Tissues were sectioned and evaluated using IR-MALDESI mass spectrometry imaging to determine the optimal tissue fixation method using ion abundance as a function of molecular class (polar metabolites, non-polar metabolites and N-linked glycans).

Contact: Mary Wang, mfwang@ncsu.edu

P10.12 Enhancing Metabolomic Coverage in Positive Ionization Mode Using Dicationic Reagents by IR-MALDESI MSI

Ying Xi (1), David Muddiman (1)
(1) NCSU, Raleigh, North Carolina, United States

Mass spectrometry imaging is a powerful tool to analyze a large number of metabolites with their spatial coordinates collected throughout the sample. However, the significant differences in ionization efficiency pose a big challenge to metabolomic mass spectrometry imaging. To solve the challenge and obtain a complete data profile, researchers typically perform experiments in both positive and negative ionization modes, which is time-consuming. In this work, we evaluated the use of the dicaionic reagent, 1,5-pentanediyl-bis(1-butylpyrrolidinium) difluoride (abbreviated to $[C_5(bpyr)_2]F_2$) to detect a broad range of metabolites in the positive ionization mode by infrared matrix-assisted laser desorption electrospray ionization mass spectrometry imaging (IR-MALDESI MSI). $[C_5(bpyr)_2]F_2$ at 10 µM was doped in 50% MeOH/H_2O (v/v) electrospray solvent to form +1 charged adducted ions with anionic species (-1 charged) through post-electrospray ionization. This method was demonstrated on sectioned rat liver and hen ovary. 73 deprotonated metabolites from rat liver tissue sections were successfully detected and putatively identified in the positive ionization polarity through adduct formation with the dicationic reagent, along with 164 positively charged metabolite ions commonly seen in positive ionization mode, which resulted in 44% increased molecular coverage. In addition, we were able to generate images of hen ovary sections showing their morphological features. Following-up tandem mass spectrometry (MS/MS) indicated that this dicationic reagent $[C_5(bpyr)_2]^{2+}$ could bond with the headgroup of lipid ions, implying a possible mechanism behind the adduct formation. The application of the dicationic reagent $[C_5(bpyr)_2]^{2+}$ on tissue sections provides a rapid and effective way to enhance the detection of metabolites in positive ionization mode.

Contact: Ying Xi, yxi3@ncsu.edu

P11: Multi-omics Methods and Applications

P11.01 Multi-omics analysis of S. pneumoniae extracellular vesicles

Stephanie Biedka (1), Rory Eutsey (2), Luisa Hiller (2), Amber Lucas (1)
(1) Impact Proteomics, Pittsburgh, Pennsylvania, United States
(2) Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

Extracellular vesicles (EVs) are cell-derived nanoparticles that can serve as carriers of DNA, RNA, and protein. For the pathogen S. pneumoniae, which causes diseases such as pneumonia and meningitis, EVs can be used to transport these macromolecules to recipient cells, facilitating communication and signaling. However, the overall composition of these EVs and how the macromolecules packaged within these EVs contribute to the pathogenicity of S. pneumoniaeis is still unknown. In this work, we have developed a novel multi-omic workflow that isolates DNA, RNA and protein from a single S. pneumoniae EV starting sample, enabling analysis of these macromolecules without the need for multiple sample preparation steps and allowing us to gain insights into the composition and functionality of S. pneumoniae EVs. EVs from a S. pneumoniae culture were isolated and the proteins in the EV homogenate were labeled with the bifunctional ProMTag. One end of the ProMTag forms a reversible, covalent bond with proteins. ProMTag's other functional group is methyltetrazine, which forms an irreversible bond with trans-Cyclooctene (TCO), allowing capture of EV proteins on TCO-agarose beads. Organic solvent was then added to precipitate the nucleic acids in the EV homogenate. Nucleic acids were released by a series of washes and subsequently separated into DNA and RNA fractions by RNase or DNase treatment, respectively. Proteins were then released from the TCO-agarose beads by reversing the ProMTag-protein linkage, yielding proteins in their original, unmodified state ready for analysis. DNA and RNA were sequenced, and proteins were identified using mass spectrometry. Using this workflow, we were able to describe the composition of multiple macromolecules within S. pneumoniae EVs and gain insights into how these macromolecules facilitate communication and infection within the S. pneumoniae population.

Contact: Stephanie Biedka, stephanie.biedka@impactproteomics.com

P11.02 Microbiome multi-omics analysis of human whole stool samples utilizing a single Bligh-Dyer sample preparation method

Stephanie Servetas (1), Tracey Schock (2)
(1) NIST, Gaithersburg, Maryland, United States
(2) NIST, Charleston, South Carolina, United States

The microbiome is a critical component of an individual's health. The structure (bacterial communities) and function (metabolic activity) of one's microbiome can elicit positive and negative physiological mechanisms either leading to a robust health state or incite a foundation for the development of disease. Microbiome composition can be readily determined using 16S rRNA gene sequencing or whole genome sequencing from stool. Metagenomic and metatranscriptomic sequencing technologies can illuminate microbes' functional capacities and states while additional measurements are needed to elucidate the microbiome-host interactions. Stool proteomic, metabolomic, and lipidomic, measurements offer the ability to simultaneously assess the complex interplay between the microbiota, host, and environment. In this study, genomic, proteomic, metabolomic, and lipidomic analyses were conducted on reference stool material as part of the National Institute of Standards and Technology (NIST) Whole Stool Interlaboratory Study (ILS). The primary goals for the ILS were to: 1)

identify metabolites in human whole stool using analytical platforms of your choice; 2) assess fold change metabolite differences in different diet types; and 3) evaluate metabolite differences in aqueous and lyophilized whole stool samples. Samples consisted of four pooled human whole stool samples comprising different diet types and storage conditions. Proteins, metabolites, and lipids were extracted in a single modified Bligh-Dyer sample preparation method. Protein identification were performed using a complied database from the vegan and omnivore genomic profiling analysis of stool samples while metabolite and lipid identifications were preformed using multiple spectral libraries. These identifications and annotations revealed a diverse range of host and microbial proteins, metabolites, and lipids as well as multiple classes of exposure-related compounds including those from dietary sources, medications, drugs, and environmentally relevant exposures. Multivariate statistics was performed to reveal biological differences between phenotypic groups for both vegan and omnivore diets.

Contact: Clay Davis, clay.davis@nist.gov

P11.03 Proteogenomic approach to studying the functional effect of cellular aneuploidy in cancer

Lizabeth Katsnelson (1), Xin Zhao (1), Pan Cheng (1), Teresa Davoli (1), David Fenyö (1)
(1) NYU School of Medicine, New York, New York, United States

Aneuploidy, a hallmark of cancer, is the presence of an abnormal number of chromosomes. While rare and typically detrimental to cellular fitness in normal cells, aneuploidy occurs in ~90% of solid tumors in the form of somatic copy number alterations (SCNAs). It is believed that aneuploidy arises from defective regulation during mitosis, which leads to chromosomal instability. However, its role in tumorigenesis remains unknown. Tumors with high levels of aneuploidy (> 8-10 gains or losses) correlate with poor prognosis, increased immune evasion (decreased immune infiltration), and lowered response to immunotherapy. This suggests that aneuploidy may be both a predictive and prognostic biomarker, as well as a therapeutic target for cancer. Since proteins perform the majority of essential cellular tasks and are the targets of most drugs, it is crucial to study the downstream effects of aneuploidy at the proteomic level. Using publicly available data from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) and the Cancer Cell Line Encyclopedia (CCLE), we reveal differences in the cancer proteome among low and high aneuploid tumors and cell lines across several cancer types. Upregulated pathways in highly aneuploid tumors include translation and protein regulation, mitochondrial regulation, and cellular respiration and metabolism, while downregulated pathways are enriched for immune responses. Utilizing multi-omics data, we also compare immune profiles of low and high aneuploid tumors, which we may then compare to proteomics data to find possible mechanisms of immune evasion in aneuploid tumors. Together, these findings may uncover therapeutic vulnerabilities in highly aneuploid, aggressive cancers.

Contact: Lizabeth Katsnelson, lizabeth.katsnelson@nyulangone.org

---

P11.04 Spatially Targeted Proteomics and Multimodal Imaging Distinguish Molecular Signatures of Amyloid Plaque-Containing Islets from Type 2 Diabetic Donors

Angela R. S. Kruse (1), Josiah C. McMillen (1), John T. Walker (2), Audra M. Judd (1), Nathan Heath Patterson (1), Danielle B. Gutierrez (1), Jeremy L. Norris (1), Jeffrey M. Spraggins (3), Alvin C. Powers (4), Richard M. Caprioli (3)
(1) Vanderbilt University, Nashville, Tennessee, United States
(2) Vanderbilt University School of Medicine, Nashville, Tennessee, United States

(3) Vanderbilt University, Nashville, Tennessee
(4) Vanderbilt University Medical Center, Nashville, Tennessee, United States

The pancreas is a complex organ with a crucial role in digestion and blood glucose regulation. Approximately 2% of the pancreas consists of islets, which are groups of hormone-producing cells. The islets of individuals with type 2 diabetes (T2D) are dysfunctional and may undergo cell death. Amyloid plaques contribute to islet cell death but the molecular environment supporting plaque formation remains poorly understood. Since amyloid plaques vary in morphology and location, comparing molecular signatures of islets with and without plaques is essential. To accomplish this, we developed a multimodal platform combining microscopy, micro-Liquid Extraction Surface Analysis (microLESA), and imaging mass spectrometry (IMS) to study individual islets in pancreas tissue from non-diabetic and T2D human donors. Autofluorescence and polarized light microscopy were used to localize islets and amyloid plaques. Microscopy approaches were integrated with microLESA, in which trypsin droplets were targeted to individual islets for on-tissue digestion. Liquid chromatography-tandem mass spectrometry was performed on digested islet proteins to generate chemically rich data from non-diabetic and T2D islets. Additionally, matrix-assisted laser desorption/ionization-IMS was performed to spatially analyze peptides from non-diabetic and T2D pancreas tissue sections. Differential expression analysis of amyloid-containing versus non-amyloid-containing islets in T2D pancreata identified 15 proteins associated with amyloid plaques, including collagen IV and myristolated alanine-rich C-kinase substrate protein. Further comparison of all islets from non-diabetic and T2D donors identified 85 differentially abundant proteins including peroxiredoxin, carboxypeptidase, and calmodulin. These results suggest that specific enzymes and chaperones are associated with amyloid plaques, potentially as a stress response or an aberrant signal involved in plaque formation. Additionally, peptide IMS showed changes in protein abundance consistent with microLESA data, but also identified peptides differing in localization but not abundance. Together, these findings underscore the importance of a multimodal approach to begin unraveling the molecular basis for amyloid plaque formation in T2D.

Contact: Angela Kruse, angela.kruse@vanderbilt.edu

P11.05 Precipitate to Fractionate: Molecular Weight-Based Fractionation with the ProTrap XG

Jessica Nickerson (1), Venus Baghalabadi (1), Alan Doucette (1)
(1) Dalhousie University, Halifax, Nova Scotia, Canada

Purpose: Multi-omics approaches have emerged to understand disease pathology and drug resistance, probing both the proteome and metabolome of the sample. We herein present a rapid precipitation-based approach to selectively fractionate cellular components, isolating intact proteins, the low-mass peptidome and with opportunity for metabolome profiling by MS. Methods: A test sample of S. cerevisiae lysate was subject to sequential precipitation from 20, 40, 60, 80, and 97% acetone with zinc sulfate. Recovered protein fractions were quantified and visualized by SDS PAGE, and subject to bottom-up LC-MS/MS. A low-molecular weight sample was prepared by trypsin digestion of yeast. Peptides were precipitated with 97% acetone and zinc sulfate in the ProTrap XG. Pellets were re-solubilized quantified and characterized by bottom-up LC-MS/MS. Results: Intact protein pellets recovered from 2 min precipitation in 80% acetone revealed > 95% yield. Bottom-up LC-MS/MS characterization of the pellet and supernatant demonstrated a bias towards the rapid recovery of high molecular weight species ($p < 0.001$). Recovery of low molecular weight peptides revealed 90 ± 10% precipitation yield in 97% acetone and zinc sulfate (LC-MS/MS confirmed a median molecular weight of 2.25 kDa). SDS PAGE analysis of the sequentially precipitated sample revealed a molecular weight trend in the fractions collected from increasing acetone content. Future experiments will apply a sequential precipitation approach to isolate the proteome, and peptidome from a

pancreatic cancer cell line. We will subsequently explore the metabolome isolated in the supernatant fraction from the ProTrap XG, with MS profiling of each sample. Conclusions: We demonstrate a rapid and robust approach for quantitative precipitation of intact protein as well as low molecular weight proteins and peptides facilitated by the ProTrap XG.

Contact: Jessica Nickerson, jessica.nickerson@dal.ca

P12: Post-Translational Modifications

P12.01 Phosphoproteomic Analysis of HeLa Digest using Optimized Zeno EAD IDA Workflow

Alexandra Antonoplis (1), Christie Hunter (1), Nick Morrice (2)
(1) SCIEX, Redwood City, United States
(2) SCIEX, Alderley Park, United Kingdom

Introduction: In data dependent acquisition workflows, collision induced dissociation (CID) is commonly used to fragment peptide ions for analysis. While CID enables identification of many high-confidence peptides and proteins, preserving labile post-translational modifications (PTMs) can be challenging. Electron activated dissociation (EAD) is a fragmentation mode that involves capture of electrons by molecular ions to form radicals that dissociate into ions. EAD often preserves PTMs and provides complimentary sequence information to CID. This work focuses on information dependent acquisition (IDA) method development and analysis of HeLa digest using EAD fragmentation to localize peptide phosphorylation sites. Methods: 100 µg of digested HeLa cell lysate was fractionated using high pH RP-HPLC (43 fractions). Fractions were separated using a 20 min microflow gradient with a flow rate of 6 µL/min. Zeno EAD IDA was performed using the ZenoTOF 7600 system, with EAD reaction time, electron beam current and electron kinetic energy optimized for analysis. The Zeno trap was activated for increased MS/MS sensitivity. MASCOT was used for all database searching and search results were imported into Scaffold for analysis. Results: For Zeno EAD IDA, several method parameters were optimized for data acquisition including EAD reaction time, electron beam current and electron kinetic energy. Analysis of 43 high pH fractions of HeLa digest resulted in 5825 protein identifications and 36,588 peptide identifications, including 135 phosphopeptides. EAD spectra from phosphopeptides identified the sequence and site of phosphorylation using complete z+1 ion series, including localization in peptides with multiple serines. Conclusions: Zeno EAD IDA analysis of 43 HeLa digest fractions using microflow chromatography enabled large scale protein identification, with over 5000 proteins identified and confident identification and site assignment of peptide phosphorylation sites.

Contact: Alexandra Antonoplis, alexandra.antonoplis@sciex.com

P12.02 O-GlcNAc-omic Analysis of Triple Negative Breast Cancer Cells

Bernice Agana (1), Susana Comte-Walters (1), Baylye Burnette (1), Jennifer Bethard (1), Mi-Hye Lee (1), Kim Norris-Caneda (1), Mauricio Reginato (2), Lauren Ball (3)
(1) Medical University of South Carolina, Charleston, South Carolina, United States
(2) Drexel University, Philadelphia, Pennsylvania, United States
(3) Medical University of South Carolina, Charleston, South Carolina

The dynamic, stress and nutrient-sensitive modification of protein Ser/Thr residues with B-N-acetylglucosamine (O-GlcNAc) impacts multiple regulatory processes enhancing proliferation, survival, and metastasis of cancer cells. While global O-GlcNAc modification and O-GlcNAc transferase (OGT) expression is elevated in a number of cancers including breast cancer, characterization of the sites of O-GlcNAc modification has been challenging due to the difficulty in efficiently enriching O-GlcNAcylated peptides and the labile nature of the O-linkage by collisional dissociation during tandem mass spectrometry. To uncover mechanisms regulated by O-GlcNAc cycling enzymes, immunoaffinity enrichment was performed using a newly available anti-O-GlcNAc antibody (Cell Signaling) prior to LC-MS/MS. O-GlcNAcylated peptides were immunoprecipitated from trypsin digested MDA-MB-231 cell lysates following high pH RP fractionation and

concatenation. Alternating HCD and ETD MS/MS were acquired on a ThermoScientific Orbitrap Fusion Lumos with Easy1200 nLC. Spectra were searched for S/T HexNAc, accounting for potential neutral loss of HexNAc by HCD MS/MS, and for phosphorylation using MaxQuant and Protein Prospector which provided complementary results. To date over 1000 putative HexNAc sites were detected from 345 proteins. As previously reported proteins associated with chromatin remodeling, transcription, and translation were heavily O-GlcNAcylated. 233 HexNAc sites occurred at known sites of phosphorylation (PhosphoSitePlus) and 28 HexNAc peptides were also phosphorylated. Spectra from novel, putative O-GlcNAcylated proteins (~80) will be manually interrogated to confirm the presence of diagnostic GlcNAc fragment ions and the site(s) of modification. Ongoing studies are aimed at coupling this approach to quantitative proteomics to identify differentially modified sites of O-GlcNAcylation from tissue samples and breast cancer cells.

Contact: Lauren Ball, ballle@musc.edu

P12.03 Discovery of the Post-Translational Modification Mechanisms of the Antibacterial Peptide TCLE, to Lead to the Creation of Novel Antibiotics

S. Anisha Daley (1), Diana Calvopina (1), Kimberly Wagstaff (1), Katherine Brown (1), Joel Griffitts (1), John C. Price (1)
(1) Brigham Young University, Provo, Utah, United States

Bacteria are evolving, and in order to keep ourselves from being left behind, we need to improve our antibiotics as well. The TCLE peptide is a bacteria-produced antibiotic that is heavily modified after translation. By learning about the modification processes, we will be able to purposefully alter it to create novel anti-bacterial compounds effective against many diverse bacteria. The TCLE peptide is made up of two parts: the core portion and the leader portion. The core portion is modified by several protein complexes, and the leader portion is where enzyme complexes attach while modifying the core. The core has six cysteines, and the first post-translational modification changes those cysteines into thiazoline intermediates then thiazoles. This is done by a protein complex called TCLIJN. TCLIJN grabs on to the leader portion of the peptide and uses it to find the core and modify the cysteines. The objective of this study is to learn how much of the leader peptide TCLIJN needs to modify the core cysteines. To discover this, we truncated amino acids off the leader portion, starting at 6 and going up to 18 by intervals of three. After co-expressing the peptide with TCLIJN, we then measured how many cysteine to thiazole modifications occurred by using mass spectrometry. The data suggested that at truncations of 15 and beyond, complete cysteine modification is significantly decreased, causing the number of thiazolines and thiazoles to increase. This data explains how much of the leader peptide is essential for TCLIJN's modifications. As we continue to discover the characteristics of the TCLE peptide, we will learn how to more purposefully modify this peptide to create the desired products. Understanding the TCLE peptide and all its facets will pave the way to greater and more purposeful creation of antibiotics.

Contact: S. Anisha Daley, anisha@mdaley.net

P12.04 Linking Dynamic Lysine Acylation with Metabolic Regulation in Bacteria

Janine Y. Fu (1), John M. Muroski (1), Robert P. Gunsalus (1), Rachel R. Ogorzalek Loo (1), Joseph A. Loo (1)
(1) University of California, Los Angeles, Los Angeles, California, United States

The post-translational modification (PTM) N-ε-lysine acylation has gained attention due to its impact in regulating diverse biological mechanisms. Lysine acylation is tightly linked to metabolism as the production of intrinsically reactive metabolites,

reactive acyl-CoA species (RACS), can give rise to non-enzymatic acyl modification. Different metabolic pathways generate various RACS that can lead to diverse acyl-modifications. To fully understand how these different acyl-PTMs affect metabolic regulation, it is crucial to characterize these modifications through a system-wide and unbiased approach. Here, we investigate the acyl-ome in syntrophic bacteria without using PTM-specific enrichment, allowing us to identify a variety of acyl-proteoforms. Using tandem mass spectrometry (LC-MS/MS) on the Orbitrap Exploris 480, we have identified 4 types of acyl modifications (acetyl-, butyryl, 3-hydroxybutyryl-, and crotonyl-lysine) and over 800 sites of acylation. One metabolic pathway in particular, butyrate degradation, is heavily modified with over 100 acyl-PTMs on the enzymes involved. The number and type of modification changes with the microbe's growth conditions and correlates with the buildup of RACS at known metabolic bottleneck points in the microbe's degradation pathway. Many of these sites show extreme heterogeneity in the types of acylation observed. Targeted and quantitative parallel reaction monitoring (PRM) MS experiments measuring acyl-PTM levels also show evidence of acylation cross-talk between multiply modified peptides. We found that the abundance of each acyl-peptide variant significantly changes with growth condition, thus revealing the intimate link between acyl-PTMs and metabolism. As several metabolic proteins have modification hot spots, we also developed an intact MS approach to study these acyl-proteoforms and characterized lysine site reactivities with RACS intermediates. These findings not only demonstrate our comprehensive approach to capture the acyl-ome, but also highlight the potential regulatory role of protein acylation in metabolic function.

Contact: Janine Fu, janinefu@chem.ucla.edu

P12.05 Compared dda-PASEF and prm-PASEF approaches for quantification of 2000 RAS induced Phosphopeptides.

Elizabeth GORDON (1), Matthew Willetts (1), Diego Assis (1), Rodrigo M Ferreira (2), Uma K. Aryal (2)
(1) Bruker Daltonics Inc, BIllerica, Massachusetts, United States
(2) Department of Comparative Pathobiology, Purdue University, West Lafayette, Illinois, United States

Introduction Understanding molecular factors and biological processes of cellular senescence provides important insights into the intrinsic cellular mechanisms for cancer prevention, and organismal aging. Herein, we have used dda-PASEF and prm-PASEF for quantification of RAS induced phospho-peptides, our analysis is added by the ability of the trapped ion mobility to give greater separation to phospho position specific isomers. Methods Human diploid fibroblast strain IMR90 (CCL-186; ATCC, USA) cells were transduced with ER:RAS lentivirus and treated with 100 nM of (Z)-4-Hydroxytamoxifen (4-OHT) for ER:RAS activation and induction of oncogene induced senescence (OISCells were harvested after 6 days of 4-OHT activation for nuclear extraction, trypsinization, and enrichment of phosphopeptides using Polymer-based Metal-ion Affinity Capture (PolyMAC) spin tips. Samples were run on a nanoElute LC (Bruker Daltonics) using a 25 cm Aurora nano column (IonOpticks, Australia) at 400 nl/min with a 70 min gradient, and a longer 120min gradient for global samples. LC-TIMS MS/MS data were obtained from a timsTOF Pro instrument. dda-PASEF data were analyzed using PEAKS OnLine (Bioinformatics Solutions). PRM PASEF data was analyzed with Skyline. Results Global proteome analysis identified 55,000 peptides and 4500 proteins, but less than 1% total identifications came from phosphopeptides. Analysis of 400ng of PolyMAC enriched IMR90 cell digest identified around 20,000 unique phosphopeptides and over 2800 phosphoproteins. This study shows over 1,200 protein groups added quantifiable changes in phosphorylation state. Unlike dda-PASEF, targeted prm-PASEF of selected phosphopeptides, showed that phosphorylation could be seen with RAS induction in both enriched and unenriched samples. Conclusions In IMR90 cells induction of the Estrogen receptor RAS fusion shows that 3,984 peptides were observed to be phosphorylated using dda-PASEF and that they then could be targeted with PRM-PASEF and quantified in a label free approach.

Contact: Elizabeth Gordon, elizabeth.gordon@bruker.com

P12.06 Abscisic Acid Controlled Redox Proteome of Arabidopsis and Its Regulation by Heterotrimeric G-proteins

Amanda Smythers (1), Nikita Bhatnagar (2), Chien Van Ha (2), Evan McConnell (1), Boominathan Mohanasundaram (2), Sona Pandey (2), Leslie Hicks (1)
(1) University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States
(2) Donald Danforth Plant Science Center, St. Louis, Missouri, United States

Reactive oxygen species (ROS) are highly efficient, membrane permeable cell signals that are essential for cellular regulation. These signals can mark macromolecules for degradation, act as a retrograde signal for transcriptional regulation, as well as directly modulate enzymatic activity via the reversible oxidation of cysteine thiols. Although this reversible cysteine oxidation is an essential regulatory mechanism, the resulting modifications are labile and infrequent compared to the global proteome, making them challenging to maintain ex vivo and detect reproducibly. In our lab, we have optimized an approach for oxidized resin assisted capture (OxRAC) that enables thorough, reproducible analysis of the full reversible cysteine redoxome. By employing iodoacetamide as a blocking reagent during cell lysis, we ensure that cysteines reduced in vivo are excluded before non-specifically reducing reversibly oxidized cysteines for enrichment and analysis via label free quantitative LC-MS/MS. In this work, our optimized OxRAC platform was implemented to probe the dependence of heterotrimeric G-proteins on reversible cysteine oxidation for the mediation of abscisic acid (ABA) signaling pathways in Arabidopsis thaliana. By cross comparing the wildtype Col0 plants with the G-beta protein (AGB1) null mutant, agb1, we were able to quantify 6,891 unique oxidized cysteine-containing peptides and reveal 923 significant changes in oxidation following ABA treatment. Divergent pathways, including primary metabolism, ROS response, translation, and photosynthesis, exhibited both ABA and G-protein dependent redox changes, many of which occurred on proteins not previously linked to either ABA or G-proteins. Together, these data uncover a complex network of reversible oxidations that allow ABA and G-proteins to rapidly adjust cellular signaling to adapt to changing environments and suggest that a functional G-protein complex is required to maintain intracellular redox homeostasis and fully execute plant stress responses.

Contact: Amanda Smythers, asmyther@email.unc.edu

P12.07 High-throughput protein modification quantitation analysis using intact protein MRM and its application on hENGase inhibitor screening

Dingyin Tao (1), Miao Xu (1), Atena Farkhondeh (1), Andrew Burns (1), Steven Rodems (2), Matthew Might (3), Wei Zheng (1), Christopher LeClair (4)
(1) National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland, United States
(2) Travere Therapeutics, San Diego, California, United States
(3) University of Alabama at Birmingham, Birmingham, Alabama, United States
(4) National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland

The emerging demand for protein quantitation has been increasing in multiple fields, which includes protein biomarker screening and validation, protein therapeutics (antibody and antibody-drug conjugation), and drug screening and development. Currently, there remains a large gap for performing high-throughput protein quantitation at the intact protein level using label-free methods. In our effort to develop such a high-throughput protocol, we selected ribonuclease B (RNB),

the substrate for human endo-β-N-acetylglucosaminidase (hENGase), as a model as it is a promising drug target for the treatment of N-Glycanase deficiency. Intact protein level multiple reaction monitoring (MRM) methods were initially developed and optimized to quantify RNB and deglycosylated RNB (RNB-deg) with an observed S/N ratio improvement of nearly 20-fold compared to traditional full MS scan methods. The protein MRM methods were utilized on the RapidFire-MS/MS system as a means of increasing throughput to enable hENGase inhibitor screens where we were able to achieve a 12-fold throughput improvement. The assay was further optimized into a 384-well plate format for compound screening purposes with S/B ratio >37-fold and Z' factor >0.7. This is suitable for high-throughput screening of compound collections at a speed of 2 hours per 384-well plate and equating to >3,000 compounds per day at a single concentration dose. This 384-well plate based automated SPE-MS/MS assay is efficient and robust for compound screening and the assay format has a wide applicability to protein targets for other disease models.

Contact: Dingyin Tao, dingyin.tao@nih.gov

P13: Protein Complexes and Interactomics

P13.01 Reconstructing the mammalian brain protein interactome using high-throughput proteomics

Vy Dang (1), Ophelia Papoulas (1), Claire McWhite (2), Edward Marcotte (1)
(1) University of Texas at Austin, Austin, Texas, United States
(2) Princeton University, Princeton, New Jersey, United States

Understanding the complexity of the mammalian brain has been a long-standing challenge, not least because of the many specialized proteins and assemblies required for proper neuron function and connectivity. In particular, protein-protein interactions (PPIs) are critical to the proper functioning of neural proteins, and neurons from different mammals share many stable macromolecular protein complexes that perform conserved cellular functions. However, such data are only partially known due to the challenge of defining PPIs at high throughput directly from mammalian brains. In this project, we determined neuronal PPIs and “protein neighborhoods” at large scale by applying a tag-less proteomics technique called co-fractionation/mass spectrometry (CF-MS) to primary brain tissues of several mammals, including humans. CF-MS involves non-denaturing extraction of endogenous proteins in their native assembly states, followed by chromatographic separation into biochemical fractions, then mass spectrometry to define co-eluting proteins. Given sufficient independent separations and the appropriate statistical framework, stable protein complexes can be reliably identified based on protein co-elution patterns. We will describe our progress in combining this integrative proteomic approach with available brain PPI data and information from sequence orthology across mammals in order to define a large-scale map of conserved protein complexes in the mammalian brain.

Contact: Vy Dang, vyqtdang@utexas.edu

P13.02 System Wide Profiling of Protein Interaction Dynamics Links Host Innate Immunity and DNA Damage Responses

Joshua L. Justice (1), Josiah E. Hutton (1), Michelle A. Kennedy (1), Brett Phelan (1), Bokai Song (1), Dawei Liu (1), Ileana M. Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Protein-protein interactions (PPIs) underly the cellular capacity to sense viral infection and subsequently mount an immune response. Viruses have acquired mechanisms to counteract cellular immunity by suppressing, subverting, or hijacking the very interactions that drive the antiviral program. This interplay between infection and host defense occurs at multiple sub-cellular arenas through at the interface of dynamically shifting and interchanging virus-virus, virus-host, and host-host PPIs—the scope of which has not yet been fully captured. Using Thermal Proteome Profiling coupled to Mass Spectrometry (TPP-MS), we constructed a cell-wide portrait of the PPI dynamics that coordinate the innate immune response to Herpes Simplex Virus 1 (HSV-1) infection in primary fibroblasts. We found that IFI16, a nuclear DNA sensor that serves as a central platform for HSV-1 immune responses, coordinates with the DNA damage response (DDR) by activating the master DDR kinase, DNA-PK. Microscopy and molecular virology assays revealed that IFI16 recruits DNA-PK to the genome at the nuclear periphery. DNA-PK was necessary to suppress viral spread and mount an inflammatory antiviral response. The discovery of DNA-PK as a major contributor to the nuclear immune response further underscored a need to characterize its kinase-substrate network. Coupling phospho-peptide enrichment with DNA-PK chemical inhibition revealed hundreds of DNA-PK dependent modifications after both viral infection and DNA damage. Integrating the phospho-peptide and TPP-MS

analyses revealed extensive DNA-PK driven modification of the cellular and viral proteomes, uncovering the reciprocal regulation of IFI16 by DNA-PK phosphorylation at IFI16 residue T149. Mutation of this phospho-site showed that DNA-PK modifies IFI16 to promote type-1 interferon secretion while suppressing the expression of other inflammatory cytokines. Altogether, our study not only represents the first cell-wide characterization of PPIs during HSV-1 infection, but also uncovers a missing link in the immune signaling pathway that places IFI16 and DNA-PK central to herpesvirus innate immunity.

Contact: Joshua Justice, jjustice@princeton.edu

P13.03 Thermal proximity coaggregation mass spectrometry reveals shared and distinct rewiring of interactomes in alpha, beta, and gamma herpesvirus infection

Tavis Reed (1), Olga Troyanskaya (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Determining protein-protein interactions (PPIs) and their dynamics is key to uncovering the function and regulation of proteins in both homeostatic and perturbed states. Viruses rely on temporal remodeling of PPIs, including virus-virus, virus-host, and host-host PPIs, in order to successfully replicate within host cells. Defining dynamic interactomes at a system level is key to understanding the biology of viral infections and identifying possible targets for therapeutic intervention. An unbiased method for global profiling of dynamic changes in interactions is Thermal Proximity Coaggregation (TPCA) mass spectrometry. TPCA relies on the principle that interacting proteins co-aggregate under temperature-induced stress, resulting in overlapping aggregation curves. Analysis of the distance between curves, their melting temperatures, and their clustering are used to predict PPIs. Previously, our lab has leveraged TPCA to study the nuclear-replicating human cytomegalovirus (HCMV; a beta-herpesvirus) and herpes simplex virus type I (HSV-1; an alpha-herpesvirus). Here, we 1) optimize the TPCA sample preparation and 2) apply this to a model gamma-herpesvirus, Kaposi's Sarcoma-associated herpesvirus (KSHV). First, we performed comparative experimental and computational analysis of a multitude of different cell lysis conditions and set of denaturation temperatures, identifying a workflow that improves the coverage of proteins from different functional and subcellular localization classes and the prediction of PPIs. Next, we used this workflow to perform TPCA during the reactivation of KSHV in a latency model. Integration of the KSHV dataset with our previous HCMV and HSV-1 TPCA datasets revealed conserved PPI rewiring events involving both host-host and viral-host PPIs. We show that known homologs of viral proteins share similar interactions and temporal interaction dependence, highlighting the promise of these comparative analyses to reveal functions and properties of poorly characterized viral proteins. Altogether, this work helps to guide TPCA experimental designs and reveals temporal virus-host interactions critical for the replication of these ancient viruses.

Contact: Tavis Reed, tjreed@princeton.edu

P13.04 The protein organization of a red blood cell

Wisath Sae-Lee (1), Caitlyn L. McCafferty (1), Eric J. Verbeke (1), Pierre C. Havugimana (2), Ophelia Papoulas (1), Claire D. McWhite (3), John R. Houser (1), Kim Vanuytsel (4), George J. Murphy (2), Kevin Drew (5), Andrew Emili (2), David W. Taylor (1), Edward M. Marcotte (1)
(1) University of Texas at Austin, Austin, Texas, United States
(2) Boston University, Boston, Massachusetts, United States

(3) University of Texas at Austin, Austin, Austin, Texas, United States
(4) Boston University, Boston, Massachusetts
(5) university of Illinois at chicago, Chicago, Illinois

Erythrocytes (red blood cells; RBCs) are the simplest primary human cells, lacking nuclei and all major organelles. Despite their simplicity, RBCs dynamically change cellular morphology and physiology in response to pathophysiological cues, mediated in the absence of gene expression. While hemoglobin accounts for 98% of expressed RBC proteins, the full proteome is predicted to include >1,000 distinct proteins of which the roles and those of complexes they form remain elusive. In this study, we first defined a canonical RBC proteome of 1,202 proteins (1% FDR) using machine learning and data from quantitative mass spectrometry and RNA-seq on RBCs and other blood cell types. We then determined the soluble stable protein complexes in mature RBCs, based on mass spectrometry of 1,944 native biochemical fractions of hemoglobin-depleted hemolysate and detergent solubilized membrane protein complexes. Our data reveal an RBC interactome dominated by protein homeostasis, redox biology, cytoskeletal dynamics, and carbon metabolism, and give insights into the biophysical organization of integral membrane proteins, channels, cytoskeletal proteins, and metabolic enzymes at the molecular level. We validated protein complexes through electron microscopy and chemical crosslinking, and with these data, built 3D structural models of the ankyrin/Band 3/Band 4.2 complex that bridges the spectrin cytoskeleton to the RBC membrane to confer the characteristic RBC cell shape and morphological flexibility. The model provides detailed interactions at the amino acid level for the first time and suggests spring-link behavior of ankyrin in linking the cytoskeleton to the membrane. Taken together, our study provides a comprehensive view of the protein organization of the simplest primary human cell and serves as a comprehensive resource for future research.

Contact: Wisath Sae-Lee, mwsaelee@utexas.edu

P13.05 Temporally resolved protein-protein interactions to guide therapeutic intervention in protein misfolding diseases

Madison Taylor Wright (1), Bibek Timalsina (1), Valeria Garcia Lopez (1), Lars Plate (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

Cells are under pressure to properly fold, assemble, and maintain the integrity of thousands of proteins throughout changes in cellular conditions and stress. The proteostasis network (PN), consisting of folding, trafficking, and degradation components, helps facilitate the proper processing, trafficking, localization, and secretion of these proteins in a highly coordinated process referred to as protein quality control (PQC). Protein misfolding diseases such as thyroglobulin-related congenital hypothyroidism (CH) are a result of improper partitioning of proteins between PQC pathways. Coordination of PN-client interactions are critical for mediating proper PQC decisions. Recently, the PN has become an attractive target for therapeutic intervention to combat such disease states. Yet, one of the bottlenecks in the implementation of such strategies is a lack of understanding in the timing and coordination of PN-client interactions as there are no methodologies available to identify and quantify transient protein-protein interactions with time-resolution at an organelle-wide scale. We are therefore developing a quantitative mass spectrometry method, time-resolved interactome profiling (TRIP), to characterize the temporal interactions between client proteins and the PN to better understand PQC processes. TRIP is achieved with a two-stage enrichment strategy which utilizes unnatural amino acid incorporation through biorthogonal noncanonical amino acid tagging (BONCAT) and streptavidin-biotin enrichment to label and subsequently purify nascent client proteins into time-resolved fractions for affinity purification – mass spectrometry (AP-MS) analysis. We have coupled this methodology with functional genomic screening to identify key PQC regulators that control folding and secretion of WT thyroglobulin and CH-

associated mutants. Ultimately, our results may reveal PQC mechanisms available for therapeutic targeting and provide insight into the coordination of the PN that have broader applicability in other protein misfolding diseases.

Contact: Madison Wright, madison.t.wright@vanderbilt.edu

P14: Proteomics in Immunology and Infectious Diseases

P14.01 Disrupting Host-Pathogen Interactions Using Host-Targeted Protein Folding Regulators

Katherine Almasy (1), Jonathan Davies (1), Lars Plate (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

The COVID-19 pandemic has highlighted the importance of rapidly responding to emerging pathogens. The use of Remdesivir, initially investigated as a broad-acting RNA virus therapeutic, as a treatment for SARS-CoV-2 shows the advantage of having existing antiviral drugs which may potentially be used against future outbreaks. However, many viruses that pose a large global health burden have no current therapeutic options available. Among these are the flaviviruses, which include dengue, Zika, and yellow fever virus. Direct-acting antiviral compounds developed against these pathogens have been examined, but not successfully translated into therapeutics due to rapid emergence of resistance in the viral genome. An alternative option for broad-spectrum antiviral agents is to target host processes required for the viral life cycle and disrupt macromolecular interactions to block viral replication. To this end, we showed that two proteostasis regulator compounds (147 and 263) can act as antivirals against multiple strains of dengue and Zika virus. These compounds were initially developed to enhance host protein folding, but we discovered that they also potently inhibit flavivirus infection. Mechanistic investigations show 147 works in a thiol-dependent manner, and we use chemoproteomic approaches to identify endoplasmic-reticulum (ER) resident chaperone as the predominant protein targets of this molecule. Despite compound 147 and 263 possessing similar chemical structure, mechanistic studies and target ID approaches show the modes of action of the two molecules likely diverge. Both molecules are effective against a wide panel of viruses, providing a potential new avenue for host-centered therapeutics by targeting conserved ER proteostasis processes to act as broad-spectrum antivirals.

Contact: Katherine Almasy, katherine.m.almasy@vanderbilt.edu

P14.02 Applying proteomics to address a conundrum in virology: how mitochondrial fragmentation can be linked to increased bioenergetics

Cora Betsinger (1), Connor Jankowski (1), William Hofstadter (1), Joel Federspiel (1), Clayton Otter (1), Pierre Jean Beltran (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Alterations to mitochondrial functions and cellular metabolism are hallmarks of nearly all viral infections. As obligate intracellular parasites, viruses rely on mitochondria for the production of biosynthetic precursors and energy necessary for generating new viral particles. However, how the widely-spread human cytomegalovirus (HCMV) remodels mitochondrial structure and function had presented a conundrum. It remained unclear how HCMV increases mitochondrial bioenergetics despite triggering mitochondrial fragmentation. Employing a multidisciplinary approach integrating proteome and interactome datasets with super-resolution confocal microscopy, LC-MS based metabolite profiling and metabolic assays, we identify a previously uncharacterized viral protein, pUL13, which targets the mitochondria and increases oxidative phosphorylation during infection. We use targeted mass spectrometry analysis of the HCMV proteome during infection, coupled with molecular virology techniques to establish that pUL13 is required for productive HCMV replication. We then quantify temporal cellular proteome changes during infection and demonstrate that pUL13 alters electron transport chain (ETC) protein abundances. Using LC-MS based metabolite profiling and live-cell Seahorse metabolic assays to monitor

cellular respiration, we establish pUL13 as necessary and sufficient to increase cellular respiration, not requiring the presence of other viral proteins. To mechanistically define the function of pUL13 in regulating cellular respiration, we characterize the spatio-temporal pUL13 functional interaction network during infection. We discover and validate that pUL13 targets the MICOS complex, a critical regulator of mitochondrial architecture and ETC function. We then use stimulated emission depletion (STED) super-resolution microscopy analysis to visualize the impact of pUL13 on mitochondrial ultrastructure. Our findings address the outstanding question of how HCMV modulates mitochondria to increase bioenergetic output and expands the knowledge of the intricate connection between mitochondrial architecture and ETC function. Importantly, this is the first known instance of a virus protein targeting the MICOS complex to increase bioenergetic output, highlighting a mechanism that other virus pathogens might also possess.

Contact: Cora Betsinger, cnbetsinger7@gmail.com

P14.03 Membrane protein composition of S. pneumoniae extracellular vesicles provide insights into extracellular vesicle formation and shedding

Stephanie Biedka (1), Rory Eutsey (2), Luisa Hiller (2), Amber Lucas (1)
(1) Impact Proteomics, Pittsburgh, Pennsylvania, United States
(2) Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

Extracellular vesicles (EVs) are complex, cell-derived nanoparticles generated by all cell types. EVs are composed of lipid bilayer membranes and their associated membrane proteins, nucleic acids, and luminal proteins. The mechanism by which Gram-positive bacteria shed EVs is still unknown. EVs from the Gram-positive human pathogen S. pneumoniae, which is a major cause of otitis medi and pneumonia, are of particular interest because of how they EVs modulate the host immune response. To uncover possible mechanisms for EV production and shedding in S. pneumoniae, we have performed a comparative proteomics analysis of EV membrane proteins versus whole-cell membrane proteins. Membrane proteins were enriched from intact S. pneumoniae cells or their EVs using a ProMTag labeling and capture workflow. ProMTag is a bifunctional protein tag where one moiety of the tag is able to form a reversible, covalent link to primary amines on proteins. The other moiety is methyltetrazine, which can form an irreversible, covalent bond with trans-Cyclooctene (TCO) on the surface of beads to capture ProMTagged proteins for cleanup and elution. Using this workflow plasma membrane proteins can be tagged, captured, washed to remove non-plasma membrane proteins, and then eluted in their original, unmodified state. In this study, intact cells and EVs from S. pneumoniae cultures were separated and the extracellular domains of membrane proteins in these two fractions were labeled with ProMTag. The membrane proteins were then enriched, washed, and eluted using the ProMTag workflow. These membrane protein populations were then TMT labeled and analyzed using mass spectrometry. Comparative analysis revealed membrane proteins that are concentrated or absent in EV membranes relative to bulk plasma membrane from whole cells, indicating a selective process for EV formation in S. pneumoniae. With this information, we gained new insights into extracellular vesicle formation and shedding.

Contact: Stephanie Biedka, stephanie.biedka@impactproteomics.com

---

P14.04 Human Viruses Rewire Membrane Contact Sites for Pro-Viral Organelle Remodeling

Katelyn Cook (1), Elene Tsopurashvili (1), Jason Needham (2), Sunnie Thompson (2), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States
(2) University of Alabama Birmingham, Birmingham, Alabama, United States

Membrane contact sites (MCSs) link organelles in dynamic intracellular networks that coordinate cellular functions across space and time. All organelles use MCSs, which use protein tethers to facilitate biomolecule transfer at nanometer scales to finely-tune subcellular organization. Although organelle contacts have emerged as critical regulators of cellular homeostasis, translating MCS biology to disease contexts has remained challenging. This is largely due to a lack of methodologies able to capture the spatial, temporal, and functional facets of MCSs at a whole-cell scale. Here, we design a targeted mass spectrometry platform using parallel reaction monitoring for detecting and quantifying MCS proteins at all organelles simultaneously (MCS-PRM). We then apply MCS-PRM to human infections, defining functional virus-driven MCS alterations by both ancient and rapidly evolving human viruses: human cytomegalovirus (HCMV), herpes simplex virus type 1 (HSV-1), influenza A (Infl. A), and the beta-coronavirus HCoV-OC43. We find that viral regulation of organelle contacts is time-sensitive and organelle-specific, a molecular basis for switching anti- to pro-viral organelle structures and functions. Integration with live super-resolution microscopy and molecular virology reveals a previously unreported mitochondria-ER contact structure induced by infection, stable membrane encapsulations that we term MENC. We also determine that premature ER-mitochondria tethering activates STING and elevates interferon secretion, priming cells against infection. During HCMV infection, MCS-PRM coupled to proximity ligation assays and 3D image analysis demonstrate that ER contacts are increased and enriched at infection-derived enlarged peroxisomes, which plaster along ER membranes and require the ER-peroxisome tether ACBD5 to form. We find that ACBD5 abundance controls the balance between pro-viral peroxisome biogenesis and membrane expansion, exhibiting a broadly anti-viral role in each infection examined. By characterizing the global virus-directed modulation of MCSs, our study defines the molecular fingerprints of organelle remodeling linked to infection progression and pathogenesis.

Contact: Katelyn Cook, katelync@princeton.edu

P14.05 Viral Remodeling of the Host Proteome and Host Factors Associated with Viral Genomes

Joseph Dybas (1), Krystal Lum (1), Katarzyna Kulej (1), Emigdio Reyes (1), Richard Lauman (2), Matthew Charman (1), Caitlin Purman (1), Robert Steinbock (3), Nicholas Grams (2), Alexander Price (4), Lydia Mendoza (2), Benjamin Garcia (2), Matthew Weitzman (2)
(1) Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States
(2) University of Pennsylvania, Philadelphia, Pennsylvania, United States
(3) University of Pennsylvania, Philadelphia, Pennsylvania
(4) Children's Hospital of Philadelphia, Philadelphia, Pennsylvania

Interactions between viral and host proteins determine the ultimate outcome of virus infection. Cellular antiviral proteins present the first line of host defense, while viral proteins act to establish an environment conducive to infection. Viruses remodel the cellular proteome by altering protein abundance, rewiring protein-protein or protein-nucleic acid interactions, and redirecting protein post-translational modifications. Global proteomics provides an ideal approach to investigate virus-host interactions within the infected host cell. Adenovirus (AdV) is a ubiquitous human pathogen. AdV is known to induce degradation of host antiviral factors during infection and is an important model system to study viral manipulation of the host proteome. However, the full repertoire of virus-host interactions, host antiviral factors, and viral countermeasures are not fully understood. We describe an integrated proteomics strategy designed to define host proteins that target replicating AdV viral genomes, and quantify AdV-induced manipulations to the host proteome. We used Isolation of Proteins on Nascent DNA (iPOND) combined with mass spectrometry, to quantify the proteome associated with AdV genomes. We compared wildtype AdV to a mutant lacking viral proteins known to counteract host antiviral factors. Our rationale was that cellular proteins enriched on viral genomes, in the absence of viral countermeasures, may include host antiviral factors. We used

whole cell proteomics (WCP) to quantify changes in protein abundances during AdV infection and uncovered a subset of newly identified cellular degradation targets. We integrated iPOND and WCP datasets to define host proteins and pathways manipulated by AdV. We uncovered novel AdV targets, including SWI/SNF-related proteins (SMARCA6 and SMARCAD1) and SMC proteins (SMC5 and SMC6), and showed that these proteins have inhibitory impacts on viral infection. Our integrative, systems-level proteomics analysis reveals novel host targets of AdV. Elucidating how the host proteome is remodeled during infection reveals novel antiviral proteins and identifies viral countermeasures.

Contact: Joseph Dybas, dybasj@chop.edu

P14.06 Direct antigen epitopes mapping in human serum samples against AVA vaccine-elicited polyclonal antibodies

Mulin Fang (1), Kenneth Smith (2), Zhe Wang (3), Kellye Cupp-Sutton (3), Kathleen Haley (4), Judith James (4), Si Wu (1)
(1) University of Oklahoma, Norman, Oklahoma, United States
(2) Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, United States
(3) University of Oklahoma, Norman, United States
(4) Oklahoma Medical Research Foundation, Oklahoma City, United States

Anthrax vaccine adsorbed (AVA) containing protective antigen (PA) is the only FDA-licensed anthrax vaccine in the US. Characterization of the binding of AVA-induced anti-PA polyclonal human antibodies to the PA antigen after vaccination is crucial to understanding the mechanisms of AVA-elicited humoral immune response as well as to help refine the vaccine for protection against anthrax. However, techniques that measure conformational epitope binding in complex mixtures of antibodies (e.g., serum) remain elusive. We have recently developed and optimized a long-gradient (e.g., 90 minutes) sub-zero temperature ultra-high performance liquid chromatography hydrogen deuterium exchange mass spectrometry (UPLC-HDX-MS) platform for the analysis of protein interactions in complex biological samples including human plasma or sera. Here, we applied this novel platform to characterize PA epitopes against 4 fully human monoclonal antibodies. We found that p1C03 demonstrates partial neutralizing activity in vivo and binds to a flexible loop region on domain 1A of PA. p6C04 and p1A06 bind to the same epitope in domain 3 in a manner that prevents oligomerization but is not neutralizing. p6C01 binds to an epitope in domain 3 in adjacent helices and is highly neutralizing. The epitopes we identified to be associated with neutralizing performance will contribute to the successful development of novel therapeutics and subunit vaccines against anthrax. The performance of the UPLC-HDX-MS platform was further evaluated for direct epitope mapping in total IgG pooled from healthy controls without AVA vaccination (spiked with 2% p1C03 and p6C01) and total IgG pooled from AVA-vaccinated donors. The epitopes against p1C03 and p6C01 were identified in the control sample and two protected regions were identified in the vaccinated sample. Our results demonstrated the feasibility of the application of our novel platform for direct mapping of antigen epitopes in vaccine-elicited serum samples.

Contact: Mulin Fang, Mulin.Fang-1@ou.edu

P14.07 Peroxisomes are remodeled into dual structural populations to promote pro-viral cellular reprogramming during herpesvirus infection

William Hofstadter (1), Katelyn Cook (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

Organelle functions are tightly linked to the regulation of their protein composition, shape and subcellular localization. This is well-illustrated by dynamic changes to peroxisomes during virus infections, whereby virus-host control of peroxisome biogenesis and structure has emerged as a critical determinant of the metabolism-immunity axis required for infection progression. We recently showed that human cytomegalovirus (HCMV)—a herpesvirus that infects >70% of the world population and a leading agent of birth defects—increases peroxisome numbers and remodels peroxisomes into distinct fragmented and enlarged populations. These changes are accompanied by depressed anti-viral signaling and upregulation of plasmalogen lipid synthesis for virus assembly. Despite the clinical importance of understanding HCMV-induced peroxisome alterations, the molecular underpinnings of peroxisome remodeling remain largely unknown. Here, we integrate interaction-based and targeted mass spectrometry techniques with microscopy and molecular virology to define distinct viral strategies of peroxisome proliferation, enlargement, and fragmentation. We find that Drp1 and pUL37, host and viral fission factors, respectively, exhibit preferential localization to fragmented peroxisomes during HCMV infection. We found that pUL37 interacts with peroxisomal fission proteins upstream of Drp1, PEX11β and MFF. Further genetic manipulations reveal that both Drp1 and pUL37 are required for HCMV-driven peroxisome proliferation and establishment of the fragmented peroxisome population. HCMV infection concurrently produces an enlarged peroxisome fraction, at which we find preferential accumulation of the ER-peroxisome tethering protein ACBD5. We establish that ACBD5 is required for peroxisome growth during infection and that its overexpression suppresses pro-viral peroxisome proliferation. Moreover, both Drp1 and ACBD5 are necessary for efficient HCMV replication, highlighting the importance of regulating peroxisome dynamics for the progression of infection. Altogether, these data indicate that peroxisome subsets develop distinct proteomes during HCMV infection, leading to the disparate effects on peroxisome metabolism and immune signaling required for virus replication and spread.

Contact: William Hofstadter, wah2@princeton.edu

P14.08 Dynamic nuclear organization and protein translocation underlie nuclear viral DNA sensing and intrinsic immune signaling

Dawei Liu (1), Krystal Lum (1), Joshua Justice (1), Peter Metzger (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

In mammalian cells, DNA sensors are critical players in innate immunity, serving to detect the presence of pathogenic DNA and activate innate immune responses. The interferon inducible protein 16 (IFI16) has emerged as a key sensor of foreign DNA in the nucleus. Its function is especially critical in the defense against nuclear-replicating viruses, such as herpes simplex virus 1 (HSV-1). During viral infection, IFI16 quickly localizes to the nuclear periphery, where it binds to incoming viral DNA and oligomerizes. Upon viral DNA binding, IFI16 inhibits virus replication by inducing cytokine expression and inhibiting viral gene expressions. However, little is known about how IFI16 transmit its downstream signals to activate intrinsic and innate immune responses upon binding to viral DNA. Here, we first investigated the reorganization of nuclear composition early in infection. Using biochemical fractionation and quantitative mass spectrometry, we defined the dynamic proteomes of the nuclear periphery compared to the nuclear core in uninfected and infected primary fibroblasts. This provided insights into which host factors are specifically recruited to the nuclear periphery in response to the deposition of viral genome. Next, to identify IFI16-dependent protein complex formation changes during infection, we performed thermal proximity coaggregation (TPCA) mass spectrometry analyses in control and IFI16 CRISPR-knockout cells. Finally, we investigated whether, upon binding to viral DNA, IFI16 induces nuclear-cytoplasmic shuttling of yet unknown host factors to communicate with the ER-localized STING adaptor protein. We infected cells with wild type HSV-1 or a strain that lacks the ability to inhibit IFI16, and performed nuclear-cytoplasmic fractionation and tandem mass tagging (TMT)-MS to predict

alterations in translocation events. Cross-referencing the nuclear periphery interactions and proteome datasets, we identified dynamic IFI16 associations with kinases at the nuclear periphery. Through targeted MS, mutagenesis, and virology assays, we discovered a role for IFI16 phosphorylation in intrinsic immune response.

Contact: Dawei Liu, daweiliu@princeton.edu

P14.09 Architecture of Mammalian Innate Immunity: How Interferon Classes Tune the Protein Interactions and Abundances of Interferon-Stimulated Genes

Krystal Lum (1), Ileana Cristea (1)
(1) Princeton University, Princeton, New Jersey, United States

The mammalian innate immune system has been shaped by the coevolution of host-pathogen interactions over millions of years. To control the replication and spread of infection by a vast diversity of pathogens including viruses, the innate immune response has evolved to converge on a core cellular activity: the expression of several interferon (IFN) classes, which induce the transcription of hundreds of interferon-stimulated genes (ISG). While the significance of ISGs in stimulating adaptive immunity and promoting antiviral functions is well accepted, and mechanisms of action have been characterized for some ISGs, several fundamental aspects of this sophisticated host response remain unknown. Our knowledge of the cellular co-factors associating with ISGs in immune cells and how canonical IFN classes (Type I, II, and III) organize these ISG protein interaction networks are unclear. We also lack methods that reliably and specifically monitor ISG protein abundances en masse. Here, we adapt thermal proximity co-aggregation coupled to mass spectrometry to capture the system-wide dynamics of ISG protein complexes upon stimulation of immune cells with different IFN types. We identify de novo ISG-ISG interactions and protein assemblies that are commonly engaged across several IFN treatments, or specific to an IFN class. We further develop a targeted mass spectrometry assay based on parallel reaction monitoring to quantify protein abundances of ISGs representative of shared and distinct IFN classes throughout treatments with different IFN types. We observe that ISGs exhibit several distinct protein expression profiles, suggesting ISG subsets are temporally and differentially regulated during an innate immune response. Altogether, our assays and findings establish a cartographic protein map of mammalian innate immunity as it is shaped by canonical IFN classes. Our framework can be applied to any biological question involving ISGs, including assessing clinical samples and applications in viral infections and immune-related disease states such as autoimmunity.

Contact: Krystal Lum, klum@princeton.edu

P14.10 Proteomic characterization of human small airway epithelial cells and the effect of the BRD4-specific inhibitor ZL0454 on RSV infection by diaPASEF-MS

Morgan Mann (1), David Roberts (1), Ying Ge (1), Allan Brasier (1)
(1) UW - Madison, Madison, Wisconsin, United States

Respiratory Syncytial Virus (RSV) is a ubiquitous airway virus that causes severe inflammation and airway pathology in children and the elderly. RSV infects the epithelial cells of the upper and lower respiratory tract, resulting in toll-like receptor and NF-kB mediated innate inflammation. In the process, the NF-kB subunit RelA/p65 translocates into the nucleus and interacts with the epigenetic scaffold BRD4. The resulting complex traffics to NF-kB early-intermediate genes to trigger the coordinated expression of pro-inflammatory chemokines and interferons to signal nearby immune cells and form an early

antiviral defense. However, in severe infections, hyperstimulation and chronic inflammation can result in tissue damage and airway obstruction, often requiring hospitalization. Bromodomain (BRD) inhibitors have been shown to block virus-induced inflammation and fibrosis, and are a promising tool for addressing the clinical consequences of chronic inflammation. However, BRD inhibitors are often highly toxic and associated with poor clinical outcomes, and even next-generation BRD inhibitors can lead to unexpected and off-target effects. Accordingly, we applied high-throughput PASEF mass spectrometry to deeply interrogate the proteomes of human small airway epithelial cells infected with RSV and treated with our in-house BRD4-specific BRD inhibitors (e.g. ZL 0454). Samples were analyzed using a nanoElute LC system coupled to a timsTOF Pro mass spectrometer (Bruker Daltonik) utilizing the PASEF acquisition mode. MaxQuant was used to process the data and perform label-free quantitation. Samples will be further probed by Data-independent Analysis (dia) – PASEF using the FragPipe proteomics pipeline. At the conclusion of this study, we expect to observe significant reductions to virus-induced signaling proteins, corresponding to blockage of the BRD4/NF-kB signaling axis; alterations to other pathways are also expected. Together with previously obtained transcript-level data, this work will provide additional insights into the mechanisms of clinically relevant BRD4 inhibitors.

Contact: Morgan Mann, morgan.w.mann@gmail.com

P14.11 Ontogeny of the Human Thymus Proteome during Childhood

Anais MEZIANI (1), Simon Van Haren (2), Ceren Uncu (1), Meenakshi Jha (1), Kyle Higgins (1), Kenneth Parker (3), Christoph Schlaffner (4), Hanno Steen (5), Judith A Steen (1)
(1) F.M. Kirby Neurobiology Center, Department of Neurobiology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States
(2) Precision Vaccines Program, Division of Infectious Diseases, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States
(3) 1 F.M. Kirby Neurobiology Center, Department of Neurobiology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States
(4) Department of Pathology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States
(5) Department of Pathology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts

The thymus is an important site to the development of an immune-competent and self-tolerant T cell repertoire. Studies suggest that stromal cells in the thymus express a wide range of the human proteome. Cleaved peptide/epitopes of this proteome are displayed on these cells via MHC1 receptors. The thymus selects against T cell receptors that either bind too tightly or do not bind to these displayed epitopes. The thymus encounters profound anatomical changes throughout its lifespan, where organ atrophy begins at 1 year old and continues into adolescence. The decline of naïve T cell output was correlated with thymus involution. While Age-related morphological changes were broadly characterized, molecular and cellular mechanisms driving this regression are still elusive. Despite this key role in promoting immune system homeostasis, the thymus has not been studied using deep proteomics in humans. Here, we present a global proteomics investigation of fresh-frozen human thymi (n=73), obtained from patients undergoing corrective cardiac surgery. The ages ranged from 1 day to 15 years of life. Using the SP3 method coupled with the tims-TOF Pro we identified and quantified >8500 proteins. Our analysis revealed significant enrichment of proteins associated with positive regulation of T cell migration and proliferation in the thymi from newborns and infants until 4 months. In thymi from older children, we measure significant enrichment of proteins associated with tissue remodeling as well as progesterone and testosterone metabolic process. Additionally, we develop a workflow for proteomics analysis of 14 FACS sorted thymic cells populations derived from pooled human tissues. In summary, using the largest collection of fresh-frozen thymi, we are the first to characterize the human

thymus proteome and its global age-dependent changes during childhood. Our result represents a grounding dataset to better understand the dynamics of thymus involution in the context of tissue and cell-specific protein expression.

Contact: Anais MEZIANI, anais.meziani@childrens.harvard.edu

P14.12 Identifying prognostic signatures in the serum proteome of recuperated COVID-19 patients.

Smruti Pushalkar (1), Shaohuan Wu (1), Matthew Pressler (1), Justin Rendleman (1), Burcu Vitrinel (1), Elodie Ghedin (2), Ted Ross (3), Hyungwon Choi (4), Christine Vogel (1)
(1) New York University, New York, New York, United States
(2) National Institute of Health, Bethesda, Maryland, United States
(3) University of Georgia, Athens, Georgia, United States
(4) National University of Singapore, Singapore

The COVID-19 pandemic affected >260 million people worldwide with >5 million deaths. Presented by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the disease displays extensive clinical heterogeneity, ranging from asymptomatic to severe symptoms. While specific examples are known, e.g., elevated blood cytokine levels correlating with severe symptoms, the more general changes in the patient serum proteome are still under investigation. We used quantitative shotgun proteomics to characterize serum proteins of 29 recuperated COVID-19 patients and age-, race-, and gender-matched healthy controls. COVID-19 patients were 9 to 70 days post diagnosis and symptom-less at the time of sampling. We quantified ~330 proteins involved in inflammation, activation of complement pathways, blood coagulation and antimicrobial response. Of these, 22 and 15 serum proteins were significantly up- and down-regulated, respectively in COVID-19 patients compared to healthy controls (q-value<0.05). Recuperated COVID-19 patients had significantly elevated immunoglobulins for antigen recognition, Peroxiredoxin-2, hemoglobin subunits, Carbonic Anhydrase1, and Apolipoprotein A1. Interestingly, they also showed significant depletion of actin cytoskeleton signaling proteins, such as filamin, profilin and cofilin, suggesting possible links to hampered T cell signaling as well as manipulation of host actin cytoskeleton by the virus. Moreover, the COVID-19 infected individuals showed significant downregulation of thymosin ß, transgelin-2 and plasma serine protease inhibitor critical for lung functioning and enzymes that activate inflammatory cascades consistent with the skewed immune and organ function in COVID-19 patients. Comparing our results with other analyses, we identified intriguing similarities with acute COVID infections such as elevated hemolysis markers providing potential etiological insight into hemolytic complications, e.g., anemia or irreversible organ damage.>

Contact: Smruti Pushalkar, sp117@nyu.edu

P15: Proteomics of Diseases in Precision Medicine

P15.01 Parallel Integrated Analysis of Alzheimer Disease Cerebrospinal Fluid and Plasma Using 3 Proteomics Platforms for Biomarker Discovery and Validation

Eric B Dammer (1), Lingyan Ping (1), Duc M Duong (1), James J Lah (1), Allan I Levey (1), Nicholas T Seyfried (1), Erik CB Johnson (1)
(1) Emory University, Atlanta, Georgia, United States

Alzheimer's disease (AD) biofluid biomarkers reflecting brain pathophysiology beyond amyloid-β and tau (A/T) are needed. We analyzed the proteomes of paired blood plasma and cerebrospinal fluid (CSF) of the same AD and A/T biomarker-negative control subjects using three highly specific and sensitive quantitative assays to compare protein measurements across platforms and provide a consensus atlas of biofluid protein changes in AD that may serve as mechanistic, therapeutic, or diagnostic biomarkers. Blood plasma and cerebrospinal fluid from patients positive for A/T biomarker AD status with cognitive impairment (n=18) were assayed for quantitative protein differences relative to age-, sex-, and race-matched A/T biomarker-negative cognitively normal control subjects (n=18). Assays were discovery-mode mass spectrometry (MS), Olink antibody proximity extension assay (PEA) (>1100 proteins), and SomaScan aptamer-based fluorescent array quantitation (>6600 proteins). Biofluids were assayed without depletion on all three platforms, and also with immunodepletion of the 14 most abundant proteins prior to additional mass spectrometry. All samples were analyzed from equal biofluid volume. Proteomic changes were summarized using integrative analysis incorporating knowledge from brain tissue proteomic systems biology. Measurements of protein changes across the platforms were consistent overall and robust to depletion of the most abundant proteins in either CSF or plasma. Overall, more proteins were decreased in AD CSF and plasma. In AD CSF, increased proteins included neuron/synapse-derived proteins of the 14-3-3 family, neurogranin (NRGN), and neurofilament light polypeptide (NEFL). Metabolic enzymes were found increased or decreased. Neuronal marker protein neuronal pentraxin-2 (NPTX2), and vascular/blood signature proteins such as alpha-1-antichymotrypsin (SERPINA3), were found decreased. In AD plasma, many of the same neuronal marker proteins decreased that were increased in CSF, including 14-3-3 proteins and NRGN, with strong anticorrelation. Multi-platform CSF and plasma AD biomarkers reflect consistent measurements, biased towards decreased levels, and reflecting negative correlation of some markers across the two biofluids.

Contact: Eric Dammer, edammer@emory.edu

P15.02 Proteomic Analysis of Dorsal Root Ganglia of Type II Diabetic Donors with Peripheral Neuropathic Pain

Megan Doty (1), Yan Wang (1), Bradford Hall (1), Margaret Cassidy (1), Sijung Yun (2), Ashok Kulkarni (1)
(1) NIH/NIDCR, Bethesda, Maryland, United States
(2) Yotta Biomed, LLC., Potomac, Maryland, United States

Diabetic peripheral neuropathy (DPN) is characterized by spontaneous pain often in the extremities. Incidence of DPN is anticipated to rise with the global increase in diabetes. However, there remains a lack of safe, effective analgesics to control this painful condition. The dorsal root ganglia (DRG) contain soma of sensory neurons and modulate sensory signal transduction into the central nervous system. The aim of this study was to gain a deeper understanding of changes in molecular pathways in the DRG of DPN patients with chronic pain. Towards this goal, we used bottom-up proteomics approach to compare the proteome and phosphoproteome profiles of DRGs from donors with DPN and non-diabetic

controls. Pathway analysis was performed using proteins and phosphopeptides with significantly different abundance between the two groups. We found alterations in structural proteins and proteins involved in modulation of the immune system. Additionally, pathway analysis suggests interplay between differentially regulated proteins and phosphoproteins. These results provide new insights into the molecular mechanisms underlying DPN.

Contact: Megan Doty, megan.doty@nih.gov

P15.03 Urinary Protein Signatures for the Detection of Antiretroviral-Induced Acute Kidney Injury: A Retrospective Study

Ireshyn Govender (1), Previn Naicker (1), Neil Martinson (2), Stoyan Stoychev (1)
(1) Council for Scientific and Industrial Research, Pretoria, South Africa
(2) Perinatal HIV Research Unit, Johannesburg, South Africa

Introduction: Approximately one in five South Africans are HIV positive, which accounts for ~10% of the global HIV/AIDS burden. South Africa administers the largest antiretroviral therapy (ART) program worldwide. However, 8-10% of first-line ART patients experience acute kidney injury (AKI). Current AKI tests are unreliable and reflect following significant kidney damage. Hence there is a need to identify sensitive and accurate markers for early detection of AKI. This study employed mass spectrometry-based urinary proteomics for biomarker discovery in ART-AKI in two distinct South African cohorts. Methods: Two, case-control (age-race-gender) matched, cohorts (n=74, n=80) were analysed. A simple and robust sample collection procedure was coupled with reproducible, high-throughput, automated magnetic microparticle-based protein capture, clean-up, and digestion. Proteome profiles were generated via data independent acquisition using an UltiMate 3000 UHPLC coupled to a TripleTOF 6600. Raw data was processed using a study-specific spectral library in Spectronaut 14. Proteins common to both cohorts were filtered as reporters of ART-AKI and processed using Perseus (v1.6.12.0) and Cytoscape (v3.8.2). Machine learning (ML) was applied to protein abundance data using OmicLearn (v1.1.0) Results: Two-hundred and two common proteins (inclusion criteria: q-value ≤ 0.01, unique peptides ≥ 2) showed the same directional change in abundance. These indicated that damaged kidneys may undergo renal hepatisation showing increased levels of liver-associated proteins. Immune system, complement cascade and serine protease inhibitors, were the dominant protein families showing increased abundance in AKI patients. The XGBoost algorithm showed ≥ 95% sensitivity and specificity in classifying AKI and severity. Conclusions: The workflow applied here showed clinical robustness and can find utility as a routine workflow for urinary proteomics studies. ML offers promise for use with complex urinary proteomics datasets for clinically applicable outcomes. Validation of the identified protein biosignatures is currently on-going in a large cohort (n = 2000).

Contact: Ireshyn Govender, ireshyng@gmail.com

P15.04 Proteomics to understand effect of maternal diet during pregnancy on fetal adult life

Zeeshan Hamid (1), Cun Li (2), Peter Nathanielsz (2), Laura Cox (1), Michael Olivier (1)
(1) Wake Forest University School of Medicine, Winston Salem, North Carolina, United States
(2) Southwest National Primate Research Center, San Antonio, Texas, United States

Introduction: Maternal diet during pregnancy has an important impact on growth and development of the fetus. Epidemiological and animal studies have shown that quality of diet during pregnancy affects health of the child, with

hypercaloric diet combined with maternal obesity during pregnancy increasing the risk for various offspring disease conditions in adulthood, particularly metabolic and cardiovascular diseases. While several controlled experimental studies in rodents show effects of maternal diet during pregnancy on adult life, experimental studies in species like baboons that have a closer genetic and evolutionary relationship to humans are very limited. In this study, we compared the liver proteome from baboon offspring of mothers fed a chow diet or a Western high-fat diet during pregnancy. Methods: Our study included 32 liver biopsies from baboons aged 3-7 years (n=13 fed with normal chow diet and n=19 fed a maternal western diet during pregnancy). For proteomics analysis, samples were homogenized in Tris buffer followed by reduction, alkylation and trypsin digestion. Peptides were TMT labelled and pooled into four 10-plex TMT groups with a pooled reference channel. Resulting peptides were offline fractionated using high pH zip tips and separated on 75µm×15cm RSLC C18 easy spray column using a 2-hour gradient and analyzed on Fusion Lumos Orbitrap mass spectrometer. Data were analyzed using MaxQuant and normalization and post-processing were done using MSstats TMT workflow. Results and conclusions: We obtained quantitative data for 3736 proteins, of which 673 proteins were significantly different between two diet groups (p-value <0.05). Among other important pathways, DAVID pathway analysis shows strong enrichment of proteins involved in protein folding and regulation of mRNA splicing. Molecular functions specific to differential proteins include pathways specific to poly(A) RNA binding and nucleotide binding. Overall, our study shows a clear long-term impact of maternal western diet during pregnancy on offspring liver proteome.>

Contact: Zeeshan Hamid, zhamid@wakehealth.edu

P15.05 Divergent proteostasis interactions reveal distinct pathways attenuated by correctors in CFTR misfolding mutations

Minsoo Kim (1), Eli McDonald (1), Carleen Sabusap (1), Lars Plate (1)
(1) Vanderbilt University, Nashville, Tennessee, United States

Cystic fibrosis (CF) is one of the most prevalent lethal genetic diseases that has over 2000 identified genetic variants in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Corrector drugs such as lumacaftor (VX-809), tezacaftor (VX-661) and elexacaftor (VX-445) combat mutation-induced defects by stabilizing the structure of CFTR. The stabilization enables proper folding and thus facilitates processing and trafficking to increase the amount of functional CFTR on the cell surface. Here, we identified differential responses to correctors by CFTR variants and investigated the underlying cellular mechanisms of how CFTR biogenesis is altered for these variants. We employed affinity purification-mass spectrometry (AP-MS) multiplexed with isobaric tandem mass tags to define the remodeling of CFTR proteostasis interactions for multiple CFTR mutant variants in response to correctors. We identified several dysregulated pathways in the CFTR interactome in misfolding variants. We determined that correctors restore early proteostasis pathways such as translation and folding as well as degradation machinery. In particular, degradation fates diverged for mutants in which inhibiting the proteasome attenuated VX-809 response in responsive mutations but inhibiting the lysosome attenuated F508del mutant VX-809 response. Our data suggest a previously unidentified role for proteasomal degradation in VX-809 correction of CFTR. Moreover, VX-445 treated CFTR exhibited unique interactomics profiles compared to VX-809 treatment, indicating how critical pathways and proteostasis factors differ for individual correctors. We narrowed down several proteins as novel targets of interest for validation via specific pathway modulators and knockdowns. Our approach is unique in quantitatively identifying proteostasis characteristics associated with mutant-specific therapeutic response to correctors. Our results could provide a better understanding of the corrector mechanism of action and reveal novel cellular targets for therapeutic approaches.

Contact: Minsoo Kim, min.soo.kim.1@vanderbilt.edu

P15.06 Plasma Proteomic Changes in Response to a High Salt Diet in Female Baboons

Sumaiya Nazli (1), Angelica Riojas (1), Zeeshan Hamid (1), Kimberly D. Reeves (1), Laura A. Cox (1), Michael Olivier (1)
(1) Wake Forest School of Medicine, Winston-Salem, United States

Hypertension is a complex medical condition that affects 45% of U.S. adults, and the prevalence increases with age. The kidneys play a key role in regulating blood pressure, predominantly by regulating salt homeostasis. Non-human primates are an ideal animal model to study physiological responses to diet due to their similarity to humans. In our study, female baboons (n=8) were maintained on a low-sodium chow diet and fed on a high-sodium diet for six weeks. Blood samples were collected before and after the high-salt challenge. Blood pressure was monitored throughout the study using telemetry. We analyzed plasma by using mass spectrometry. For initial comparisons, we included proteins that were present in at least 50% of the samples at each time point (e.g., high vs low-sodium intake). We identified 180 proteins that overlapped between groups, while 43 proteins were unique to the samples collected after the high-sodium diet and only 3 proteins were unique to the samples collected on the low-sodium diet. We computed logistic regression with each protein to test for differential detection rates between groups. False Discovery Rate (FDR) was computed to adjust for multiple comparisons. No results met an FDR < 0.05, but 3 proteins met an unadjusted p-value < 0.05 (F13A1, KRT2, LAMP1). For proteins that had less than 30% missingness across groups, we computed linear regressions to test for differential abundance between groups. No proteins met an FDR < 0.05, but 17 proteins met an unadjusted p-value < 0.05. Many of these proteins were related to heme signaling pathways. The plasma proteomics differences between the two groups could play a key role in functional differences in the hypertensive group. This represents a promising avenue of investigation that may lead to more development of these proteins in relation to possible impairments in kidney function with regulation of hypertension.

Contact: Sumaiya Nazli, snazli@wakehealth.edu

P15.07 Laser Capture Microdissection Assisted Spatial Resolved Proteomics Analysis of FFPE Human Tissues for Fibrosis-Related Disease MOA Study

Yu Tian (1), Xue Wang (1), Jan Schejbal (1), Liang Jin (1), Baoliang Cui (1), Yupeng (David) He (2), Yonghao Cao (3), Annette Schwartz Sterman (4), Robert Dunstan (5), Chenqi Hu (6)
(1) DMPK-BA, Abbvie Bioresearch Center, Worcester, Massachusetts, United States
(2) Discovery Dermatology and Fibrosis, Abbvie, North Chicago, Illinois, United States
(3) Transformational and Translational Immunology Discovery, Abbvie Bioresearch Center, Worcester, Massachusetts, United States
(4) Immunology Pathology, Abbvie Bioresearch Center, Worcester, Massachusetts
(5) Immunology Pathology, Abbvie Bioresearch Center, Worcester, Massachusetts, United States
(6) DMPK-BA, Abbvie Bioresearch Center, Worcester, Massachusetts

Traditional omics with bulk sampling poorly defines disease heterogeneity and lack the sensitivity to identify true disease targets. Therefore, for diseases with a high degree of heterogeneity, like idiopathic pulmonary fibrosis (IPF), and hidradenitis suppurativa (HS), spatially resolved omics analysis is critical. laser capture microdissection (LCM) is a powerful technology for spatial omics analysis, which can achieve spatial resolution as well as morphology-directed area selection based on histo/path classification with strong biology information association. However, it's very challenging to execute downstream

omics analysis due to the limited quantity of tissue after high-resolution LCM area selection. To overcome these challenges, an LCM proteomics workflow was developed and optimized, including FFPE tissue sectioning, tissue harvesting by LCM, proteomics sample handling and preparation, LC/MS instrument settings, and proteomics database search strategy. A technical POC study was conducted using FFPE idiopathic pulmonary fibrosis samples. regions from 3 IPF FFPE blocks, 8 regions were collected by LCM based on 4 histologic features “Normal” (2 regions), fibrotic (3 regions), fibroblastic foci-rich (2 regions), and honeycombing (1 region). From these samples, ~2,900 proteins were identified and quantified. Fibroblastic foci-rich vs. “normal” had more differentially-expressed proteins compared to the fibrotic vs. “normal” regions. These findings may be related to the hypothesis that fibroblastic foci are a leading indicator of early IPF progression. In addition, the LCM proteomics platform was also applied in an HS technical pilot study to prepare for the multi-omics analysis of HS skin samples. 12 HS skin LCM samples were prepared designated with 4 distinct regions, HS epithelium, epidermis, inflammation, fibrosis. A total of ~1,300 proteins were identified and quantified. PCA analysis showed a clear separation of these four regions. Multiple differential expression analysis was performed between regions, with a panel of extracellular matrix-related targets identified with a relationship to fibrosis mechanism.

Contact: Yu Tian, yu.tian@abbvie.com

P15.08 Towards understanding the comprehensive extracellular matrix proteome in teeth

Yan Wang (1), Priyam Jani (1), Marian Young (1), Olivier Duverger (1), Janice Lee (1)
(1) National Institute of Craniofacial and Dental Research, National Institutes of Health, Bethesda, Maryland, United States

Background: The human tooth is a unique organ composed of mineralized tissues enamel and dentin, that have a prominent extracellular matrix (ECM). Several rare diseases affect teeth, but the proteomic outcome is unknown. Our goal was to establish new protocols for analysis of teeth proteome to provide a foundation to investigate diseased teeth. Methods: Wisdom teeth were extracted at the NIH-DIR Dental Clinic and processed for protein extraction using protocols established previously. Participants consented to clinical protocol (NCT01805869) before collecting teeth. In this work, we attempt to optimize conditions for storage as well as protein extraction and analysis. For this purpose, teeth from healthy volunteers were processed in separate conditions, including storage in RNALater vs Trizol, protein extraction using Guanidine vs urea, and benefit of multiple steps with demineralization and limited trypsin digestion. Tryptic peptides were analyzed using standard nanoLC-MS/MS workflow with data dependent analysis using orbitrap Fusion Lumos mass spectrometer. Proteome Discoverer 2.5 software was used for protein identification and quantification. Results: More than 2000 proteins were detected from the mass spec analysis. RNALater and Trizol both preserved similar number of proteins. Guanidine had higher proteins dissolved than urea, however, some proteins, especially the non-collagenous proteins dissolved more in Urea. While most of the proteins identified were from the initial Guanidine or urea extraction, the follow up extracts after demineralization released collagens and non-collagenous proteins which make up the bulk of the ECM composition of dentin in tooth. Conclusion: We tested different conditions to test the best protocol for extracting proteins from human teeth. RNALater could be an optimal solution for preserving proteins in teeth and other mineralized tissues. Guanidine and Urea each extracted higher number of specific proteins, so they should be used depending on the affected target proteins for downstream analysis.

Contact: Yan Wang, yan.wang2@nih.gov

P15.09 A Fully Automated High-Throughput, Deep-Scale Quantitative Plasma Proteomics Workflow Enables Quantitively Profile More Than 1000 Proteins Per Sample

Yu Zhou (1), Amirmansoor Hakimi (1), Yang Liu (1), David M. Horn (1), Kristan Bahten (2), Steven L. Reeber (1), Yue Xuan (1), Kevin L. Schauer (3), Runsheng Zheng (1), Oleksandr Boychenko (3), Y. Ruben Luo (4), Kara L. Lynch (5), Andreas F. Huhmer (1), Daniel Lopez-Ferrer (1)

(1) Thermofisher Scientific, San Jose, California, United States
(2) 1. ThermoFisher Scientific, San Jose, California, United States
(3) Thermofisher Scientific, San Jose, California
(4) Stanford University, Palo Alto, California, United States
(5) University of California San Francisco, San Francisco, California, United States

Introduction: LCMS-based plasma proteomics is advancing our understanding of human molecular pathophysiology and empowering the discovery of therapeutic targets and biomarkers. However, managing throughput, proteome depth and reliability altogether represent a major gap to fully enable meaningful large cohort proteomics studies. Methods: We developed an automated and high-throughput solution to enable large-scale proteomics analysis of proteins from 1 µL of plasma. We automated the sample preparation protocol using a commercially available 96 MS Sample Prep Kit and a liquid handling robotic platform. Digested peptides were separated using a Thermo Scientific™ Easy-Spray™ PepMap™ Neo column on a Thermo Scientific™ Vanquish™ Neo UHPLC system coupled to a Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer with a High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS™). Results: We standardized a plasma profiling workflow solution using the following building blocks: 1) A liquid handler for automated sample preparation; 2) A next-generation LC that enables higher robustness and peak capacity; 3) The unified integration with FAIMS Pro Interface coupled to a Thermo Scientific Orbitrap Exploris 480 mass spectrometer. 4) Furthermore, we fully benefited from a new search engine that uses deep neural networks to maximize the retrieval of protein identifications yielding significant improvement compared to conventional search engines. The integration of these technologies and devices allowed us the scalability to profile >1000 proteins from 1µL plasma samples without compromising reliability, robustness, and data quality. 85% of 1014 identified protein group has %CV of less than 20% in triplicates injections. 82 FDA-approved biomarkers can be identified, and 66 FDA-approved biomarkers can be reproducibly quantified (%CV<20%, n=3) with our high throughput label-free LCMS method. Conclusions: We developed a highly reproducible quantitative plasma proteomics workflow with data confidence to increase the statistical power of human large cohort plasma proteomics study.>

Contact: Yu Zhou, yu.zhou5@thermofisher.com

P16: Quantitative Proteomics and Systems Biology

P16.01 A Proteomics Perspective of Postnatal Swine Heart Development

Timothy Aballo (1), Elizabeth Bayne (1), Wuqiang Zhu (2), Meng Zhao (2), Ahmed Mahmoud (1), Jianyi Zhang (2), Ying Ge (1)
(1) University of Wisconsin-Madison, Madison, Wisconsin, United States
(2) University of Alabama at Birmingham, Birmingham, Alabama, United States

Purpose: The neonatal swine heart possesses an ability to regenerate injured myocardium through the proliferation of pre-existing cardiomyocytes (CMs), but this regenerative capacity is lost shortly after birth. The mechanisms governing the proliferative capacity of CMs during this early postnatal stage are unknown; therefore, there is great need to define the proteomic landscape during postnatal development to identify regulators of this regenerative process. Methods: Left ventricular (LV) tissue was isolated from swine at postnatal days (P) 1, 7, 28 and 56 (n=3). For top-down analysis, myofilament proteins were enriched using a differential pH extraction and analyzed by reversed phase liquid chromatography (RPLC) interfaced with a Q-TOF mass spectrometer. For bottom-up analysis, LV tissue was homogenized in Azo, and proteins were digested with trypsin prior to RPLC interfaced with the timsTOF Pro. Results: Top-down proteomics identified over 75 sarcomeric proteoforms and revealed a concerted transition of fetal sarcomeric proteoforms into their more mature counterparts. More specifically, we observed a marked decrease in ssTnI expression, cTnT2 expression, and α-Tpm phosphorylation throughout postnatal cardiac development. Global bottom-up proteomics quantified over 4,000 protein groups throughout the developmental window. We identified >900 differentially expressed proteins (adjusted p<0.05) and mapped these proteins to different biological processes, revealing a steep decrease in translational processes and a sharp increase in metabolic proteins in P56 hearts compared to P1 hearts. Conclusion: Using these methods, we comprehensively defined how the proteomic landscape of swine hearts changes throughout postnatal development. We identified significant alterations in sarcomere proteoform expression patterns, which may influence sarcomere disassembly, a process necessary for CM proliferation. Additionally, we mapped global changes in protein abundance to determine there was a significant increase in metabolic activity and a substantial decrease in processes related to translation. Overall, we will use these results to guide future investigations into swine heart regeneration.>

Contact: Timothy Aballo, aballo@wisc.edu

P16.02 A complete and automated sample preparation strategy for high throughput and standardized proteomics

Nicolai Bache (3), Dorte B. Bekker-Jensen (1), Florian S. Harking (2), Angela McArdle (1), Lasse Falkenby (1), Jakob Bunkenborg (1), Ole B. Hørning (1), Jesper V. Olsen (2)
(1) Evosep Biosystems, Odense, Denmark
(2) University of Copenhagen, Copenhagen, Denmark
(3) Evosep Biosystems, Odense

Purpose Mass spectrometry-based proteomics is fast growing and provides a powerful set of technologies, with the potential to revolutionize health care and enable precision medicine. We recently developed the Evosep One, which is specifically designed for high throughput applications, with a focus on clinical analysis of large sample cohorts. For a widespread adoption of proteomics in the clinic, the entire workflow from sample preparation through LC-MS and data

analysis needs to be fast and robust to enable the required throughput. To meet these requirements, we introduce an end-to-end workflow from protein lysate to peptides loaded on Evotips ready for injection on the Evosep One. Furthermore to facilitate high throughput analysis in the clinical laboratory, we have developed a new standardized method for the Evosep One allowing for 500 samples analyzed per day. Methods HeLa cells were cultured in DMEM media and harvested in boiling 5% sodium dodecyl sulfate (SDS) buffer. A complete and automated sample preparation workflow was integrated on an Opentrons OT-2 robot utilizing protein aggregation capture (PAC) on magnetic microparticles, followed by on-bead trypsin digestion, and automatic loading of the digest onto Evotips using a specially designed pneumatic 8-channel module from Evosep (prototype). The device has a footprint of a 96 well microtiter plate in order to be compatible with most liquid handling robots and uses positive air pressure to move liquid through the Evotip. Preliminary data We benchmarked the performance against an identical set of experiments performed manually. Peptide and protein identifications are similar in both sets of experiments as both sets also show good reproducibility, while the fully automated workflow requires limited hands-on time. Conclusion A fully automated end-to-end workflow from protein lysate to peptides loaded on Evotips ready for injection on the Evosep One ensuring efficient storage of peptides until analysis.

Contact: Nicolai Bache, nb@evosep.com

P16.04 Phosphoproteomic analysis reveals kinases involved in SARS-CoV-2 induced pancreatic β-cell death

Ran Cheng (1), Chien-Ting Wu (1), Peter V. Lidsky (2), Yinghong Xiao (2), Janos Demeter (1), Raul Andino (2), Peter K. Jackson (1)
(1) Stanford University, Stanford, California, United States
(2) University of California, San Francisco, San Francisco, California, United States

While many people infected with COVID-19 were fortunate enough to not succumb to the disease, hundreds of millions of people were still afflicted by COVID-19 in the past two years and it is imperative to examine possible long-term complications that might arise due to the infection. While the airway epithelium is the initial site of infection, SARS-CoV-2 can disseminate to infect other organs. The pancreas secretes hormones that control blood glucose levels and dysfunctional pancreas is often the cause of diabetes. Pancreatic islets express ACE2, the receptor for SARS-CoV-2 Spike protein, and is believed to be a target of SARS-CoV-2. Here we examine effects of SARS-CoV-2 infection of the pancreas and how it relates to diabetes through phosphoproteomic analysis. We were able to detect on average 20,000 unique peptides with above 85% phosphoenrichment. From this dataset, our KSEA was able to predict the kinase activities of 67 different kinases. Doing a gene ontology (GO) biological process over-representation analysis of kinases with predicted upregulated activities revealed cell death pathways as the most significantly enriched pathways 30 minutes post RBD treatment. This trend was also observed in pancreatic islet cells 24 hours post SARS-CoV-2 infection, indicating COVID-19 induces cell death. TUNEL staining, an indicator of cell death, confirmed an increase in cell death in β-cells infected with SARS-CoV-2. This process is similar to the cause of Type 1 diabetes where insulin deficiency due to autoimmune response related loss of β-cells, indicating COVID-19 patients are more prone to developing diabetes. Induction of the JNK-MAPK pathway in β-cells post SARS-CoV-2 infection was predicted using KSEA, suggesting that the JNK-MAPK pathway could be involved in inducing apoptosis in infected β-cells and a candidate for potential target for therapeutic intervention.

Contact: Ran Cheng, rancheng@stanford.edu

P16.05 Quantitative proteomics of zebrafish embryos reveals wave-like expression profiles of transcription factors during maternal to zygotic transition

Fei Fang (1), Daoyang Chen (1), Liangliang Sun (1)
(1) Michigan State University, East Lansing, United States

The transition from fertilized egg to embryo is accompanied by a multitude of changes in gene expression. Zebrafish is a widely used vertebrate model organism in embryological developmental biology due to its high genomic conservation with humans and rapid development. While the transcriptional events during the development were extensively studied on zebrafish, the data on the proteome profile are scarce, especially during embryogenesis. In the present study, we coupled iTRAQ labeling with both CZE-MS/MS and LC-MS/MS to profile protein expression levels of zebrafish embryos across four stages during the maternal-to-zygotic transition between fertilization and 6 hours post fertilization (hpf). Nearly 5,000 proteins were quantified with biological replicates. The clusters of protein expression profiles clearly indicated the expression level of nucleus proteins dramatically increased in pre-midblastula transition (MBT) and MBT stages, involving in the assembly of RNA Polymerase II Complex and protein transport. Among the 110 quantified transcription factors (TFs), 41 TFs have significant abundances changes across the four stages and generate 4 waves of zygotic transcription. The wave-like expression patterns showed that the TFs containing homeobox and zf-H2C2-2 domains playing dominant roles during MBT. To further investigate how nanog, an important homeobox TF, instruct cell fate, we investigated proteomics, morphological and functional changes of zebrafish embryos upon morpholino-mediated knockdown of Nanog expression. The result demonstrated the Nanog monitor the MBT via regulating expression level of ribosome proteins. Finally, to investigate the mechanisms of zygotic genome activation across different species, the expression dynamics from our large-scale quantitative proteomics data of zebrafish (~5000 proteins) and Xenopus laevis (~4,000 proteins) were compared across four stages from fertilized egg to embryo.

Contact: Fei Fang, fangfei1@msu.edu

P16.06 Proteomic Analysis of Alcohol-associated Hepatitis Reveals Glycoprotein NMB (GPNMB) as a Novel Hepatic and Serum Biomarker

Kristofer Fritz (1), Peter Harris (1), Cole Michel (1), Youngho Yun (1), Courtney McGinnis (1), Mohammed Assiri (2), Ali Ahmadi (3), Zhaoli Sun (3), James Roede (4), Matthew Burchill (1), David Orlicky (4), Rebecca McCullough (1)
(1) University of Colorado Anschutz Medical Campus, Aurora, Colorado, United States
(2) King Saud University, Riyadh, Saudi Arabia
(3) Johns Hopkins School of Medicine, Baltimore, Maryland, United States
(4) University of Colorado Anschutz Medical Campus, Aurora, Colorado

Alcohol consumption remains a leading cause of liver disease worldwide, resulting in a complex array of hepatic pathologies. Individuals who progress to a rarer form of alcohol-associated liver disease (ALD), alcohol-associated hepatitis (AH), require immediate life-saving intervention in the form of liver transplantation. While multiple mechanisms have been identified that contribute to ALD, no cures exist and mortality from AH remains high. The rapid onset of AH is poorly understood, and an early-stage serum biomarker could prove to be life-saving. To identify novel pathways associated with AH, our group utilized proteomics to investigate AH-specific biomarkers in liver explant tissues. Protein abundance and acetylomic analyses were performed utilizing nHPLC-MS/MS, revealing significant changes to proteins associated with metabolic and inflammatory fibrosis pathways. Here, we describe a novel hepatic and serum biomarker of AH, glycoprotein NMB (GPNMB). The anti-inflammatory protein GPNMB was significantly increased in AH explant liver (HPLC-MS/MS) and

serum (ELISA) compared to healthy donors by 50-fold and 6.5-fold, respectively. Further, bioinformatics analyses identified an AH-dependent decrease in protein abundance across fatty acid degradation, biosynthesis of amino acids, and carbon metabolism. The greatest increases in protein abundance were observed in pathways for focal adhesion, lysosome, phagosome, and actin cytoskeleton. In contrast with the hyperacetylation observed in murine models of ALD, protein acetylation was decreased in AH compared to normal liver across fatty acid degradation, biosynthesis of amino acids, and carbon metabolism. Interestingly, immunoblot analysis found epigenetic marks were significantly increased in AH explants, including Histone H3K9 and H2BK5 acetylation. The increased acetylation of histones likely plays a role in the altered proteomic profile observed, including increases in GPNMB. Understanding the origin and consequences of these changes will yield new mechanistic insight for ALD as well as identify novel hepatic and serum biomarkers, such as GPNMB.

Contact: Kristofer Fritz, kristofer.fritz@cuanschutz.edu

P16.07 Deep Metaproteome Analysis using a Vanquish Neo UHPLC System Coupled to an Orbitrap Eclipse Tribrid Mass Spectrometer with FAIMS Pro Interface

Amirmansoor Hakimi (1), Kevin L. Schauer (1), Runsheng Zheng (2), Alexander Boychenko (2), David Horn (1), Steven L. Reeber (1), Daniel Lopez Ferrer (1)
(1) Thermofisher Scientific, San Jose, California, United States
(2) Thermofisher Scientific, Germering, Germany

Purpose: LC-MS-based proteomics is a powerful tool for deep profiling of peptides and proteins in complex biological samples. Analyzing the proteome of microbial communities represents a challenge for current proteomics workflows due to the wide dynamic range of metaproteomes. Extensive fractionation is required to address this challenge. Here we use a newly developed high performance nLC column that provides extensive peak capacity and is hyphenated to a High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) interface coupled to a Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer to maximize the proteome coverage of metaproteome samples with very high protein diversity. Methods: 1µg of peptides were separated using a Thermofisher Scientific™ EASY-Spray™ PepMap™ Neo column on a Vanquish Neo UHPLC system coupled to an Orbitrap Eclipse Tribrid MS with and without FAIMS. Three CVs with the least overlap in the peptides were used. The final FAIMS MS method was set to switch between different CVs with a top-speed method in a 3 second total cycle time over a 140 minutes gradient. Results: The effect of gas-phase fractionation using the FAIMS was evaluated. A Vanquish Neo UHPLC system utilizing a 75µm x 75cm EASY-Spray PepMap-Neo column was used to separate 1µg load of peptides from different microbiome standards. The raw files were searched against concatenated databases downloaded from UniPort using Thermofisher Scientific™ Proteome Discoverer™ 3.0 software utilizing, SEQUEST-HT, and INFERYS re-scoring algorithm. Preliminary results showed over 10,000 proteins and 70,000 peptides in the ZymoBIOMICS Microbial Community standard for the No-FAIMS dataset. The addition of FAIMS improved protein identifications by about 19%. Similar improvement was observed in the ZymoBIOMIC Gut Microbiome standard dataset with identification of over 11,000 proteins and 80,000 peptides in the No-FAIMS experiment with a 17.5% improvement in protein identification when FAIMS is used.

Contact: Amirmansoor Hakimi, amirmansoor.hakimi@thermofisher.com

---

P16.08 Comprehensive Comparison of Serum vs. Plasma High-Throughput Proteomics Across a Healthy Gender-Matched Cohort

Ventzi Hristova (1), H. Sophia Chung1 (1), Chiung-Yun Chang (1), Junmin Wang (1), Sonja Hess (1)
(1) Dynamic Omics, Centre for Genomics Research (CGR), Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, Maryland, United States

Purpose: Clinical proteomics focuses on the discovery of novel, early biomarkers in liquid biopsy samples that can be translated into clinical and precision medicine applications. Here we aim to comprehensively compare advantages and challenges associated with serum and plasma, to guide specimen selection. This is achieved using our high-throughput end-to-end proteomics workflow of a healthy gender-matched cohort. Methods Fasting EDTA-treated plasma and matched serum were simultaneously collected from 40 healthy donors, 20 male and 20 female. To assess clot formation variability, multiple serum samples were collected from a subset of donors. Specimens were centrifuged at 1500xg for 15 minutes to remove contaminants including platelets and clots. Samples were either directly digested with EasyPep 96 MS Sample Prep Kit (Thermo Fisher Scientific) or subjected to depletion (High Select Top14 Resin Pierce) prior to digestion. Peptide analysis was performed using 45- and 20-minute gradients with data-independent acquisition (DIA) methods on an Orbitrap Exploris 480 mass spectrometer. DIA bioinformatic analysis was performed using library-based and non-library approaches. Results Proteomics analysis of plasma and serum without depletion consistently demonstrated improved proteome coverage in serum with up to 19% more protein groups identified compared to matched plasma. Replicate analysis show at least 70% of protein groups identified in each sample have <20% CV. Following depletion, the proteome coverage of matched serum and plasma is similar. This supports our hypothesis that clot formation removes select high abundance proteins and allows detection of more moderate to low abundance species. Ongoing analysis of donor-matched serum aims to address if/how variability in clot formation impacts proteomics data. Conclusion Depletion improves proteome coverage for serum and plasma, but in the absence of depletion more proteins are identified in serum compared to plasma. Specimen collection and processing must be evaluated and standardized to ensure minimal variability, in particular due to clotting.>

Contact: Ventzi Hristova, ventzislava.hristova@astrazeneca.com

---

P16.09 Fast library generation using Zeno MS/MS and microflow chromatography

Alexandra Antonoplis (1), Nick Morrice (2)
(1) SCIEX, Redwood City, California, United States
(2) SCIEX, Alderley Park, United Kingdom

Generation of spectral libraries is often thought of as a time consuming procedure, but in reality they can be generated very quickly. Using microflow chromatography with very fast MS/MS acquisition on a QTOF system, then searching the data in the cloud can enable library generation in <1 day. Here this workflow was explored using the ZenoTOF 7600 system, and the impact of larger libraries on DIA protein quantification was tested. Forty-three fractions from high pH reversed phase separations of complex digested samples (K562, Hela) were collected and then analyzed using microflow chromatography (20 min gradients). Data dependent analysis and SWATH acquisition was performed on the ZenoTOF 7600 system using Zeno MS/MS for fast, high sensitivity MS/MS analysis. Protein identification and SWATH acquisition data processing was performed in the cloud using OneOmics suite. The libraries generated from each fractionated cell line contained >8000 proteins and >175000 peptides. Comparing the proteins found from each cell line, >500 new proteins were added to the library by analyzing a second cell line. Processing SWATH acquisition data with the large, combined library enabled extraction of many more proteins from a standard cell line acquisition, resulting in a 30-40% improvement from the previously generated libraries on older platforms. SWATH acquisition experiments were performed with and without Zeno

MS/MS activated across a range of sample loadings and the impact on proteins quantified was assessed. For loads in the 100-400ng on column range, a 2-3 fold increase in proteins quantified <20% CV was observed. For lower loads, even higher gains were observed. Using fast microflow LC and Zeno MS/MS combined with cloud-based data processing, very large scale libraries can be generated in under 24 hours. Higher quality MS/MS libraries can improve information content extracted for SWATH acquisition data. Zeno SWATH can provide significant improvements in protein quantification>

Contact: christie hunter, christie.hunter@sciex.com

P16.10 Enhanced Detection of Low Abundance Proteins in Various Mouse Tissues Using Offline HPLC as a Precursor to LC-MS/MS.

David Parkinson (1), Leena Patil (1), John Price (1)
(1) Brigham Young University, Provo, Utah, United States

The human proteome is vast, and most tissues are dominated by a few distinct high abundance proteins (HAPs). In bottom-up proteomics, peptides from HAPs significantly suppress LC-MS detection of peptides from low abundance proteins (LAPs). This is problematic because detecting LAPs is essential to understanding cellular regulation of the proteome. Traditionally, offline high-performance liquid chromatography (HPLC) is used as an additional dimension to LC-MS to separate digested proteins (peptides), yielding a mild increase in LAP detection as 15–20 concatenations of interspaced fractions are analyzed by subsequent LC-MS. Since this process is time-consuming and non-specific to HAPs, a faster and better way of eliminating HAP interference is desirable. Additionally, since various proteomes contain unique compositions of high and low abundance proteins, we wanted to develop robust strategy that is generalizable to any tissue or organism. Using a nonporous reverse-phase C18 stationary phase, we developed an optimized method that separated intact proteins (prior to digestion) in the first dimension. Since we separated HAPs and LAPs prior to digestion, HAP peptides and LAP peptides were kept separate and could then be detected separately on MS, significantly reducing HAP interference, increasing detection of LAPs that were not originally seen, and increases total peptide coverage for LAPs that originally had a only few peptide matches. We used a refined computer program to dictate our concatenation scheme, separating proteins into various groups, classified by overall abundance. This process was successfully applied to various mouse tissues (liver, serum, muscle, and brain), and each showed comparable results. The methodology we have developed can be easily applied to protein mixtures from any tissue of any species. It is robust and fast, allowing for an increased detection and peptide coverage of LAPs while largely reducing the needed run time.

Contact: David Parkinson, davidhparkinson@gmail.com

P16.11 Deep Proteome Coverage by Offline HPLC Fractionation of Intact Proteins as a Precursor to Bottom-up Proteomic Analysis

Leena Patil (1), David Parkinson (1), John Price (1)
(1) Brigham Young University, Provo, Utah, United States

The human proteome includes more than 20,000 proteins ranging from <10 kDa to >1 MDa and thus the separation and identification of proteins is challenging. Bottom-up proteomics have been extensively used for protein analysis, wherein the proteins are enzymatically digested into smaller peptides prior to mass spectrometry (MS) analysis. The most significant disadvantage of this technique lies in the difficulty to detect low abundance proteins (LAPs) in the presence of high

abundance proteins (HAPs). Many disease associated proteins and drug targets are LAPs. In order to visualize these proteins, we have incorporated offline-HPLC fractionation and concatenation as a preparatory step, allowing us to separate and enrich LAPs from HAPs prior to bottom-up analysis. Protein level fractionation dramatically decreases sample complexity as proteins are diverse in their chemical nature unlike than peptides which show a uniform behavior. We have compared weak anion exchange (WAX) and reversed phase liquid chromatography (RPLC) as an offline HPLC techniques in the first dimension and LC-MS in the second dimension. Using intact protein mixtures from mouse liver tissue, we optimized separation by selecting the stationary phase and modifying mobile phase conditions, as well as flow rate and temperature. Each fraction was concatenated into one of four different groups, based on relative protein abundance. Each group was then analyzed separately with LC-MS. By applying this method as a precursor to LC-MS analysis, we have detected LAPs that would otherwise be missed due to HAPs interference. Offline RPLC fractionation resulted in a better separation at protein level and subsequently better identification in Bottom-up LC-MS analysis than WAX. This could be attributed to the better selectivity and higher interaction in RPLC. We have applied the developed method to serum, brain and muscle tissue and witnessed a substantial increase in the proteome coverage.

Contact: Leena Patil, leenapatil@chem.byu.edu

P16.12 Best practice of FAIMS with Single Compensation Voltage in Single-shot Proteomics using Data Independent Acquisition

Yue Andy Qi (1), Luke Reilly (1), Lirong Peng (1), Erika Lara (1), Daniel Ramos (1), Michael Fernandopulle (2), Caroline Pantazis (1), Julia Stadler (1), Marianita Santiana (1), Faraz Faghri (1), Mike Nalls (1), Priyanka Narayan (3), Andrew Singleton (1), Mark Cookson (1), Michael Ward (2)
(1) National Institute on Aging, Bethesda, Maryland, United States
(2) National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, United States
(3) National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland, United States

Using high field asymmetric waveform ion mobility spectrometer (FAIMS) pro interface for proteomic analyses has gained significant interests because FAIMS can reduce the background contamination peaks and boost the low intensity peaks. Although researchers have demonstrated that multiple compensation voltages (CVs) in FAIMS benefit the total proteome coverage for data dependent acquisition (DDA), this approach may increase cycle time due to switching among multiple CVs, which could affect data independent acquisition (DIA) even more because higher quality MS2 spectra are commonly required for DIA. Herein, we report the best practice of using single CV FAIMS-DIA approach for large-scale proteomics project. We have systematically evaluated multiple factors in this workflow, including (1) segmented (acquisition windows) and FAIMS-based gas phase fractionation, and liquid phase fractionation for spectral library generation; (2) a range of single CV FAIMS for both DDA and DIA; (3) comparison of spectral library and sequence FASTA library-based database search using 3 most common search engines, which are Spectronaut, EncyclopeDIA and PEAKS studio. We used standard HeLa digests, two induced pluripotent stem cell (iPSC) lines for the method development and validate the optimized method in a temporal proteomic profiling of iPSC-derived neurons during a course over 28 days differentiation. Importantly, we generated companion single-cell transcriptomic datasets in parallel to verify our findings in proteomics. We found single CV at -35V outperformed other tested CVs at DIA mode. The FAIMS-DIA (CV = -35V) approach provides the deepest proteome, less inter-sample variation and better reproducibility. Interestingly, we found FASTA-based directDIA outperforms any spectral library-based searches, of which Spectronaut provides the highest peptide identification. This extensively optimized and validated proteomic pipeline can readily identify ~8,000 proteins from any single shot run. Additionally, we innovatively used single-cell transcriptomic to validate neuron markers identified by proteomics.

Contact: Andy Qi, andy.qi@nih.gov

P16.13 Identification of RNA-Binding Proteome in Mouse Macrophages

Deepali Rathore (1), Aleksandra Nita-Lazar (1)
(1) National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, United States

RNA-protein interactions play key roles in cell metabolism and gene expression. Proteins bind to certain structural motifs in RNA using RNA binding domains (RBDs), however, unconventional binding can also occur that does not require conventional RBDs. This binding results in a ribonucleoprotein complex (RNP) which are dynamic in nature. Any perturbations in RNA-protein interactions can leads to cellular dysfunction and disease. Multiple approaches exist to catalogue RNA binding proteins (RBPs), but none comprehensively capture those due to inherent limitations. We developed a mass spectrometry-based approach to identify RNA associated proteome at a global level in mouse macrophages, which entails 254nm ultraviolet crosslinking of RBPs to RNA in live cells, followed by their capture on silica and selective elution of peptides that represent those RBPs. The RBPs are then enzymatically digested, cleaned up of any non-peptide impurities and analyzed with liquid chromatography online (LC) with orbitrap tandem mass spectrometry (MS/MS). The UV dosages are optimized as per the protein ID readout from MS. LC-MS/MS analysis identified 689 proteins with a 60% match with murine database. Gene ontology term analysis showed highest enrichment levels for RNA proteins, mRNA processing, and nucleic acid metabolic processes in the UV crosslinked samples. The crosslinked RBPs were then compared in lipopolysaccharide (an endotoxin that initiates inflammatory response through Toll-like receptor (TLR) 4) activated vs. non-activated mouse macrophages, which identified 27 novel candidates not previously annotated as RBPs but could be possible regulators in TLR signaling pathways. The most enriched KEGG pathways were found to be RNA transport and Spliceosome. The interaction network analysis revealed 59 proteins that are also involved in the MyD88 signaling pathway and 14 proteins with possible involvement in immune response to LPS via spliceosome. This research was supported by the Intramural Research Program of NIAID, NIH.

Contact: Deepali Rathore, deepali.rathore@nih.gov

P16.14 Protein Folding Stabilities are a Major Determinant of Oxidation Rates for Buried Methionine Residues

ETHAN WALKER (1), John Bettinger (1), Kevin Welle (1), Jennifer Hryhorenko (1), Sina Ghaemmaghami (1)
(1) University of Rochester, Rochester, New York, United States

The oxidation of protein-bound methionines to form methionine sulfoxides has a broad range of biological ramifications and it is therefore important to delineate factors that influence methionine oxidation rates within a protein. This is especially important for biopharmaceuticals, such as antibody-based therapeutics, where slow oxidation over long term storage can lead to their eventual deactivation, or degradation. Understanding the factors that drive methionine oxidation can allow one to predict whether a given Met residue will be labile, and thus streamline the development of drug candidates. Previously, neighboring residue effects and solvent accessibility (SA) had been shown to impact the susceptibility of methionine residues to oxidation. In this study, we provide proteome-wide evidence that oxidation rates of buried methionine residues are also strongly influenced by the thermodynamic folding stability of the domains where they reside. We surveyed the E. coli proteome using several proteomic methodologies and globally measured oxidation rates of methionines in the presence and absence of tertiary structure as well as folding stabilities of methionione containing domains. The data indicate that buried methionines have a wide range of protection factors (PFs) against oxidation that correlate strongly with folding

stabilities. Concordantly, we show that in comparison to E. coli, the proteome of the thermophile T. thermophilus is significantly more stable and thus more resistant to methionine oxidation. These results indicate that oxidation rates of buried methionines from the native state of proteins can be used as a metric of folding stability. Additionally, to demonstrate the utility of this correlation, we used native methionine oxidation rates to survey the folding stabilities of E. coli and T. thermophilus proteomes at various temperatures and suggest a model that relates the temperature dependence of the folding stabilities of these two species to their optimal growth temperatures.

Contact: Ethan Walker, ethan_walker@urmc.rochester.edu

P17: Single-Cell Proteomics

P17.01 Application of a library-free dia-PASEF approach for high throughput and high sensitivity proteomics

Diego Assis (2), Stephanie Kaspar-Schoenefeld (1), Elizabeth Gordon (2), Matthew Willetts (2), Christoph Krisp (1), Nagarjuna Nagaraj (1), Gary Kruppa (3)
(1) Bruker Daltonics GmbH & Co. KG, Bremen, Germany
(2) Bruker Scientific LLC, Billerica, Massachusetts, United States
(3) Bruker S.R.O, District Brno-City, Czech Republic

dia-PASEF (Meier et.al., 2019) takes advantage of the additional dimension of separation provided by trapped ion mobility (TIMS). The dia-PASEF cycle time can be reduced to make it compatible with short gradient separation while preserving high selectivity. This combination of DIA and PASEF compensates for the traditional DIA pitfalls: by using a pattern of m/z isolation windows within consecutive tims events, the percentage of ions used in the dia-PASEF can be greatly increased. Here, we evaluate benefits of dia-PASEF including library-free data processing for very short gradients enabling ultra-high sample throughput proteomics. Moreover, we demonstrate the performance of dia-PASEF on low sample amounts. K562 tryptic digests (Promega) were analyzed by coupling an EVOSEP One (EVOSEP) to a trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro 2 for short gradient and timsTOF SCP for low sample loads) using optimized dia-PASEF schemes. Data was processed using a library-free approach with DIA-NN (Demichev et.al, 2021) incorporated into the PASER software (Bruker) and Spectronaut (Biognosys) using default settings. Using a 300 samples per day method (300 SPD) with 200 ng on column, 3,600 protein groups and 23,500 peptides were identified with a run time of only 4.8 min. To demonstrate the potential of dia-PASEF for high sensitivity proteomics we used the timsTOF SCP mass spectrometer in combination with low flow rate delivery from the Evosep system (Whisper100, 40 samples per day with gradient flow of 100nl/min). With this setup and loading only 125 picograms, 1,249 (±123, n=6) protein groups could be reproducibly identified using dia-PASEF using the library-free based approach. To further increase proteome coverage for both, high throughput, and high sensitivity proteomics, we will investigate additional new developments. Our results demonstrate the benefits of using dia-PASEF acquisition for high throughput and high sensitivity proteomics studies using library-free approaches.

Contact: Diego Assis, diego.assis@bruker.com

P17.02 Standardized and easy to deploy nanoLCMS for ultra-sensitive and high-throughput analysis of limited sample amounts and single cells

Oleksandr Boychenko (1), Runsheng Zheng (1), Christopher Pynn (1), Xuefei Sun (2), Alec Valenta (3)
(1) Thermo Fisher Scientific, Germering, Germany
(2) Thermo Fisher Scientific, Sunnyvale, California, United States
(3) Thermo Fisher Scientific, Michigan, Michigan, United States

Proteomics profiling of limited sample amounts like single-cells or cells populations, laser capture microdissected tissue, etc. requires the maximum sensitivity from ESI interface and MS. In addition, the analysis of individual targets like single cells needs sufficient throughput to draw meaningful statistically relevant conclusions. The existing approaches for nanoLCMS analysis of limited samples are not widely adopted due to special setups, non-standard hardware configurations, advanced technical skills required to make connections and operate nanoLC systems at ultra-low nano-flow rates below 100 nL/min.

We developed a set of standardized methods based on the novel nanoLC-MS hardware configuration that allows to make up to 48 ultra-sensitive nanoLC-MS runs per 24 hours using columns with 20 µm internal diameter and flow rates of 100 nL/min and below. We systematically investigated the effect of the flow rate as well as column diameter on the signal intensity, separation efficiency, and the number of protein and peptide identifications. The optimized configuration and nanoLC-MS methods allowed to achieve 60-80% MS utilization and to handle low sample volumes (from 10 to 100 nL). The performance of the approach was tested on the label-free quantification of HeLa samples (0.2-5 ng). The sensitivity improvements up to 10-folds in comparison with nanoLCMS analysis at 300 nL/min led to the identification and quantification of > 2000 protein groups in the single-shot injection of 1 ng of HeLa protein digest. The robust system operation at low-flow rates and high-pressure capabilities opened possibilities for automated high-throughput analysis of limited samples on large scale.

Contact: Oleksandr Boychenko, oleksandr.boychenko@thermofisher.com

---

P17.03 Multiplexed nanoLC System Produces a Label-Free Single-Cancer-Cell Proteome Every 15 Minutes

Kei Webber (1), Sebastian Zapata (1), Thy Truong (1), Jacob Davis (1), Alex Buttars (1), Yiran Liang (1), Madisyn Johnston (1), Fletcher Smith (1), Arianna Mahoney (1), Hailey Jones (1), Jacob Heninger (1), Andikan Nwosu (1), Richard Carson (1), Greg Nordin (1), Ying Zhu (2), Ryan Kelly (1)
(1) Brigham Young University, Provo, Utah, United States
(2) Pacific Northwest National Laboratory, Richland, Washington, United States

Cellular phenotypes can be measured directly using single-cell proteomics (SCP) by mass spectrometry (MS). Such measurements are more sensitive with extremely low-flow separations (<100 nL/min) because of improved ionization efficiency. However, the time required for the sample and mobile phase to clear extra-column volumes in these low-flow systems can lead to inefficient duty cycles wherein the mass spectrometer sits idle for significant amounts of time between analyses. The resulting low-throughput has hindered the adoption of label-free SCP compared to approaches using isobaric multiplexing. As label-free SCP has some distinct advantages over isobaric-labeling methods, we sought to address the low-throughput problem. We here present a multicolumn liquid chromatography (LC) system that more than doubles the throughput of label-free SCP by using two parallel systems to multiplex sample loading, online desalting, analysis, and column regeneration steps. Reliance on feature mapping of high resolution MS1 spectra to generated libraries for peptide identification allowed us to compensate for the Orbitrap's slower acquisition speed and achieve faster nanoLC separations. This high-throughput column system showed good reproducibility between the parallel columns with an 8% difference in median abundances across 15-minute (n = 100) replicate analyses of single-cell-sized aliquots of commercial human cell lysate tryptic digest, with an average of 937 protein groups identified and a CV of 18.5%. Altering the total analysis times to 7, 10, 15, and 30 minutes resulted in an average of 621, 774, 952, and 1622 protein groups identified, respectively. When applied to single HeLa cells, we identified an average of 986 protein groups per cell (n=53) using 30-minute cycles. Upon reducing the cycle time to 15 minutes, we identified an average of 660 protein groups per cancer cell (n=150). This throughput is easily double that of any previous label-free SCP approach, and begins to approach parity with isobarically-labeled approaches.

Contact: Kei Webber, keiwebber3@gmail.com

P17.04 Ultra-Sensitive Proteome Quantification on the timsTOF SCP Mass Spectrometer

Matthew Willetts (1), Kristina Marx (2), Renata Blatnik (2), Verena Telstrom (2), Markus Lubeck (2), Oliver Raether (2), Nagarjuna Nagaraj (2), Gary Kruppa (3)
(1) Bruker Daltonics, Billerica, Massachusetts, United States
(2) Bruker Daltonics GmbH & Co. KG, Bremen, Germany
(3) Bruker S.R.O, Brno

Introduction Single-cell proteomics is a relatively young niche of proteomics compared to single-cell genomics. In recent years, significant progress has been made in sample handling and boosting the sample signal by multiplexing with mass spectrometers. timsTOF SCP is the first commercially introduced mass spectrometer for single-cell proteomics. The modified front-end (orthogonal ion-guiding) of the instrument increases the ion transfer up to five times keeping ultra-high robustness – the default attribute of the timsTOF platform. Here we demonstrate the performance of the instrument for low sample loads in the range of 250 pg to 1 ng combined with robust low flow rate delivery from the Evosep system. Methods Commercially available K562 tryptic peptide digests (Promega) were loaded on the Evotips according to the recommended protocol provided by the vendor. Peptide amounts from 125 pg to 25.6 ng were used to evaluate the performance of the instrument. The Whisper low flow method – 40 samples per day (SPD) with a gradient time of about 28 minutes was used for separating the eluting peptides on a 15 cm column with 75 µm ID coupled to captive spray ionization source using a 10-micron ID zero-dead volume emitter. Samples were analyzed using a dia-PASEF method tuned for high sensitivity measurements. Data were processed in Spectronaut 15. Preliminary Data diaPASEF data processing was performed with Spectronaut 15 using a spectral library of 5,200 protein groups and 54,000 peptides and a further improved library with 9,300 protein groups and 116,000 peptides, respectively. Using the combination of Whisper 40 SPD and the high sensitivity methods on the timsTOF SCP, we could quantify around 1,500 protein groups from 250 pg sample loads and close to 2,000 protein groups from 500 pg loads. Performing a dilution series experiment, around 4,500 protein groups could be quantified in 6.4 ng peptide loads.

Contact: Matthew Willetts, matt.willetts@bruker.com

P18: Top-Down Proteomics and Proteoform Biology

P18.02 Quantitative analysis of Intact Plasma Proteome in Systematic Lupus Erythematosus (SLE) using Protein-Level TMT Labeling and Top-Down Proteomics

Yanting Guo (1), Walter Galie (1), Kellye Cupp-Sutton (1), Dahang Yu (1), Ken Smith (2), Xiaowen Liu (3), Si Wu (1)
(1) University of Oklahoma, Norman, Oklahoma, United States
(2) Oklahoma Medical Research Foundation, OKC, Oklahoma, United States
(3) Tulane University, New Orleans, LA, United States

Systematic Lupus Erythematosus (SLE) is an autoimmune disease that leads to widespread inflammation. SLE biomarker discovery at an early stage is critical for effective treatment, and avoidance of irreversible tissue/organ damage. Plasma is a key clinical sample type for biomarker discovery because its collection is non-invasive, and it circulates among all organs to reflect downstream damage. High-throughput, quantitative MS-based proteomics would be advantageous for plasma biomarker discovery, but this analysis is challenging due to the wide dynamic range of plasma protein abundance. Top-down proteomics techniques can discern unique proteoforms that may also be relevant biomarkers but could become indistinguishable using typical bottom-up proteomics. Here, we present a novel quantitative top-down proteomics platform that integrates simple, precipitation-based protein depletion, protein-level TMT labeling, and top-down MS for intact plasma proteome profiling for SLE biomarker discovery. Our results indicated that a 1:2 (v/v) plasma-to-isopropanol depletion precipitated most of the albumin and enriched low abundance proteins between 10-50 kDa. Depleted SLE and healthy control samples were TMT-labeled using our recently optimized protein-level TMT labeling conditions and analyzed by nano-LC-MS/MS. All proteoforms were deconvoluted by BioPharma Finder and identified by TopPIC Suite. In total, 2727 unique proteoforms were detected. Reporter ion intensities were also extracted from the raw files using an in-house python code and a volcano plot was constructed that showed 47 proteoforms were upregulated in SLE samples compared to controls (cutoff: P-value ≤ 0.05, fold-change ≥ 2). Some PTMs were enriched in SLE samples, such as glycation that has been reported to be related to chronic inflammation in SLE and the disturbance of the oxidative/antioxidative balance. This oxidative stress leads to intensified advanced glycation end products (AGEs) formation. Overall, our results demonstrate that optimized organic precipitation and top-down TMT protocols can be applied to plasma samples for potential biomarker discovery in SLE.

Contact: Yanting Guo, Yanting.Guo-1@ou.edu

P18.03 Single-Molecule Detection of Isoform-Specific Tau Phosphorylation

Parag Mallick (1), Raymond Mak (1), Steven Tan (1), James Joly (1), Lionel Rouge (2), David Arnott (3), Cassandra Stawicki (1), Deborah Park (1), Brittany Nortman (4), Julia Robinson (1), Taryn Gillies (4), Christina Inman (1), Courtney McCormick (1), Seung-Hye Lee (5), Joanna Lipka (5), Nikhil Pandya (6), Thorsten Wiederhold (6), Jarrett Egertson (1), Greg Kapp (1), Alexis Rohou (2)
(1) Nautilus Biotechnology, San Carlos, California, United States
(2) Department of Structural Biology, Genentech, South San Francisco, California, United States
(3) Department of Proteomics, Lipidomics and Next Generation Sequencing, Genentech, South San Francisco, California, United States
(4) Nautilus Biotechnology, San Carlos, California

(5) Department of Neuroscience, Genentech, South San Francisco, California, United States
(6) Department of Biomarker Discovery, Genentech, South San Francisco, California, United States

Post-translational modifications (PTMs) contribute to the vast diversity of proteoforms that drive biological processes. Currently, very little is known about the molecular heterogeneity of proteoforms. This gap can be dominantly attributed to a lack of methods to analyze PTMs on intact single protein molecules on a proteome-scale. Here, we introduce an antibody-based single-molecule analysis platform that includes novel biochemistry for single-protein molecule immobilization, instrumentation for highly sensitive detection, and computation for data interpretation. We use this platform to study Tau proteoforms. Aberrant PTMs on Tau generate a wide diversity of proteoforms that lead to Tau misfolding and aggregation and are associated with neurodegenerative disorders collectively referred to as Tauopathies. First, we validate the platform by measuring defined mixtures of recombinant tau proteins. Next, we use the platform to examine tau protein enriched from human induced pluripotent stem cell (iPSC)-derived neurons. As PTMs can be assigned to individual protein molecules, the platform is able to reveal the molecular heterogeneity of tau proteoforms that is missed by both bulk measurements and by peptide-centric proteomics approaches. Ultimately, quantitative analysis of proteoform molecular heterogeneity will lead to a more detailed view of the tau proteoform landscape towards advancing precision medicine of tauopathies, such as Alzheimer's disease.

Contact: Parag Mallick, paragm@nautilus.bio

P18.04 A survey of truncated proteins across humans and plants reveals novel proteoforms and proteoform-dependent protein complexes

Claire McWhite (1), Wisath Sae-Lee (2), Nicholas Gort Frietas (3), Anna Mallam (2), Edward Marcotte (2)
(1) Princeton University, Princeton, New Jersey, United States
(2) The University of Texas at Austin, Austin, Texas, United States
(3) Harvard Medical School, Cambridge, Massachusetts, United States

Variability of proteins at the sequence level creates an enormous potential for proteome complexity. Here, we systematically survey human and plant high-throughput native proteomics data for abundant truncated variants of protein, where substantial regions of the full-length protein are missing from an observed protein product. In all surveyed species, approximately one percent of proteins across show a short form, which we can assign by comparison to RNA isoforms as either deriving from transcriptional processes or limited proteolysis. Some truncated proteins are observed in both humans and plants, suggesting that they date to the last eukaryotic common ancestor. While many of our detected protein fragments align with known splice forms and protein cleavage events, multiple examples are previously undescribed. We find that truncations occur almost entirely between structured protein domains, even when short forms are derived from transcript variants. Intriguingly, multiple endogenous protein truncations of phase-separating translation-related proteins are similar to fragments produced by enteroviruses during infection. We describe common features of proteins with clear short forms, finding new classes of proteins with potential bi-functionality or switch-like behavior. Finally, we describe novel proteoform-specific protein complexes.

Contact: Claire McWhite, cmcwhite@princeton.edu
