ENABLING TECHNOLOGY FOR BIOLOGICAL DISCOVERY

19th Annual US HUPO Conference

www.ushupoconference.org • www.ushupo.org

Abstract Book

US HUPO 2023

Chicago, IL | March 4-8

Table of Contents

Oral Presentations

Poster Presentations

Oral Presentations

US HUPO 2023

Chicago, IL | March 4-8

Parallel Session 01: Innovation in Mass Spec Imaging and Spatial Omics

IS01.01 | Chemical Imaging of Amyloid Peptide Aggregation and Plaque Formation Dynamics in Alzheimers Disease

Jorg Hanrieder, *University of Gothenburg, Gothenburg, Sweden*

As the Worlds population is getting older the prevalence of age associated diseases such as Alzheimer's disease (AD) is steadily increasing accordingly, currently affecting 12% over the age of 65. As the underlying mechanism remain unclear there is still no curative treatment. It is therefore critical to understand how key pathological factors of AD including beta-amyloid (A?) plaque formation are interconnected and implicated in nerve cell death, clinical symptoms and disease progression.

The main goal of our research is to elucidate the biochemical processes underlying early A? plaque formation in brain tissue.

We use advanced chemical imaging modalities including hyperspectral confocal microscopy, electron imaging and mass spectrometry imaging to delineate in vivo A? build up and deposition at cellular length scales. Together with stable isotope labelling, these tools allow us to visualize A? aggregation dynamics within single plaques across different brain regions and to follow the fate of aggregating A? species from before and throughout the earliest events of precipitating plaque pathology.

We further integrate these experiments with functional amyloid microscopy using structure sensitive fluorescent probes that allow delineate the chain of events underlying amyloid aggregation, deposition and maturation.

These data, provide a detailed picture of the earliest events of precipitating amyloid pathology at scales not previously possible. The results from these studies bring considerable novel information about the deposition mechanism of A? and its toxic interactions with the surrounding. This will open up for developing tailored strategies to affect AD pathology prior to any neurodegenerative mechanisms as well as to develop new biomarkers for AD.

IS01.02 | Multimodal Imaging Mass Spectrometry: Connecting Omics and Imaging to Discover Molecular Drivers of Health and Disease

Jeff Spraggins, *Vanderbilt University, Nashville, TN*

Cellular interactions within the tissue micro-environment form the basis of health and disease for all organisms. Exposure to nutrients, toxins, and neighboring cells trigger coordinated molecular responses that impact cellular function and metabolism in a beneficial, adaptive, or detrimental manner. Acquiring molecular information at cellular resolution is thus crucial for developing a comprehensive understanding of the biology in an organism. Matrix-assisted laser desorption/ionization (MALD) imaging mass spectrometry (IMS) addresses this need by combining the spatial fidelity of classical microscopy with the molecular specificity of a mass spectrometer. This presentation will highlight our work developing new, high-performance technologies for improving the spatial resolution, sensitivity, and specificity of MALDI IMS for metabolite, lipid, and protein mapping. This will include the utilization of high mass resolution FT-ICR MS and high spatial resolution Q-TOF platforms to address the molecular complexity associated with direct tissue analysis. Further, I will describe recent advances in integrated, multimodal methods that enable molecular signals to be correlated to specific biological tissue features. These technologies will be demonstrated through applications that include understanding the molecular drivers of host-pathogen interactions and the construction of comprehensive molecular tissue atlases.

OA01.01 | Combining Hi-PLEX IHC with spatial lipidomics to identify unique lipid metabolic activity in stimulated neutrophils in a lung model

Gargey Yagnik[1], Mark Lim[2], Kate Stumpo[3], Robert K. Ernst[4], **Alison Scott[4] (aliscott@umaryland.edu)**
[1]AmberGen Inc., Billerica, MA, [2]AmberGen, Inc., Billerica, MA, [3]Bruker Scientific, Billerica, MA, [4]University of Maryland, Baltimore, MD

We identified phospholipid (PL) remodeling activity in a mouse pulmonary infection model using the Gram-negative opportunistic pathogen *Pseudomonas aeruginosa* (*Pa*). Lung infiltrating neutrophils during the early innate immune response are an important feature of *Pa* pulmonary infection and their function in the cellular response is a discriminator for the outcome. Our previous studies demonstrated regio-specificity with respect to acyl configuration in response to *Pa* infection. Using a model of neutrophil induction and activation in wildtype mouse lungs, we combined cellular scale spatial lipidomics with Hi-PLEX immunohistochemistry (IHC) to

parse the cell-specific lipid response during neutrophil activation. Briefly, adult female C57Bl/6 mice were treated intranasally with vehicle control (phosphate buffered saline, PBS), cyclophosphamide (CYP), or with a combination of lipopolysaccharide (LPS, from *E. coli* BORT strain) with N-Formyl-Met-Leu-Phe (fMLP). Lungs were prepared according to previously published methods at multiple timepoints up to 2 days post-treatment. Sections (13 microns) were prepared with matrix for dual-polarity PL imaging as we have previously described (PMIDs: 31882724, 32924474). In the histology-guided approach, activated neutrophils were mapped by IHC with anti-Ly6G and feature-positive pixels from the lung parenchyma were binned into a neutrophil category to be compared against feature-negative parenchyma. The regional analysis identified significant enrichment of phosphatidylserine (PS) lipids containing polyunsaturated fatty acid (PUFA) moieties. Future directions for this work will target acyl transferase and deacylase enzymes which are known components of neutrophil cargo to confirm PL remodeling activities in the *Pa* infection model.

OA01.02 | Ambient Mass Spectrometry Imaging of Proteoforms in Biological Tissues with High Spatial Resolution using Nanospray Desorption Electrospray Ionization

Manxi Yang[1] **(yang1832@purdue.edu)**, Daisy Unsihuay[1], Hang Hu[1], Julia Laskin[1]
[1]*Purdue University, West Lafayette, IN*

Mass spectrometry is a powerful tool for the characterization of proteoforms in biological systems. Top-down proteomics has already provided unique insights into the role of posttranslational modifications (PTM) in biology. We have recently combined intact protein detection with on-tissue top-down proteomics to enable spatially-resolved proteofom characterization. In our studies, we use nanospray desorption electrospray ionization (nano-DESI), a sensitive ambient ionization technique, which relies on a localized liquid extraction and electrospray-like ionization of intact proteins. Nano-DESI generates multiply charged proteoform ions, making it a powerful tool for their identification using top-down approaches and mapping their localization in tissue sections. We used nano-DESI mass spectrometry imaging (MSI) to obtain spatial distributions of ~40 proteoforms in mouse brain tissue used as a model system. Our results reveal the differential expression of the individual proteoforms in the tissue. We have also substantially improved the spatial resolution of nano-DESI MSI of proteoforms to 9 μm and enhanced the sensitivity of the experiment by optimizing the design of the capillary-based probe and the thickness of the tissue section. Untargeted imaging of individual proteoforms with cellular resolution will facilitate cell typing, which will provide new insights into biochemical pathways when combined with spatial lipidomics and metabolomics.

Parallel Session 02: Top-Down Proteoform Biology

IS02.01 | Understanding Proteoform Biology Enabled by Novel Top-down Proteomics Technologies

Ying Ge, *UW Madison, Madison, WI*

Top-down mass spectrometry (MS)-based proteomics is arguably the most powerful technology to comprehensively characterize proteoforms that arise from genetic variations, alternative splicing, and post-translational modifications (PTMs), but myriad challenges still remain. We have been developing novel technologies to address the challenges in top-down proteomics in a multi-pronged approach including new cleavable surfactants for protein solubilization, new strategies for multi-dimensional chromatography separation of proteins, novel nanomaterials for enrichment of low-abundance proteins. Additionally, we have been developing a new comprehensive user-friendly software package for native top-down proteomics. Moreover, we have made major advances in top-down proteomics for analysis of intact proteins directly purified from heart tissue, blood, and human pluripotent stem cell-derived cardiomyocytes (hPSC). Importantly, we have linked altered cardiac proteoforms to contractile dysfunction in heart diseases using animal models and human clinical samples. Furthermore, we are harnessing the power of innovative top-down proteomics with patient specific hPSC-derived cardiomyocytes (CMs) in engineered cardiac tissue to understand proteoform biology in cardiac diseases.

IS02.02 | The Blood Proteoform Atlas and the discovery of biomarker candidates in liver transplanted recipients

Rafael Melani, *Northwestern Proteomics, Chicago, IL*

Over the past decades, large-scale proteomics studies have allowed us to begin to better understand human phenotype at a protein level. However, most proteomics measurements only identify peptides and do not precisely determine alternatively spliced sequences or posttranslational modifications of the proteins. The Blood Proteoform Atlas (BPA) presents the primary structures of ~30,000 unique proteoforms, nearly ten times more than in previous studies, expressed from 1,690 human genes across 21 cell types and plasma from human blood and bone marrow. The results indicate that proteoforms better describe protein-level cell biology and are more specific indicators of differentiation than their corresponding proteins (genes), which are more broadly expressed across cell types. Finally, we demonstrate the potential for clinical application by interrogating the BPA in the context of liver transplantation and identifying cell and proteoform signatures that distinguish normal graft function from acute rejection and other causes of graft dysfunction.

OA02.01 | Building Proteoform Maps of Human Tissues by Integrating Top-Down Proteomics with MALDI Imaging

Kevin Zemaitis[1] (kevin.zemaitis@pnnl.gov), Mowei Zhou[1], James Fulcher[1], Dusan Veličković[1], David Degnan[1], Yen-Chen Liao[1], Camryn Pettenger-Willey[1], Lisa Bramer[1], Ying Zhu[1], Ljiljana Paša-Tolić[1]
[1]*Pacific Northwest National Laboratory, Richland, WA*

Spatial distribution of proteoforms is essential for understanding cellular context and single cell biology in human diseases. Most protein imaging methods use antibody probes with selected targets. However, they may not fully capture the diverse post-translational modifications (PTMs) at the proteoform level. Mass spectrometry imaging (MSI) techniques serve as complementary method for untargeted protein detection. Among them, matrix-assisted laser desorption/ ionization (MALDI) is a commonly used method for detecting metabolites and lipids near single cell resolution. The application of MALDI to proteoform imaging is limited due to difficulty in assigning intact protein peaks solely based on mass. Most reports focused on in-situ digested peptides as a proxy to identify proteoforms with ambiguity in defining PTM stoichiometry and close homologs. We developed a top-down proteomics (TDP) workflow to map healthy human kidney and pancreas tissue sections at spatial resolution of 20~30 µm by MALDI-MSI on a modified Orbitrap UHMR for small proteoforms (<20 kDa). In parallel, functional tissue units (TFUs), such as glomerulus, medullary, and tubule in kidney, or islet and acinar in pancreas, were laser microdissected from serial sections for LCMS TDP to generate associated databases of proteoforms. On average, we characterized ~400 quantifiable proteoforms from TFUs corresponding to ~200 cells. The identifications provided the molecular formulae to assign the isotopically resolved peaks in MALDI-MSI using a custom R package “ProteoMatch”. As expected, insulin and pro-glucagon proteoforms were highly localized to pancreas islet. Ongoing analysis will examine spatial distributions of histone proteoforms for unique epigenetic features among different cell types. Our next step is to apply this workflow to generate 3D proteoform images of human pancreas tissue by stacking 2D proteoform

maps of serial sections. The 3D proteoform image will complement other molecular maps (metabolites, lipids, transcripts, etc.) for creating a comprehensive atlas to improve our understanding of cellular heterogeneity.

OA02.02 | Exploring and Comparing the Landscape of Tau Proteoforms

Steven Tan[1], Deborah Park[1], James Joly[1], Brittany Nortman[1], Julia Robinson[1], Sanjib Guha[1], Cassandra Stawicki, Jarrett Egertson[1], Hunter Boyce[1], Taryn Gillies[1], Christina Inman[1], Lionel Rouge[2], David Arnott[2], Joanna Lipka[2], Nikhil Pandya[2], Thorsten Wiederhold[3], Don Kirkpatrick[4], Alexis Rohou[2], **Greg Kapp[1] (gregk@nautilus.bio)**, Parag Mallick[1]
[1]*Nautilus Biotechnology, San Carlos, CA,* [2]*Genentech, South San Francisco, CA,* [3]*Cell Signaling Technology, Danvers, MA,* [4]*Interline Therapeutics, South San Francisco, CA*

Introduction: A wealth of information exists about the array of post-translational modifications (PTMs) on the Tau protein. Changes in specific modifications, like phosphorylation at position 214, have even been linked with development or progression of Tauopathies like Alzheimer's Disease. It is unlikely, however, that the analysis of a single modification or a single modification site is sufficient to truly understand the biological complexity of Tau and its involvement in normal brain function and disease. It is more likely that complex patterns of modifications produce a multitude of different proteoforms and that specific proteoforms should be investigated as they key indicators for health and disease states.
Methods: We are using a single-molecule protein analysis platform and coupling it with existing Tau antibodies directed at splicing variants or phosphorylation events. We bioconjugate Tau molecules to nanoparticles to ensure single-molecule analysis and bioconjugate the Tau antibodies to highly-fluorescent nanoparticles to enable the optical detection of single antibodies. By sequentially probing the Tau molecules with the full set of Tau antibodies, the quantity and diversity of proteoforms can be analyzed in different samples.
Results: The Nautilus platform and analytical approach is validated by first measuring defined mixtures of recombinant Tau proteins. Next, the same method is applied to cell lines producing Tau protein in the presence or absence of phosphorylation modifying conditions. The pattern of Tau modifications can be compared to identify specific proteoforms enriched due to drug treatment.
Conclusions: Post-translational modifications result in a largely unexplored level of heterogeneity for proteins like Tau (and likely for many other disease-relevant proteins). Mapping the landscape of proteoforms in healthy and diseased conditions promises to aid understanding of disease, point to biomarkers linked to disease onset and progression, and possibly even identify molecular tools to monitor and assess therapies to treat disease and improve human health.

Parallel Session 03: Translational MS for Neurological Disorders and Other Diseases

IS03.01 | Capturing Organelle Dynamics with Multimodal Proteomic Strategies

Ling Hao, *George Washington University, Washington, DC*

Lysosomes and mitochondria are membrane-bound organelles in the cell, responsible for trash-disposal and ATP production, respectively. Perturbations in lysosomes and mitochondria have been linked to numerous human diseases, such as cancer, neurodegeneration, and cardiovascular diseases. These organelles are highly dynamic and frequently interact with other cellular components in a transient fashion, which are difficult to capture and visualize in a high-throughput fashion. Here, I will describe our recent efforts in developing and improving proximity labeling, organelle isolation, and dynamic SILAC proteomic methods to characterize sub-organelle microenvironment and protein dynamics in live cells. Patient skin fibroblasts can be reprogrammed into induced pluripotent stem cells (iPSCs) and further differentiated into neurons, enabling the study of live human neurons that were previously inaccessible. Using multimodal proteomic strategies in human neurons, we aim to decipher the molecular mechanisms underlying lysosomal and mitochondrial dysfunctions in brain diseases. I will also introduce our newly established universal and sample type-specific contaminant libraries for both DDA and DIA proteomics that can benefit the broad proteomics community.

IS03.02 | Judith Steen Presentation

Judith Steen, *Boston Children's Hospital, Boston, MA*

Presentation details coming soon!

OA03.01 | Age/Senescence-related Exosome Biomarkers for Biosurveillance of Age-related Diseases

Sandip Kumar Patel[1] **(sandipsinghpatel@gmail.com)**, Jacob P. Rose[1], Rebecca Beres[2], Joanna Bons[1], Roland Bruderer[3], Lukas Reiter[3], Erin Baker[2], Judith Campisi[1], Birgit Schilling[1]
[1]*Buck Institute for Research on Aging, Novato, CA,* [2]*Department of Chemistry, University of North Carolina, Chapel Hill, NC,* [3]*Biognosys AG, Schlieren, Shiselweni, Swaziland*

Aging is the predominant risk factor that limits health. By 2030, 1/6th of the worldwide population will be over 60 years old, leading to a severe health crisis. Aging-related cellular senescence triggers a profound “senescence-associated secretory phenotype” (SASP), including exosomes. SASP exosomes can cause secondary senescence in a paracrine and endocrine manner and thus have tremendous potential for biosurveillance of age-related diseases, which is largely unexplored. Here, we report the first comprehensive and high-throughput analytical platforms to screen a panel of exosome-specific proteins, lipids, and miRNA for age/senescence and health biomarkers. The study design comprised exosome isolation and enrichment by size-exclusion chromatography/Ultrafiltration and antibody-based methods from plasma comparing old and young and investigating primary lung fibroblasts for which senescence was induced by irradiation, doxorubicin and mitochondrial dysfunction-induced senescence (MiDAS). Profiling of exosome-specific proteins (~2,300) and lipids (~350) generated a resource for the fast-growing gerontology community. Using a machine learning model, exosome-specific proteins, lipids, and miRNA were independently grouped for aging/senescent and control samples. Exosome proteomics using DIA-MS (Orbitrap) resulted in 144 differentially regulated proteins in aged plasma and ~1,300 changing proteins from the senescent fibroblasts. In addition, 88 exosome miRNAs were unique to old plasma. Exosome lipidomics on an ion mobility spectrometry-MS platform identified >300 lipid species, of which 23 were differentially regulated in the aged plasma, and 156 lipids were differentially expressed in senescent fibroblasts. Ten potential aging/senescence biomarkers emerged, including proteins (Peroxidasin, Hemopexin, Plasminogen activator inhibitor 1, SPARC, Transforming growth factor beta-1-induced transcript 1 protein), lipids (ceramide and sphingolipids), and miRNA (hsa-miR-532, miR-654-3p, and miR-409-3p). Future studies will validate these molecules in age-related disease cohorts.

OA03.02 | Large-scale, deep plasma protein profiling: An 1800 sample study of Alzheimer's disease

Harendra Guturu[1], Guhan R. Venkataraman[1], Matthijs B. De Geus[2], Sudeshna Das[2], Pia Kivisäkk[2], Serafim Batzoglou[1], Steven E. Arnold[1], **Asim Siddiqui[1] (asiddiqui@seer.bio)**
[1]Seer, Inc., Redwood City, CA, [2]Massachusetts General Hospital, Cambridge, MA

Biofluids, like serum or plasma are a rich source of protein biomarkers for early detection of diseases, but the large dynamic range of protein concentrations in plasma can require complex workflows and trade-offs between coverage and throughput. Here we use an unbiased, deep, and rapid proteome interrogation approach, which leverages multiple physicochemically distinct nanoparticles to provide broad coverage of the plasma proteome at scale. We previously reported on a 200-plasma sample study of 100 Alzheimer's disease and 100 non- Alzheimer's disease with healthy age- and sex- matched controls (HC), which were analyzed with the Proteograph™ Product Suite (Seer Inc.) and liquid-chromatography mass-spectrometry analysis. Across these samples, data-independent acquisition (DIA) yielded 39,699 peptides and 5,060 plasma proteins. Data-dependent acquisition (DDA) yielded 36,496 peptides and 4,706 proteins Here we are reporting an extended study of an additional 1800 samples (including reprocessing of the previous 200 samples) including longitudinal samples from the same individuals. In this study, we plan to investigate a variety of aspects related to Alzheimer's disease including:

- A better classifier for neurocognitive disorders of aging including Alzheimer's disease, frontotemporal dementias, Lewy body disorders, healthy cognitive aging.
- Identifying potential biomarkers and pathophysiological pathways associated with heterogeneity in Alzheimer's disease including, age of onset and rate of progression.

We demonstrate the Proteograph platform's ability to perform unbiased, deep, and rapid interrogation of the plasma proteome, enabling large-scale studies to detect novel insights with clinically relevant potential.

Parallel Session 04: Computational Proteomics: From MS Data to Biological Insights

IS04.01 | Challenges for AI for MS imaging

Kylie Bemis, *Northeastern University, Boston, MA*

While mass spectrometry (MS) imaging allows unparalleled insight into the spatial distribution of molecules in a sample, its data pose challenges for traditional machine learning methods. Spatial statistics and computer vision methods are not suited for such high-dimensional imaging datasets, and machine learning methods must be adapted to deal with MS imaging's imperfect class labels, noisy data, and relatively small sample sizes. Furthermore, instrumentation improvements continue to produce larger datasets and greater file sizes, compounding the computational challenges.

It is more important than ever that MS imaging data is appropriately processed for downstream analysis, but the impact of preprocessing decisions on machine learning results is often unclear. Evaluating this impact is challenging due to imperfect pixel-level annotations for MS imaging. For classification, the goal is to correctly label subregions of a sample (for example, into healthy and disease classes). However, ground truth labels are rarely available on a pixel-by-pixel level, so machine learning methods that require complete labels can produce invalid results. We present our recent work using multiple instance learning with a convolutional neural network (miCNN) in the presence of uncertain class labels. We further examine the impact that different preprocessing workflows can have on the accuracy of the results.

Many of the methods we present are implemented in the open-source R package Cardinal which provides a full software workflow for MS imaging, including data import, pre-processing, visualization, and statistical and machine learning. We will describe how recent updates to Cardinal help provide accessible and scalable computing infrastructure for processing and analyzing MS imaging experiments.

IS04.02 | TIMS DIA-NN: Deep Learning and CCS-enabled DIA at Speed and Scale

Robin Park, *Bruker, Billerica, ME*

Since data-independent acquisition (DIA) was introduced in 2004, DIA acquisition and data analysis tools have been continuously improved, making DIA a crucial technology to identify and quantify thousands of proteins with high reproducibility and deep proteomics coverage. DIA data analysis, in general, relies on spectral libraries constructed from data-dependent acquisition (DDA). Alternatively, the library-free method searches DIA data directly against a protein database. We have developed CCS-enabled TIMS DIA-NN software to support both spectral library and library-free strategies. The tool was integrated into PaSER (Parallel Search Engine in Real-time) to analyze DIA data in real time. Moreover, we implemented and incorporated multiple deep-learning models in the software to improve identification, quantification, and performance.

OA04.01 | Machine Learning to Predict the 3D Architecture of Protein Complexes

Erin Claussen[1], **Kevin Drew[1] (ksdrew@uic.edu)**
[1]*University of Illinois Chicago, Chicago, IL*

Protein complexes are essential to biologic function and when disrupted can cause adverse health outcomes and disease. Knowledge of the 3D structure of protein complexes illuminates their function and disease etiologies. Unfortunately, the 3D structures of many protein complexes are unsolved. Here we describe Direct Contacts 2, our machine learning model for predicting the 3D architecture of protein complexes. Direct Contacts 2 was informed by high throughput mass spectrometry data, including thousands of affinity purification experiments and co-fractionation mass spectrometry experiments. We evaluated our method using experimentally solved structures from the Protein Data Bank as well as a large dataset of Alphafold2 structure predictions(1). We show improved performance over previous methods for predicting direct contacts including the Y2H network HuRI. Direct Contacts 2, integrated with the existing hu.Map 2.0 co-complex network, allows for guidance in identifying protein complexes to explore further. Here we illustrate the usefulness of our model in investigating complexes associated with ciliopathies, specifically proteins OFD1, FOPNL, PIBF1, and KIAA0753. Our model had a high confidence in the direct contact interaction of OFD1 FOPNL and KIAA0753 FOPNL. We used Alphafold multimer to predict the structure of this trimer and were able to see mutations associated with orofacial digital syndrome were predicted to be at the interface. This work informs future research on ciliopathy disease as well as identifying possible approaches for treatment.

1. *Burke DF, Bryant P, Barrio-Hernandez I, Memon D, Pozzati G, Shenoy A, Zhu W, Dunham AS, Albanese P, Keller A, Scheltema RA, Bruce JE, Leitner A, Kundrotas P, Beltrao P, Elofsson A (2021) Towards a structurally resolved human protein interaction network. bioRxiv 2021.11.08.467664; doi: https://doi.org/10.1101/2021.11.08.467664*

OA04.02 | SEC-TMT Enables Dynamic Investigation of Protein-Protein Interaction Networks

Ella Doron-Mandel[1] (ed2853@columbia.edu), Benjamin J. Bokor[1], Yanzhe Ma[1], Lena A. Street[1], Ahmed A. Abdou[1], Marko Jovanovic[1]
[1]*Columbia University, New York, NY*

The majority of cellular proteins interact with at least one partner or assemble into molecular-complexes to exert their function. This network of protein-protein interactions (PPIs) and the composition of macromolecular machines can differ between cell types and physiological conditions. Thus, characterizing the PPI network and how it dynamically changes is a major avenue to discovering biological function and gaining mechanistic understanding of cellular processes. However, producing a global-snapshot of PPIs from a given specimen using LC-MS/MS based techniques requires measuring dozens to hundreds of samples. Consequently, while recent work made seminal contributions by mapping PPIs at great depth, almost all focused on just 1-2 conditions, generating comprehensive but static PPI networks. To study how PPI changes drive cellular processes, we developed a tool enabling their investigation in dynamic systems and at scale. This was accomplished by incorporating tandem mass tag (TMT) multiplexing to a size exclusion chromatography mass spectrometry (SEC-MS) work-flow. SEC-TMT reduces measurement time 10-fold while maintaining resolution and coverage equivalent to the gold standard in the field, identifying >5000 PPIs from cell-lines and tissues, independently. This development opens new horizons in the field, making it feasible to address questions that previously required months within a few weeks, such as conducting differential analyses of the entire PPI network and its architecture, studying how molecular complex composition changes and discovering assembly-state changes of individual proteins under various biological conditions.

Parallel Session 05: Post-Translational Modifications

IS05.01 | Spatial SILAC - Exploring Spatially-Defined Changes to the Proteome

Amanda Hummon, *Comprehensive Cancer Center, The Ohio State University, Columbus, OH*

Three-dimensional cell cultures are attractive models for biological research. They combine the flexibility of cell culture with some of the spatial and molecular complexity of tissue. For example, colon cancer cell lines form spheroids, in vitro mimics of poorly vascularized tumors. The spheroids are composed of a central necrotic core, a middle quiescent layer and an outer proliferative layer of cells, similar to a rapidly growing colon tumor. Our laboratory has characterized the distribution of endogenous proteins via MALDI imaging mass spectrometry in colon spheroids and determined that the molecular gradients correlate with the pathophysiological changes in the structure. We have developed Spatial SILAC, a method to selectively label the cells in the distinct chemical microenvironments in the spheroids and assess the proteomic changes in response to drug treatment. In this presentation, the spatially-localized proteomic changes in response to the kinase inhibitor Regorafenib will be described.

IS05.02 | A Multi-PTM Workflow for Simultaneous Analysis of Protein Phosphorylation, Glycosylation, Acetylation, and Ubiquitination

Hui Zhang, *Johns Hopkins University, Baltimore, MD*

Protein post-translational modifications allow cells to perform physiological functions greater than the coding capacity of human genome. Dysregulation of PTMs is shown to relate to pathological dysfunctions. Studying multiple PTMs is undoubtedly fundamental in understanding protein functions and multi-PTM crosstalk, which is gaining traction in functional proteomic studies. Despite the importance of multi-PTM proteomics, limited studies investigate more than one PTM at a time. In this study, co-enrichments, multiplexing, fractionation, hybrid data acquisition (DDA + DIA), as well as combined data analysis were investigated for the application on tissue samples, focusing on phosphorylation, glycosylation, acetylation, and ubiquitination.

OA05.01 | Hypoxia-mediated ubiquitination signaling in DNA damage response

Haiping Ouyang[1] **(ouyan137@umn.edu)**, Yue Chen[1]
[1]*University of Minnesota, Minneapolis, MN*

Oxygen availability governs the energy homeostasis and development of living organisms and is widely implicated in diseases. Cellular adaptation to the changes of oxygen concentration is regulated by diverse signaling and transcriptional mechanisms. Ubiquitination is a fundamental posttranslational modification pathway that governs cell signaling and protein homeostasis in eukaryotic cells. However, in addition to selected targets, the global ubiquitination dynamics in response to the hypoxia microenvironment remains largely unknown. Using quantitative proteomics, affinity enrichment, and new E3 ligase activity profiling analysis, we explored the hypoxia-mediated dynamic changes of ubiquitination proteome in cancer cells and identified new hypoxia-regulated pathways in protein homeostasis and DNA damage response. Our study identified nearly 12000 ubiquitination sites on more than 3300 proteins and more than 24,000 phosphorylation sites on more than 5200 proteins. Comparing the ratio distributions, we found that hydroxyurea treatment alone led to a clear shift towards the reduced global ubiquitination and phosphorylation abundance upon DNA damage for a specific subpopulation of sites. Interestingly, under hypoxia conditions, hydroxyurea-induced DNA damage inhibited and stimulated ubiquitination and phosphorylation signaling to a greater extent. Pathway enrichment analysis showed that hydroxyurea treatment under hypoxia led to significantly increased ubiquitination targets enriched in DNA damage and replication-related pathways. Most notably, PCNA ubiquitination was significantly enhanced under the hypoxia condition compared to normoxia condition upon DNA damage. Using our recently developed Ubiquitin E3 ligase Activity Profiling Analysis (UbE3-APA), Our analysis showed that several known E3 ligases were significantly enriched among sites that showed a higher level of activation upon hypoxia and DNA damage treatment. Using siRNA-based knockdown analysis, Crispr-KO study, and biochemical and cell-based phenotypic analysis, we identified the key E3 ligase that mediates the potentiation of PCNA ubiquitination under hypoxia upon DNA damage and discovered a significant role of hypoxia microenvironment in mediating the activation of specific DNA damage.

OA05.02 | High-throughput Quantification of Intact Sialylated Glycopeptides Enabled by 12-plex SUGAR Isobaric Tags

Danqing Wang[1] (dwang356@wisc.edu), Miyang Li[1], Junfeng Huang[2], Ting-Jia Gu[1], Yusi Cui[1], Lingjun Li[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Chinese Academy of Sciences Guangzhou Institutes of Biomedicine and Health: Guangzhou, Guangzhou, China

As one of the most common glycosylation type, sialylation plays an important role in cellular recognition, cell adhesion and cell signaling. The high acidity and number of hydroxyl groups in sialic acids have a large influence on surrounding species. Altered expression of sialylated glycoproteins has also been reported in many diseases. Therefore, in-depth profiling of sialylation is crucial to understand its role in disease progression and the development of early diagnostic strategies. Mass spectrometry has become a powerful tool to study intact glycopeptides that contain the information of both glycan structure and glycosylation site. However, the analysis of intact sialylated glycopeptides (SGP) is still challenging due to their micro- and macro-heterogeneities, low abundances, and poor ionization efficiencies stemmed from their negative charges. Furthermore, quantification strategies targeting at SGP are also limited. To overcome these challenges, here we developed a high-throughput quantification strategy for intact SGP by combining mild periodate oxidation and 12-plex isobaric multiplex labeling reagents for carbonyl-containing compound (SUGAR) labeling. An aldehyde group is exclusively introduced to the sialic acid after the oxidation reaction, which can then directly react with SUGAR tag. The chemical labeling strategy was evaluated using an SGP standard and achieved nearly complete reaction efficiency. The strategy was further applied to tryptic digests from more complex samples, including bovine fetuin and mouse heart homogenates. The labeled sample achieved comparable sialyl-glycoproteome coverage compared to label-free samples. In total, 306 intact SGPs were identified from mouse heart tissue sample with high confidence and Byonic score > 300. The reporter ions generated from the labeled samples showed a ratio close to the expected ratio of 1:1:1:1 or 1:2:5:10, respectively, demonstrating good quantification accuracy. Overall, this study provides an effective and highly efficient approach for profiling and quantifying site-specific sialylation changes in complex biological systems.

Parallel Session 06: Advances in Bioinformatics

IS06.01 | Machine Learning in Biomarker Discovery: It's All About Samples And Features (And Nuance)

Heather Desaire, *University of Kansas, Lawrence, KS*

Looking for disease biomarkers in omics datasets is often compared to looking for a needle in a haystack; the number of potential markers that can now be quantified is immense. The challenge for data scientists is to find the needles (true biomarkers), when they are present and to not be misled by potentially promising candidates that turn out to be no more useful than shiny pieces of hay. The presentation will address this challenge, known as feature selection, from a variety of perspectives. As a cautionary tale, we will show that an improperly executed feature selection strategy leads to identifying disease biomarkers that appear to offer ">90% accuracy", but are, in fact, no better than flipping a coin for predicting disease. Unfortunately, the improperly executed strategy is a pervasive mistake that currently occurs when researchers combine feature selection methods with proteomics datasets; we will demonstrate how the problem can be avoided. A second vignette will illustrate the unequivocal importance of using racially diverse sample sets in biomarker studies. We will show that biomarker candidates for Alzheimer's Disease, which may be optimally beneficial for a broad, racially diverse population, can be overlooked if the underlying datasets used to identify them are not racially diverse. Identifying useful disease biomarkers is a formidable challenge, but real successes are possible. Both the samples and the features are important, and carefully attending to the nuances gives the magic to the method.

IS06.02 | Scribe: Next Generation Library Searching for DDA Experiments

Brian Searle, *Ohio State University Medical Center, Columbus, OH*

Spectrum library searching is a powerful alternative to database searching for data dependent acquisition (DDA) experiments, but has been historically limited to identifying previously observed peptides in libraries. Here we present Scribe, a new library search engine developed considering "lessons learned" from building data-independent acquisition (DIA) search engines. Scribe is specifically designed to leverage deep learning fragmentation prediction software such as Prosit. Rather than relying on highly curated DDA libraries, this approach predicts fragmentation and retention times for every peptide in a FASTA database. By leveraging expected relative fragmentation and retention time values, we find that library searching with Scribe can outperform traditional database searching tools both in terms of sensitivity and quantitative precision.

OA06.01 | Illuminating the Dark Cancer Phosphoproteome through a Machine Learned Co-Regulation Map of 30,000 Phosphosites

Wen Jiang[1] (Wen.Jiang@bcm.edu), Eric Jaehnig[1], Bing Zhang[1]
[1]*Baylor College of Medicine, Houston, TX*

Mass spectrometry-based phosphoproteomics enables proteome-wide analysis of protein phosphorylation in biological samples. As a prime example, the Clinical Proteomic Tumor Analysis Consortium (CPTAC) has performed phosphoproteomic profiling for 1,191 tumors spanning 11 cancer types, generating quantitative data on 77,442 phosphosites. Unfortunately, only less than 5% of these phosphosites have annotated regulatory kinases or biological functions, limiting our ability to gain functional insights into cancer signaling. Because co-regulation is a strong indicator of functional association, we leveraged machine learning and the vast amount of CPTAC data to build a co-regulation map of phosphosites to facilitate functional interpretation of phosphoproteomics findings. Based on a ground-truth dataset including 98,402 pairs of phosphosites known to be regulated by the same kinase (positives) and 1,317,273 pairs by kinases from distant kinase families (negatives), we developed a classifier to distinguish the positive and negative pairs using CPTAC phosphoproteomic data, protein-protein interaction data, and phosphopeptide sequences. Applying the trained classifier to 3 billion phosphosite pairs identified 2,569,519 (0.08%) with high probability of co-regulation, i.e., 400 times more likely to connect positive pairs than negative pairs in an independent ground-truth dataset. These pairs constituted a co-regulation map of 30,499 unique phosphosites, called CoPheeMap.

To demonstrate the utility of CoPheeMap, we integrated its network embedding features with other routinely used features such as Position-Specific Scoring Matrices scores to develop a machine learning model to predict kinase substrate associations (KSAs). The resulted model CoPheeKSA showed superior performance with an AUROC of 0.97 and identified 12,991 high-quality novel KSAs involving 7,908 phosphosites. In another application, CoPheeMap based information propagation (CoPheeProp) assigned ~5,000 phosphosites to different signaling pathways with high specificity, increasing existing knowledge by 10-fold. Together, CoPheeMap, CoPheeKSA, and CoPheeProp provided a systematic framework to illuminate the dark phosphoproteome, paving the way to a

comprehensive understanding of human cancer signaling.

OA06.02 | Comprehensive single-cell proteomics data analysis using FragPipe

Fengchao Yu[1] (yufe@umich.edu), Kai Li[1], Kevin L. Yang[1], Daniel Polasky[1], Alexey I. Nesvizhskii[1]
[1]*University of Michigan, Ann Arbor, MI*

Single-cell proteomics is a booming topic because of its ability to analyze proteins in individual cells. Researchers have proposed label-free-based and tandem mass tag-based approaches to generate high-quality mass spectrometry data. People have also proposed the use of wide isolation windows in data-dependent acquisition (DDA) to co-fragment multiple peptides to increase sensitivity. Although many exciting results have been reported using traditional proteomics tools, these tools were not designed for single-cell data. For example, there is no false discovery rate (FDR) control for match-between-runs (MBR). Controlling FDR is critical for single-cell data because peptide transfer contributes a large proportion to the detection of peptides and proteins. We have demonstrated that without FDR control, many proteins were falsely transferred.

Recently, our lab has developed FragPipe to meet the requirements of new applications and data types. Here, we extend FragPipe to make it more suitable for single-cell data. We improved IonQuant, an MS1-based quantification tool, to make it more sensitive and accurate. To fully utilize the peptide signal, IonQuant performs MBR and re-quantifying simultaneously. MBR transfers peptides from one run to another. Re-quantifying transfers peptides within the same run according to mass differences and retention time. We showed that with both MBR and re-quantifying, FragPipe had higher sensitivity for analyzing single-cell and bulk cell data. We also developed MSFragger-DIA to search peptides directly from DIA data. The detected peptides are used to build a spectral library for quantification analysis. The database searching does not require any spectral library or feature detection, which makes it fast and robust. MSFragger-DIA, along with other DIA-related tools, have been incorporated into FragPipe. Using single-cell and low-input datasets, we demonstrated that FragPipe had higher sensitivity and faster speed than other tools.

Parallel Session 07: Protein Dynamics and Turnover

IS07.01 | Novel Gene-Specific Translational Mechanism in Human Diseases

Xian Chen, *University of North Carolina (UNC)-Chapel Hill, Chapel Hill, NC*

By largely unknown mechanisms, COVID-19 patients with pre-existing chronic inflammatory diseases are vulnerable to severe symptoms. Excessive serum cytokines and abnormally activated macrophages are found in the bronchoalveolar lavage fluid of these severe COVID-19 patients. Thus, effective therapies to minimize mortality depends on a mechanistic understanding of SARS-CoV-2-induced pathogenesis. We report a novel gene-specific translation mechanism that suppresses synthesis of proteins associated with SARS-CoV-2 transmission, impaired T cell function, blood clotting, and the host hyperinflammatory response. First, we employed our chromatin-activity based chemoproteomics (ChaC) approach to dissect the interactome/pathways associated with G9a, a histone methyltransferase whose mRNA was upregulated with virus load in COVID-19 patient peripheral mononuclear cells. This ChaC analysis revealed that constitutively (enzymatically) active G9a interacted with multiple regulators of translation and ribosome biogenesis. This finding implicated a noncanonical function of G9a in the translational regulation of COVID-19 immunopathogenesis. Accordingly, using our translatome proteomics approach, we identified and profiled proteins whose translation depended on G9a activity, that is, G9a upregulated the widespread translation of a battery of COVID-19 pathogenesis-related genes. Mechanistically, based on ChaC identification of G9a interaction with METTL3, an N6-methyladenosine (m6A) RNA methyltransferase, we found that G9a methylates the nonhistone protein METTL3 to co-upregulate m6A-mediated translation (synthesis) of specific proteins functionally associated with immune checkpoint and hyperinflammation. Therefore, we conducted a correlated multiomics study on the m6A/METTL3 transcriptome, proteome, and phosphoproteome of SARS-CoV-2 infected human alveolar epithelial cells that overexpressed ACE2 (A549-hACE2) with or without treatment with inhibitors of G9a. Results indicated that G9a inhibition reversed the multiomic landscapes established by SARS-CoV-2 infection and from which proteins that showed G9a-dependent translation unite the networks associated with viral replication, virus-host interactions, T cell activation/proliferation, and systemic cytokine response. Inhibition of G9a suppressed SARS-CoV-2 replication, which validated our discovery of G9a-regulated translational mechanism of COVID-19 pathogenesis. Further, our correlated multiomics analysis revealed a mechanism of G9a inhibition action that showed multifaceted virus- and host-directed therapeutic effects. Because translational regulation is a precise and energy-efficient step of controlling the expression of proteins, inhibiting the G9a-mediated translation regulatory mechanism can be highly specific and effective by immediately suppressing the aberrant synthesis of COVID 19-related proteins without the need for altering transcriptional activation and mRNA processing steps.

IS07.02 | Optimizing dynamic SILAC based protein turnover measurements in vivo

Marko Jovanovic, *Columbia University, New York, NY*

The last years have shown how precise measurements of protein turnover – the interplay between protein production and loss - yield important insight about gene expression regulation. Moreover, several different approaches have been introduced to provide precise turnover measurements for thousands of proteins. These include dynamic SILAC measurements which were either followed by deep fractionation and data dependent acquisition (DDA) proteomics measurements, also sometimes combined with TMT labeling, or by data independent acquisition (DIA) measurements of unfractionated samples. We are systematically testing these different approaches in two mammalian in vivo model systems. First, we are determining protein turnover changes in mouse organoids derived from pancreatic tumor and metastases. Second, we are measuring protein turnover in brain tissue from wildtype mice and mice where an important regulator of protein homeostasis is knocked out. In my talk I will share insight about our ongoing work to optimize the yield of reliable protein turnover measurements in both these systems and the advantages and challenges associated with all the tested approaches to measure protein turnover.

OA07.01 | Simultaneous Proteome Localization and Turnover (SPLAT) Analysis Reveals New Spatiotemporal Features of Unfolded Protein Responses

Jordan Currie[1] **(currie.jordanm@gmail.com)**, Vyshnavi Manda[1], Veronica Hidalgo[1], Maggie P. Lam[1], Edward Lau[1]
[1]*University of Colorado School of Medicine, Aurora, CO*

The functionality of proteins is dependent on their spatial and temporal distributions, neither of which is directly reflected in the measurement of static protein abundance in a system. The unfolded protein response (UPR) is a critical element of diverse human diseases where cells modulate the synthesis, degradation, and trafficking of endomembrane proteins in response to ER stress. Current knowledge into the regulation of this process is limited by the lack of methods to simultaneously capture the spatial and temporal contexts of proteins of interest and how they interact with each other. Here we report a mass spectrometry based proteomics workflow named Simultaneous Proteome Localization and Turnover (SPLAT). SPLAT builds on prior work in protein

turnover measurements and subcellular localization profiling, by combining dynamic stable isotope labeling, differential ultracentrifugation, and kinetic modeling to concurrently measure changes in protein turnover and subcellular localization under perturbation in one experiment.

We applied SPLAT to characterize protein spatiotemporal dynamics in two models of UPR based on thapsigargin and tunicamycin treatments in human AC16 cardiomyocytes. Both UPR models led to a global suppression in protein synthesis, but we also observed proteins (15 in thapsigargin and 23 in tunicamycin) with increased turnover including HSPA5, DNAJC3, and KIF5B, suggesting they are regulated proteins involved in the stress response. At the same time, we found 335 and 285 potential translocators induced by thapsigargin and tunicamycin, respectively, including known endomembrane trafficking proteins such as flotillins and SEC23/24 homologs. We identified instances where existing and newly synthesized proteins have differential localizations suggesting SPLAT may be useful for monitoring kinetics of protein import and assembly. Taken together, we report a method to simultaneously characterize protein spatial and temporal parameters, which may be useful for discovering signatures of dynamic cellular responses, such as responses to stress, disease, and off-target effects of pharmaceuticals.

OA07.02 | Time-resolved interactome profiling to deconvolute protein quality control dynamics

Lars Plate[1] **(lars.plate@vanderbilt.edu)**, Madison Wright[1], Bibek Timalsina[1], Valeria Garcia Lopez[1], Crissey Cameron[1]
[1]*Vanderbilt University, Nashville, TN*

Cellular processes are controlled by protein-protein interactions that require tight spatial and temporal regulations. Accordingly, it is necessary to elucidate the dynamics of these transient protein interactions to fully understand the associated processes and their dysregulation during disease states. The proteostasis network (PN), which regulates protein folding, trafficking, and turnover, has become an attractive target for therapeutic intervention to combat protein misfolding diseases. However, one challenge is the lack of understanding in the timing and coordination of PN-client protein interactions, as there are few methodologies available to quantify transient protein-protein interactions with time-resolution at an organelle-wide scale. To address this challenge, we developed a quantitative mass spectrometry method, ***t***ime-***r***esolved ***i***nteractome ***p***rofiling (TRIP), to characterize the PN dynamics and better understand protein quality control processes. TRIP combines pulsed unnatural amino acid labeling with in-situ cross-linking, followed by multi-step affinity purification of the time-synchronized fraction for the protein of interest. By capturing these labeled populations across chase timepoints and enriching them into time-resolved fractions, TMT-multiplexed quantitative proteomics can then temporally resolve protein-protein interactions. We will show that TRIP can uncover the dynamics of the secretory proteostasis network at a systems level and elucidate the specific protein quality control defects associated with secretion-deficient, mutant forms of the prohormone thyroglobulin. We are able to track mutation-specific alterations across pathways such as N-glycosylation, Hsp70/90 assisted folding, disulfide/redox processing, and degradation. Further RNAi screening revealed key protein degradation components, whose inhibition can restore mutant thryroglobulin secretion. These results uncovered new protein quality control mechanisms available for therapeutic targeting of hypothyroidism. Finally, we will highlight the broader utility of our TRIP methodology to elucidate cellular protein interaction dynamics, including host-virus interactions of coronavirus nonstructural proteins.

Parallel Session 8: Biomarkers and Precision Medicine

IS08.01 | Profiling Phosphoproteome Landscape in Circulating Extracellular Vesicles from Microliters of Biofluids

Andy Tao, *Purdue University, West Lafayette, IN*

Many biological processes are regulated through dynamic protein phosphorylation. Monitoring disease-relevant phosphorylation events in circulating biofluids is highly appealing but also technically challenging. We introduce here a functionally tunable material and a strategy, extracellular vesicles to phosphoproteins (EVTOP), which achieves one-pot EV isolation, extraction and digestion of EV proteins, and enrichment of phosphopeptides starting with only trace amount of biofluids. The streamlined, ultra-sensitive platform enables us to quantify 500 unique EV phosphopeptides with only a few ?L of plasma and over 1,200 phosphopeptides with 100 ?L of cerebrospinal fluid (CSF). We demonstrated its clinical application through evaluating the outcome of chemotherapy of primary central nervous system lymphoma (PCNSL) patients with small volume of CSF, presenting a powerful tool for broad clinical applications.

IS08.02 | Can We Democratize Precision Medicine?

Jennifer Van Eyk, *Cedars-Sinai Medical Center, Los Angeles, CA*

An underlying premise of precision medicine is that increasing the number of clinical and molecular features will improve accuracy diagnosing disease and predicting clinical outcomes (e.g., a biomarker predictive value, PPV). In other words, more is better. We propose that developing a parsimonious multi-omic model, composed of the minimal number of features able to provide similar or equal predictive performance to models produced with larger and more complex analyte compositions is required for traction in the clinical situation, especially in health care systems with economic burden. Furthermore, we propose that the features underlying a parsimonious model will represent the minimal mechanisms driving the disease outcome even in diseases with high heterogeneity in clinical manifestation. As a first test of our proposal, we analyzed 74 patients with pancreatic ductal adenocarcinoma (PDAC) obtaining >6500 features from clinical, computational pathology, and molecular (DNA, tissue RNA, tissue and plasma protein, and plasma lipid). Multiple independent machine learning models were developed and tested on curated single- and multi-omic analyte panels to determine their ability to predict clinical outcomes in patients. Interesting, the best performing multi-Omic models were comprised of different feature types even if they had equivalent PPV and accuracy for survival (0.85, 0.87, respectively), suggestive of diverse disease mechanisms achieving the same clinical outcome. The parsimonious model with 589 multi-Omic features had the same PPV while <50 features comprised of only plasma lipids and plasma protein had only slightly lower PPV. Thus, the parsimonious model is both cost effective and easily deployable in clinical practice, regardless of health care system, especially if two tier analysis is deployed with an initial plasma sample followed, if positive, by a more invasive biopsy. The adoption of remote sampling devices, where an individual can take their own blood sample and submit by mail to practitioners, could reduce barrier in medical deserts leading to broader adoption of clinical care.

OA08.01 | Proteoform-reaction-monitoring (PfRM) and the discovery of biomarker candidates in liver-transplanted recipients

Che-Fan Huang[1] (chefanhuang2015@u.northwestern.edu), Rafael D. Melani[1], Fernanda Negrão[1], Joseph Greer[1], Bryan Early[1], Ryan Fellers[1], Josh Levitsky[1], Neil L. Kelleher[1]
[1]*Northwestern University, Evanston, IL*

Acute rejection after liver transplantation is the leading cause of mortality among liver transplanted recipients (LTRs). Unfortunately, patient prognosis is based on invasive procedures like biopsy. Therefore, we are applying mass spectrometry-based top-down proteomics to detect differentially expressed proteoforms from LTR's blood cell samples. Ultimately, we aim to develop a minimally invasive, fast, and sensitive assay for diagnostic purposes. Previously, we performed a top-down quantitative analysis of peripheral blood mononuclear cells (PBMCs) and validated a set of 24 proteoforms in a second cohort of LTRs. Here we develop a new targeted quantitative assay named proteoform reaction monitoring (PfRM) to analyze a validation cohort of 100 LTR samples collected longitudinally from 25 patients at four time-points over a year. Patients were divided into three groups: transplant excellent; acute dysfunction, no rejection; and acute rejection. Proteins from PBMCs were pre-fractionated using gel electrophoresis. Standard proteins and the 0-30 kDa fraction of the PBMCs were analyzed by PfRM on an Orbitrap Eclipse. The PfRM method captures all 24 immunoproteoforms, and the obtained data were analyzed using Proteoform Finder software. Data collected on standards demonstrate that the PfRM method improves quantification metrics (low femtomolar) compared to MS1-based assay while providing fragment ion specificity. Our initial analysis of the longitudinal LTR samples reveals several proteoforms showed differential regulation over the time course samples. For example, profilin-1, platelet basic protein, and platelet factor 4 all spiked in visit 2 of

acute rejection patients and dropped in later visits, while these proteoforms remained low in patients with healthy liver transplants. The PfRM workflow also allowed us to differentiate and quantify two proteoforms of platelet factor 4 independently using proteoform-specific fragments. This work suggests that targeted analysis of blood proteoforms can enhance diagnostic methods based on the discovery and validation of protein-based biomarkers.

OA08.02 | Biomarker discovery in loss of TDP-43 function in frontotemporal dementia

Yue Andy Qi[1] (andy.qi@nih.gov), Sahba Seddighi[1], Colleen Bereda[1], Jessica Roberts[2], Mandy Guo[1], Luke Reilly[2], Andrew Singleton[3], Nicholas Seyfried[4], Michael Ward[2]
[1]NIA/NIH, Bethesda, MD, [2]National Institutes of Health, Bethesda, MD, [3]Center for Alzheimer's Disease and Related Dementias, National Institute on Aging, National Institutes of Health, Bethesda, MD, [4]Emory University, Atlanta, GA

Frontotemporal dementia (FTD) is related adult-onset neurodegenerative diseases with extensive overlap in clinical, genetic, and neuropathological features. The loss of the DNA/RNA binding protein TDP-43 occurs up to 80% of the FTD cases. However, no clinically relevant biomarkers of TDP-43 pathology exist, impeding both patient stratification and therapeutic development. Recent evidence showed that dysregulation of TDP-43 induces the formation of cryptic exons in patient brains. Therefore, we aimed to monitor the protein products of TDP-43 related mis-splicing for putative biomarker discovery for FTD. Using CRISPRi knockdown of TDP-43 in iPSC-derived neurons followed by long-read total RNAseq, we detected 384 cryptic exon formed transcripts, including well-characterized transcripts such as STMN2. We identified 65 *de novo* cryptic peptides assigned to 12 genes in iPSC-neurons using a comprehensive proteogenomic approach incorporating a customized database including mis-splicing variants called from RNAseq. Furthermore, we discovered 48 of these mis-spliced transcripts in post-mortem cortex from 10 FTD patients and validated 6 corresponding peptides using proteomics, and 2 parental proteins by immunoblotting. We next developed a proteogenomic informatic pipeline to retrieve the cryptic peptides in human cerebrospinal fluid (CSF) processed by data independent acquisition-based shotgun proteomics. In this cohort study including 85 FTD cases and 90 health controls, we identified 20 cryptic peptides, and 12 of them have been observed in our datasets of iPSC-neurons and/or human brains. Next, we found 4 peptides derived from HDGFL2 and RSF1 are significantly elevated in cases via a targeted MRM assay. Overall, we first demonstrate that cryptic exon transcripts identified in human brains are translated to protein products and can be detected in human CSF, and developed an integrated proteogenomic-based targeted proteomic approach for clinical use. Importantly, our study highlights the precision medicine where the cryptic peptides are holding the great potential as *de novo* biomarkers for FTD early diagnosis.

Parallel Session 9: Fundamental Understanding of the Nervous System

IS09.01 | Time Resolved Metabolomics in the Brain using Sampling with LC-MS

Robert Kennedy, *University of Michigan, Ann Arbor, MI*

Brain extracellular space contains a wide range of molecules including neurotransmitters, neuromodulators, and metabolites. The chemical milieu in this space is indicative of cellular activity and is involved in regulation of that activity. Samples from this space can be accessed by microdialsyis probes. Most work to date using microdialysis has focussed on measuring a small number of neurotransmitters or metabolites at a time. In this work we apply LC-MS based metabolomics approaches to monitoring the brain metabolome. Our goal is to both identify the chemicals present and begin to understand their relationship to phenotype, behaviour, drug effects, or disease state. In undirected metabolomics compounds can be identified by matching the mass spectra to a database. Often only a small fraction of the signals detected can be attributed to specific compound. For example, it is not uncommon to detect 104 “features” (signal at a given retention time and mass) but only identify a few hundred compounds. We show how deeper analysis can be performed by using advanced separations and concentrated samples. Once identified, compounds can be tracked using faster seprations for good throughput. We have identified over 300 compounds present in the brain extracellular space so far with potential for many more. We also show how directed metabolomics methods can be used to uncover chemical differences of phenotypes, in this case the HR/LR behavioural model, and relate these to differences in behavior. Finally, we describe a LC-MS based method to using stable-isotope tracing to track specific pools of glutamate as a neurotransmitter.

IS09.02 | The Role of Medium Spiny Neurons in Neurodegenerative Diseases, such as Huntington's Disease and XDP Parkinsonism

Birgit Schilling, *Buck Institute for Research on Aging, Novato, CA*

Huntington’s disease (HD) is a neurodegenerative disease caused by a CAG repeat expansion in the Huntingtin (HTT) gene. The resulting polyglutamine (polyQ) tract alters the function of the HTT protein. Although HTT is expressed in different tissues, the medium spiny projection neurons (MSNs) in the striatum are particularly vulnerable in HD. Thus, we sought to define the proteome of human HD patient-derived MSNs. We differentiated HD72 induced pluripotent stem cells and isogenic controls into MSNs and carried out quantitative proteomic analysis by two approaches. First, using data-dependent acquisitions with FAIMS (FAIMS-DDA) for label-free quantification on the Orbitrap Lumos mass spectrometer, we identified 6,323 proteins with at least two unique peptides (FDR ? 0.01). Of these, 901 proteins were significantly altered in the HD72-MSNs, compared to isogenic controls. Second, we quantitatively validated protein candidates by comprehensive data-independent acquisitions on a TripleTOF 6600 mass spectrometer quantifying 3,106 proteins with at least two unique peptides. Functional enrichment analysis identified pathways related to the extracellular matrix, including TGF-beta regulation of extracellular matrix, epithelial-mesenchymal transition, DNA replication, senescence, cardiovascular system, organism development, regulation of cell migration and locomotion, aminoglycan glycosaminoglycan proteoglycan, growth factor stimulus and fatty acid processes. Conversely, processes associated with the downregulated proteins included neurogenesis-axogenesis, the brain-derived neurotrophic factor-signaling pathway, Ephrin-A: EphA pathway, regulation of synaptic plasticity, triglyceride homeostasis cholesterol, plasmid lipoprotein particle immune response, interferon-? signaling, immune system major histocompatibility complex, lipid metabolism and cellular response to stimulus. Moreover, proteins involved in the formation and maintenance of axons, dendrites, and synapses (e.g., Septin protein members) are dysregulated in HD72-MSNs. Importantly, lipid metabolism pathways were altered, and we found that lipid droplets accumulated in the HD72-MSNs, suggesting a deficit in lipophagy. Our proteomics analysis of HD72-MSNs identified relevant pathways that are altered in MSNs and confirm current and new therapeutic targets for HD.

OA09.01 | The Neuropeptide Neuroparsin-A Regulates Caretaking Behavior in Leafcutter Ants

Michael Gilbert (michael.gilbert@pennmedicine.upenn.edu), University of Pennsylvania, Philadelphia, PA

The leafcutter ant, Atta cephalotes, is the most socially complex insect known, and offers a new powerful experimental paradigm to investigate the role of gene regulation, social influence, and epigenetic mechanisms in a phenotypically plastic system at the organismal level. A. cephalotes has 8 behaviorally and morphologically distinct worker castes with nearly identical genomic DNA, but the epigenetic mechanisms underlying these differences is completely unknown. A barrier to understand chromatin-based gene regulation and sociality has been lack of a model system. Here, we find that Neuroparsin-A (NPA), a neuropeptide previously

explored in the context of social behavior in locusts, regulates caretaking behavior in leafcutter ants. Brain knockdown of NPA results in acquisition of caretaking behavior, measured by a behavior assay using a 3D-printed chamber to reflect the physiology of the colony. Transcriptomic analysis of the NPA and its receptor (VKR) confirmed acquisition of nurse-biased genes. To show reversibility of the behavior, we performed absolute quantification with mass spectrometry to determine the amount of NPA in the brains of each caste. We then expressed and purified A. cephalotes NPA in Sf9 cells and injected into the brains of the gardener caste, which possess caretaking behavior at baseline. Strikingly, injection of NPA results in loss of caretaking behavior in the gardener caste. Transcriptomics following the NPA injection show and inverse correlation with the sequencing from the NPA and VKR KD. We translated these findings into higher order mammals by performing comparative proteomics and transcriptomics against the forager and nurse castes of the naked mole mole rat - another eusocial species. Finally, we performed in vitro experiments with addition of the NPA to naked mole primary brain cultures to explore novel exaptation of the insulin pathways from insects to mammals.

OA09.02 | Quantitative histone proteoform analysis of the Mus Musculus brain throughout lifespan and with life extension, spatial, and cell type specificity

Bethany C. Taylor[1] (bethanyt@bcm.edu), Karl F. Poncha[1], Nicolas L. Young[1]
[1]*Baylor College of Medicine, Houston, TX*

Histone post-translational modifications and proteoforms are essential epigenetic regulators of the eukaryotic genome. Histone families H4, H3, H2B, and H2A make up the core nucleosome, the essential physiological repeating unit of the genome, while the linker histone H1 family is involved in higher order genome structure and regulation. We have developed high-throughput quantitative top- and middle-down histone proteoform analysis methods that enable rigorous mechanistic biological inquiry *in vivo*. We achieve proteoform quantitation from less than 25 mg of brain. With these methods, we have systematically investigated the dynamics of histone proteoforms in neuroepigenetics. Here we reveal 1) proteoform-level mechanisms of a life extension phenotype compared to normal aging, 2) a quantitative proteoform map of mouse brain regions, 3) spatial difference in proteoform dynamics during adolescent development and 4) cell-type specific histone H1 dynamics in cerebellar granule neurons (CGNs) during post-natal development. Through nuclear AAV expression of a lysosomal transcription factor, TFEB, we obtain a 30 percent life extension phenotype in male mice. Our efforts uncovered a regulatory mechanism centered on the single molecule combination of H3.3{K27me2,K36me2} that increases with normal aging by 6 percent (from 2 months to 25 months) and decreases 7 percent in our experimental mice. We observe different transcriptional regulation of the wild-type mouse brain spatially and temporally. During adolescence histone H4 proteoforms indicating higher levels of transcription decrease in the hindbrain between 25 and 47 days of age (15 to 12 percent multiply acetylated H4). CGNs during development (7 to 60 days of age) show changes to histone H1 variants where the abundance of histone H1.4 increases during development (42 to 68 percent) but becomes less phosphorylated (60 to 31 percent). Our top-down quantitation of histone proteoforms has revealed previously inaccessible mechanisms of neuroepigenetics and the means to manipulate the neuroepigenetics of aging.

Parallel Session 10: MS and Non-MS Strategies for Structural Biology

IS10.01 | Using advanced cross-linking MS to investigate the distribution of long-lived proteins within mitochondrial protein complexes in mouse brain and cortex

Ewa Bomba-Warczak, *Northwestern University, Evanston, IL*

Intracellular long-lived proteins (LLPs) provide structural support for several highly stable protein complexes and assemblies that play essential roles in ensuring cellular homeostasis and function. Recently, using using in vivo dynamic metabolic stable isotope (15N) labeling of rodents, we showed that a subset of the mitochondrial proteome can persist for months in mammalian brains. Interestingly, our analysis revealed that mitochondrial LLPs (mt-LLPs) are concentrated at the inner mitochondrial membrane, specifically the sub-compartment of mitochondria called cristae. The exceptional longevity of the mt-LLPs over the course of months prompted a question of how the old and newly synthesized proteins are integrated within mitochondria and whether same peptide chains are recycled and intermixed, or spatially restricted within the organelles. To address this question, we performed a pioneering crosslinking experiments on intact mitochondria isolated from cortex and heart extracts of dynamically 15N-labeled mice. Through a combination of mitochondria immunocapture, advanced cross-linking and hybrid MS2–MS3 fragmentation approach we were able to directly probe for protein-protein interactions within old, new, and mixed mitochondrial complexes and showed that mt-LLPs are not randomly dispersed throughout individual mitochondria, but are rather spatially restricted and co-preserved for months is brains of mice.

IS10.02 | Matrix Landing Mass Spectrometry for 3D protein structure determination

Josh Coon, *UW-Madison, Madison, WI*

In this presentation I will describe modifications to an Orbitrap hybrid mass spectrometer that allow for deposition of protein-protein complex cations onto transmission electron microscopy grids. These samples are then directly imaged using TEM and the three-dimensional structures of the particles solved. Here we describe the roles of the chemical matrix, the use of the mass filter for sample purification, and outline how the technique has potential to advance the exciting field of cryo-EM.

OA10.01 | Structural Elucidation of Endogenous Human Cardiac Troponin Complexes by Native Top-Down Mass Spectrometry

Emily Chapman[1] (eachapman2@wisc.edu), David S. Roberts[1], Timothy N. Tiambeng[1], Jake Melby[1], Kyle A. Brown[1], Jään B. Andrews[1], Man-Di Wang[1], Song Jin[1], Ying Ge[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Purpose: The cardiac troponin heterotrimer complex (cTn; ~77 kDa) plays critical roles in cardiac contractility and is composed of three subunits (troponin I, T, C). Calcium ions (Ca^{2+}) bind to the TnC subunit to induce a conformational change in the cTn complex leading to muscle contraction. Alterations in cTn Ca^{2+} binding are associated with cardiac dysfunction and methods for the structural characterization of endogenous cTn remain challenging due to difficulties in the isolation and stability of cTn under native conditions.
Methods: Proteins were extracted from human cardiac tissues at physiological pH and surface-functionalized nanoparticles were used for cTn enrichment. cTn was directly infused to a Bruker 12 T SolariX FTICR mass spectrometer for ultrahigh-resolution characterization. Collisional cross section (CCS) values were determined by trapped ion mobility spectrometry (TIMS)-MS.
Results: Our native nanoproteomics technique revealed distinct perturbations in endogenous cTn complex stoichiometry and Ca^{2+} binding interactions. Native FTICR-MS/MS complex-down analysis was performed to sequence ejected cTn monomers from the heterotrimer complex for metal-binding localization and proteoform characterization. Our results localized three Ca^{2+} binding sites within the cTn complex, as well as their binding stoichiometry. Native TIMS-MS was used to probe the role of Ca^{2+} in stabilizing the cTn complex. EGTA was used to sequester bound Ca^{2+} ions from the three occupied Ca^{2+} binding sites to investigate the effects of Ca^{2+} binding. Removal of Ca^{2+} yielded more unfolded cTn conformers by TIMS-MS analysis, implying that Ca^{2+} binding is important for cTn stability.
Conclusion: We have developed a “native nanoproteomics” strategy incorporating surface-functionalized nanoparticles and native top-down MS analysis to enrich and characterize endogenous cTn from human cardiac tissues. The structural elucidation of endogenous cTn by this native nanoproteomics strategy provides direct insights into cTn structure and Ca^{2+} binding interactions to understand the molecular mechanisms behind cTn alterations in disease for precision medicine.

OA10.02 | Global Proteome Metastability Response in Isogenic Animals to Missense Mutations and Polyglutamine Expansions in Aging

Xiaojing Sui[1] (xiaojing.sui@northwestern.edu), Miguel A. Prado[2], Joao A. Paulo[2], Steven P. Gygi[2], Daniel Finley[2], Richard I. Morimoto[1]
[1]Northwestern University, Evanston, IL, [2]Harvard Medical School, Boston, MA

The conformational stability of the proteome has tremendous implications for the health of the cell and its capacity to determine longevity or susceptibility to age-associated degenerative diseases. For humans, this question of proteome conformational stability has the additional complexity from non-synonymous mutations in thousands of protein coding genes challenging the capacity of the proteostasis network to properly fold, transport, assemble and degrade proteins. Here, we quantify the proteome-wide capacity to such challenges using the isogenic organism *Caenorhabditis elegans* by examining the dynamics of global proteome conformational stability in animals expressing different temperature-sensitive (ts) proteins or short polyglutamine (polyQ) expansions in the context of biological aging. Using limited proteolysis of native extracts together with tandem mass tag-based quantitative proteomics, we identify proteins that become metastable under these conditions and monitor the effects on proteome solubility and abundance. Expression of different mutant proteins in the same tissue identifies hundreds to a thousand proteins that become metastable affecting multiple compartments and processes in a cell autonomous and non-autonomous manner. Comparison of the network of metastable proteins, however, reveals only a small number of common proteins. The most dramatic effects on global proteome dynamics occur in aging with one-third of the proteome undergoing conformational changes in early adulthood. These age-dependent metastable proteins overlap substantially with ts proteins and polyQ; moreover, expression of polyQ accelerates the aging phenotype. Together, these results reveal that the proteome responds to misfolding one-at-a-time to generate a metastable sub-proteome network with features of a fingerprint for which aging is the dominant determinant of proteome metastability.

Parallel Session 11: Metabolomics, Lipidomics and Glycomics

IS11.01 | Exploring the Molecular Universe of the Human Kidney with MALDI-MSI: From Spatial Metabolomics to Spatial Glycomics

Chris Anderton, *Pacific Northwest National Laboratory – PNNL, Richland, WA*

The Kidney Precision Medicine Project (KPMP) consortium, in part, aims to create a human kidney tissue atlas from evaluating healthy and diseased biopsies. We have developed and optimized matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI)-based spatial metabolomics, lipidomics, and N-glycomics assays for analysis of human kidney biopsy tissues obtained from tissue recruitment sites (TRS) across the USA. The KPMP TRS acquire tissue from patients with varying disease states (e.g., Acute kidney injury, Chronic kidney disease, and Diabetic kidney disease). We have also linked our data to other omics-based analyses being performed within the consortium. By combining data obtained within our tissue integration site (TIS) with data from other TISs, we have begun to identify key metabolite, lipid, and N-glycan markers of cell types and disease states.

IS11.02 | Protein-Protein Interactions in the Cell Glycocalyx

Carlito Lebrilla, *University of California – Davis, Davis, CA*

The glycocalyx is a highly interactive environment composed of glycans on lipids and proteins. Cell-cell interactions are mediated through the glycocalyx, however the characterization of the glycocalyx remains a considerable challenge. In this presentation, new methods for the characterization of the cell membrane will be described. We have developed enrichment methods and liquid chromatography – mass spectrometry techniques that yield the glycolipid and glycoprotein components of the cell membrane. We have employed glycosyltransferase inhibitors that allow us to modify the glycocalyx and create specific glycoforms and glycotypes. We then developed new methods for characterizing glycan-mediated protein-protein interactions. Through these tools, we explore the most comprehensive characterization of interactions in the glycocalyx.

OA11.01 | Biochemical Implications of the TMEM97/ Histatin-1 Interaction

Dominick Pierre-Jacques[1] (dpierr4@uic.edu), Kyung-No Son[1], Dhara Shah[1], Sang Min Lee[1], Vinay K. Aakalu[1], Stephanie M. Cologna[1]
[1]*University of Illinois Chicago, Chicago, IL*

Histatins are a family of endogenous antimicrobial peptides, where Histatin-1 (H1) has been shown to specifically promote corneal epithelial migration and corneal wound healing. Recent work from our group has identified that H1 is an endogenous ligand for TMEM97, also known as the sigma-2 receptor. TMEM97, is a key membrane protein that is primarily located in the endoplasmic reticulum and is involved in integral cellular processes such as, cell migration and cholesterol processing and is also implicated in numerous diseases, including multiple types of cancer and neurological disease. In this study, we sought to understand the biochemical consequences of applying H1 to corneal epithelia using untargeted proteomics and lipidomics. In this study, we used human corneal epithelial (HCE) cells which were cultured using standard procedures followed by phenotypic assays. Both wild-type and TMEM97 knock out cells were treated with H1 or a vehicle control. Lipids were extracted using methyl-tert-butyl ether (MTBE) extraction, where protein lysates were digested using S-Trap. Proteomics and lipidomics were performed using a Thermo Fisher QExactive or Agilent QTOF 6550 respectively. Protein identifications were obtained from Proteome Discoverer and Lipid identifiers were obtained from Lipid Annotator. To establish the methods for H1 treatment in cell culture, we first evaluated cytotoxicity and cell viability and found that H1 is not toxic to HCE upto high micromolar levels. Moreover, HCE cells exhibited enhanced migratory speed when tested in standardized scratch assays after exposure to H1. This finding required the normal expression of TMEM97. Lipidomic analysis comparing WT and TMEM97 KO cells, with and without H1 treatment, revealed changes in fatty acids, and ceramides in addition to modulation of phosphatidylinositol and cardiolipins content. Our current efforts are focused on analysis of proteomic changes, combined with lipidomic changes, as well as understanding how H1 contributes to the altered proteome and lipidome.

OA11.02 | MS-AutoQC – Interactive Dashboard for Realtime Quality Control During Mass Spectrometry Data Acquisition

Wasim Sandhu[1], **Brian DeFelice[1] (brian.defelice@czbiohub.org)**, Sarah Lin[1]
[1]*Chan Zuckerberg Biohub, Santa Clara, CA*

Ensuring collection of high-quality LC-MS/MS data is a time-consuming hurdle in untargeted metabolomics. The use of stable isotope labeled internal standards, spiked into all samples, has become commonplace. Internal standards make it easy to distinguish failed injections or poor sample preparation, despite vast biological diversity between samples. However, interpreting internal standard results in real-time – with minimal person-hours – is a challenging hurdle. Here we present MS-AutoQC, an interactive real-time dashboard for at-a-glance quality control of mass spectrometry data. Within 2 minutes of sample completion, numerous parameters are evaluated to determine sample fitness. If user-defined criteria is not met, the relevant stakeholders for the instrument run are notified automatically.

MS-AutoQC offers interactive visualizations that allow monitoring of nuanced changes such as retention time (RT) shifts, mass accuracy (m/z) drifts, and deviations in overall instrument sensitivity. The RT across samples plot allows rapid visualization of RT shifts, and can be tailored to look at any number of internal standards or samples in parallel or separately. In a similar fashion, the Δm/z across samples plot alerts users to drifts in mass accuracy. Intensity across samples plot reveal instrument sensitivity perturbations. Lastly, MS-AutoQC offers a module for longitudinal comparison of instrument benchmark samples to quantify instrument changes from study to study.
The MS-AutoQC package is easy to install and highly configurable. Its backend infrastructure is built atop open-source Python libraries, and highly generalized to adapt to any workflow or pipeline. MS-AutoQC is installed locally – instrument or server – with the option to be synced to a user's Google Drive. Drive syncing enables remote data QC, secured by Google authentication.
MS-AutoQC offers a fast, straightforward approach to ensure collection of high-quality untargeted LC-MS/MS data, allowing for less time extracting ion chromatograms and more time conducting experiments.

Parallel Session 12: Chemical Proteomics and Drug Recovery

IS12.01 | Network Systems Biology: Chemical Proteomics-Based Mapping of Macromolecular Interactions for Drug Discovery

Andrew Emili, *Knight Cancer Institute - Oregon Health & Science University, Portland, OR*

Knowledge of protein–metabolite interactions enhances understanding of biochemical processes and facilitates drug discovery, but the generation of unbiased data is challenging. We developed a sensitive, high-throughput ligand-purification/mass spectrometry approach to map endogenous small-molecule metabolites associated with essential enzymes, putative transcription factors, and functionally-unannotated proteins in the Gram-negative bacterium, Escherichia coli. We applied structure-based computational modeling to define high-confidence ligand binding pockets, and assessed metabolic pathway relationships and evolutionary conservation to determine functional significance. The resulting interaction network includes hundreds of chemically diverse compounds including reaction substrates, products, analogs and cofactors. This ligand-interactome landscape illuminates gene function, reveals unexpected pathway crosstalk and feedback mechanisms, and suggests scaffolds for antimicrobial leads. Our approach is scalable and may be equally informative for mapping similar interactions, functional dependencies and drug design in other organisms, including human.

IS12.02 | Post-Translational Modification Proteomics in 4D

Amy Weeks, *University of Wisconsin-Madison, Madison, WI*

Post-translational modifications control the structure, activity, localization, and lifetime of nearly all proteins and are often dysregulated in human disease. However, identification of post-translational modifications has far outpaced assignment of their biological functions, an endeavor that requires detailed information about where and when these modifications occur within the cell. We are integrating principles from organic chemistry, protein engineering, and mass spectrometry-based proteomics to develop innovative tools for spatially and temporally resolved mapping of protein modifications in living cells. These technologies will advance our understanding of how post-translational modifications program biological function, leading to the development of new therapeutic hypotheses for the treatment of human disease.

OA12.01 | Deciphering the Therapeutic Accessibility of the Human Cysteineome using Experimental Quantitative Chemoproteomics

Lisa M. Boatner (lisaboatner@ucla.edu), University of California Los Angeles, Los Angeles, CA

Small molecule chemical probes are useful tools for modulating protein function and can serve as leads for future medications. Cysteine chemoproteomics has emerged as an enabling technology for closing the druggability gap, namely achieving pharmacological manipulation of ~99% of the human proteome that remains untargeted by FDA approved small molecules. While recent interactive datasets have allowed users to interface more readily with cysteine chemoproteomics data, these databases remain limited to single studies and therefore do not provide a mechanism to perform cross-study analyses. Here we present CysDB as a curated community-wide repository of cysteine chemoproteomics data. CysDB is a SQL relational database that features experimental measures of identification, hyperreactivity, and ligandability for 62,888 cysteines (24% of all cysteines in the human proteome) and 11,621 proteins, representing a >50% increase in total number of identified cysteine residues from individual prior studies. In addition, the CysDB web application (https://backuslab.shinyapps.io/cysdb/) includes annotations of functionality (UniProtKB/Swiss-Prot, Pfam, Panther), known druggability (FDA approved targets, DrugBank, ChEMBL), disease-relevance and genetic variation (ClinVar, Cancer Gene Census, Online Mendelian Inheritance in Man), as well as structural features (Protein Data Bank). Showcasing the utility of CysDB, we discovered an enrichment of ligandable cysteines in undruggable classes of proteins, a subset of cysteines showing a marked preference for specific classes of electrophiles (chloroacetamide vs acrylamide), and that ligandable cysteines are present in numerous undrugged disease-relevant proteins. Most importantly, we have specifically designed CysDB for the incorporation of new datasets and features to support the continued growth of the druggable cysteineome. We expect CysDB will prove highly enabling or investigating the functions of a wide variety of biologically important and disease-relevant proteins and guiding lead chemistry for drug development programs aimed at targeting novel small molecule binding sites.

OA12.02 | Pathway-Scale Biophysical Mapping Reveals the Proteome–Metabolome Interaction Landscape of Escherichia coli

Andrew Emili (aemili@bu.edu), Knight Cancer Institute - Oregon Health & Science University, Portland, OR

Knowledge of native protein–metabolite interactions enhances mechanistic understanding of biochemical processes and facilitates drug discovery, but the generation of comprehensive data is challenging with existing methods. To address this gap, our group has developed a sensitive high-throughput ligand-purification / mass spectrometry approach to map the physical associations of chemically diverse endogenous small-molecule ligands. We deployed this platform to identify reaction substrates, products, cofactors and other ligands of essential enzymes and transcription factors in *Escherichia coli*. We next performed structure-based computational modeling, evolutionary projections, and independent validation experiments to define high-confidence ligand binding pockets, metabolic pathway relationships, and broader functional conservation. The ligand-interactome landscape we defined reveals unexpected mechanistic associations, modes of biochemical crosstalk, and scaffolds for antimicrobial leads. Our approach is scalable and may be equally informative for illuminating ligand-protein interactions, biomolecular dependencies, and drug design in other organisms.

Parallel Session 13: Advances in Single-Cell MS

IS13.01 | Automated Container-Less Cell Processing Method for Single-Cell Proteomics

Yu Gao, *University of Illinois at Chicago, Pharmaceutical Sciences, Chicago, IL*

Single-cell genomics and transcriptomics studies enabled us to characterize cell heterogeneity in various tissues, which helped us to better understand the biological system and disease progression. Single-cell proteomics, which directly measures the protein expression level, has the potential to further enhance our knowledge by providing not only a more direct measurement but also crucial information cannot be captured by genomics or transcriptomics study, such as protein activation states and post-translation modification events. One of the main challenges of single-cell proteomics is the large sample loss during sample preparation, which is largely unavoidable using standard proteomics protocols. Protein and peptide loss to the container surface is a well-known phenomenon but often overlooked in larger-scale (>1 µg) proteomics experiments. When it comes to single-cell proteomics with only picograms of protein samples, this loss becomes non-negligible and often dictates the outcomes of the experiment. More importantly, sample processing through multiple pipette tips and containers often introduces random errors, which undermine the ability to detect true heterogenous cellular events. To solve these problems and further improve the throughput and reproducibility of the single-cell proteomics experiments, we developed an automated container-less cell processing platform, utilizing acoustic levitation to process cell samples in the air. Our platform automatically performs cell lysis, protein reduction, alkylation, digestion, and peptide labeling in the air, without any sample transfer step or container. The digested and labeled peptides are then directly injected into the capillary LC-MS/MS system for analysis, eliminating manual steps and conserving most of the sample materials for proteomics analysis. Our initial test shows at least 30% improvement in peptide signals over conventional methods. This process can be performed in parallel to further improve sample processing throughput.

IS13.02 | The Journey From Low ng- to Sub-ng-Level Samples and From Small Cell Populations to Single Cells in Enabling Techniques for Bottom-Up, Top-Down, and Native Proteomic and PTM Profiling

Alexander Ivanov, *Northeastern University, Boston, MA*

The field of proteomics research has been rapidly evolving during the last two decades. During the last few years, the field of single-cell proteomics has progressed substantially. However, deep proteomic profiling of limited samples (e.g., small populations of rare cells, individual cells, microneedle biopsies, subpopulations of extracellular vesicles (EVs) isolated from minute volumes of biofluids) and especially, characterization of post-translational modifications (PTMs) and non-covalent protein interactions at such sample amount levels have been a major challenge because of very low abundance and high heterogeneity of complex biological matrices. In this presentation, we will overview a combination of advanced sample preparation, pressure- and electric field-driven ultra-low flow (ULF) high-efficiency nanoscale liquid phase separations coupled to mass spectrometry (MS) via alternative interfacing techniques to evaluate the potential applicability for high sensitivity, robust and reproducible proteomic profiling of individual cells and low ng-/sub-ng-level complex biological samples, using bottom-up, top-down, and native proteomic and PTM profiling approaches.

OA13.01 | Label Free Proteomics Applied to Viable Single Cardiomyocytes

Aleksandra Binek[1] (aleksandra.binek@cshs.org), Blandine Chazarin[1], Ali Haghani[1], Simion Kreimer[1], Jae Hyung Cho[1], Thassio Ricard Ribeiro Mesquita[1], Eduardo Marban[1], Jennifer Van Eyk[1]
[1]*Cedars-Sinai Medical Center, Los Angeles, CA*

Introduction: The cardiomyocyte is a specialized heart cell that is responsible for heart tissue contraction and is easily damaged. These large rod shape cells (average diameter ~25um, length ~100um) are challenging to isolate in large numbers using conventional cell sorting. Our aim is to develop a method to monitor proteome changes of single cells in disease and control state and evaluate their heterogeneity. We optimized isolation of cardiomyocytes using CellenONE system, label free methodology on Bruker manufactured time-of-flight instrument with the ion mobility cell and ion source optics optimized for applications requiring high sensitivity such as single-cell proteomics using a short turn-around time.
Methods: After sacrifice, hearts were digested and perfused with enzyme mix containing collagenase. Cell suspension was sorted on CellenONE instrument to isolate cardiomyocytes into a 384-well plate. Trypsin digested peptides were analyzed using data acquisition with parallel ion accumulation and ion mobility separation (PASEF) on SCP-TIMS-TOF (Bruker). The data was analyzed in FragPipe 17.1 and DIA-NN 18.1.1 using the match between runs label-free quantitation workflow.
Preliminary data: Benchmarking of the proteomics analysis was performed using the n-dodecyl β-D-maltoside (DDM) -lysed mouse cardiac cells resulted in detection of 15654 peptides (1374 proteins) with an average identification of 920 proteins per mouse

cardiomyocyte. Based on the average number of proteins detected in rodent cells we evaluated the single cardiomyocyte protein content is in the 1-10 ng range allowing for adequate quantification of the cardiac proteome with high throughput. Hierarchical clustering analysis of the whole proteome segregated the cardiomyocyte population into 2 main clusters grouping them by cellular diameter into 50-65µm and 76-112 µm groups. Cardiomyocytes with a 66-75µm cellular diameter size were equally segregated between the 2 main populations. Those single cell proteomics applications developed in cardiomyocytes are now being tested on and translated to human iPS cell lines.

OA13.02 | Capillary Electrophoresis Data-independent Acquisition Mass Spectrometry Enables Subcellular Proteomics

Bowen Shen[1] **(bwshen@umd.edu)**, Leena R. Pade[1], Fei Zhou[1], Peter Nemes[1]
[1]*University of Maryland College Park, College Park, MD*

Our goal was to enable the proteomic characterization of a cell undergoing cell division in a live vertebrate (frog) embryo. The commitment of the pluripotent cell to specific tissue fates is essential to cell and embryonic development. Although various approaches have been developed during the last few years to measure proteomic differences between embryonic cells in amphibian and mammalian cells, how a precursor cell cleaves into two heterogeneous descendent cells during early embryonic development is still unknown. In addition, current single-cell proteomics analysis based on liquid-chromatography high-resolution mass spectrometry (LC-HRMS) requires too large starting materials for analysis. To overcome this technical limitation, we coupled microanalytical capillary electrophoresis (CE) with electrospray ionization (ESI) time-of-flight (TOF) mass spectrometry executing data-independent acquisition parallel accumulation–serial fragmentation (diaPASEF). This CE-diaPASEF approach identified ~1,100 protein groups from 200 pg of HeLa digest in 45 min separation, which is ~50-times higher sensitivity and three times faster analytical throughput than reported by HPLC-HRMS. This sensitivity allowed us to explore the cellular and subcellular proteomic content of identified cells in live embryos of the South African clawed frog (*Xenopus laevis)*. Opposite locations (future D11 and D12 poles) within the same D1 cell of an 8-cell *Xenopus* embryo were analyzed, measuring 1,400+ protein groups from ~5 ng, viz <0.025% of the subcellular proteome content. A total of 171 proteins were significantly (p=0.05) differently enriched between the cellular poles, suggesting subcellular heterogeneity leading up to cell division. The following single-cell proteomics analysis of the D11 and D12 cells demonstrated that the protein enrichment either persisted or reversed after cell differentiation. Our CE-diaPASEF approach lays the foundation for subcellular proteomics analysis and provides valuable information on cell formation of molecular cell heterogeneity during early embryonic development.

Parallel Session 14: Quantitative Proteomics for Disease Research

IS14.01 | Global Protein-Turnover Quantification in Escherichia coli Reveals Cytoplasmic Recycling under Nitrogen Limitation

Martin Wuhr, *Princeton University, Princeton, NJ*

Protein turnover is a critical regulatory mechanism for proteostasis. However, proteome-wide turnover quantification is technically challenging, and, even in the well-studied E. coli model, reliable measurements remain scarce. Here, we quantify the degradation of ~3.2k E. coli proteins under 12 conditions by combining heavy isotope labeling with complement reporter ion quantification and find that cytoplasmic proteins are recycled when nitrogen is limited. We used knockout experiments to assign substrates to the known ATP-dependent proteases. Surprisingly, we find that none are responsible for the observed cytoplasmic protein degradation in nitrogen limitation, suggesting that a major proteolysis pathway in E. coli remains to be discovered. Lastly, we show that protein degradation rates are generally independent of cell division rates. Thus, we introduce broadly applicable technology for protein turnover measurements. We provide a rich resource for protein half-lives and protease substrates in E. coli, complementary to genomics data, that will allow researchers to decipher the control of proteostasis.

IS14.02 | Chemistry and Biology of Protein Poly-ADP-Ribosylation

Yonghao Yu, *Columbia University, New York City, NY*

PARP1 is a nuclear enzyme that is critically involved in mediating DNA damage response (DDR). Its enzymatic function is to catalyze a protein posttranslational modification (PTM) known as Poly-ADP-ribosylation (PARylation). During genotoxic stress, PARP1 is recruited to nicked DNA and is rapidly activated, resulting in the synthesis of many PARylated proteins and initiation of the DNA repair mechanisms. It has been proposed that PARP1 inhibitors (PARPi) kill tumors via two distinct but interconnected mechanisms (i.e., PARP1 inhibition and PARP1 trapping). However, all FDA-approved PARPi possess both of these two activities, and their relative contribution to the cytotoxicity of PARPi is poorly understood.
We recently developed a large-scale mass spectrometric approach for the proteome-wide and site-specific characterization of Asp/Glu-PARylation (Zhang et al., Nature Methods, 2013; Gibson et al., Science, 2016; and Zhen et al., Cell Reports, 2017). Using this technology, we comprehensively characterized the downstream signaling network of PARP1 and found that the PARylated proteins are involved in not only DDR, but a wide array of other nuclear functions.
To understand the role of PARP1 trapping, we recently developed a series of small molecule degraders of PARP1 (Wang et al., Nature Chemical Biology, 2019). Treatment with one such compound, iRucaparib-AP6, results in highly efficient and specific PARP1 degradation in primary neonatal rat cardiomyocytes. iRucaparib-AP6 blocks the enzymatic activity of PARP1 in vitro, and PARP1-mediated PARylation signaling in intact cells. These “non-trapping” PARP1 degraders mimic PARP1 genetic depletion, which enables the pharmacological decoupling of PARP1 inhibition from PARP1 trapping. Using these unique compounds, we showed that PARP1 trapping is a key determinant of the DNA damage, cytotoxic and immunomodulatory effects of PARP1 inhibitors (Kim et al., eLife, 2020 and Kim et al., Cell Chemical Biology, 2021). In summary, these compounds represent ‘non-trapping’ PARP1 degraders that block both the catalytic activity and scaffolding effects of PARP1, providing an ideal approach for the amelioration of the various pathological conditions caused by PARP1 hyperactivation.

OA14.01 | Preclinical validation of molecular markers in circulating extracellular vesicles for non-invasive detection and monitoring of renal cell carcinoma

Marco Hadisurya[1] (mhadisur@purdue.edu), Zhuojun Luo[1], Xiaofeng Wu[2], Anton B. Iliuk[3], Ronald S. Boris[4], W. Andy Tao[1]
[1]Department of Biochemistry, Purdue University, West Lafayette, IN, [2]Department of Chemistry, Purdue University, West Lafayette, IN, [3]Tymora Analytical Operations, West Lafayette, IN, [4]Department of Urology, Indiana University School of Medicine, Indianapolis, IN

Currently, most cancer diagnosis procedures include a diagnostic imaging process, such as a CT scan, followed by a tumor biopsy. Tissue biopsy is an invasive and painful procedure and may pose health risks for patients such as those with kidney diseases. Liquid biopsy, the ability to detect and monitor disease through biofluids, is highly promising and may replace tissue biopsy with an immense potential public health impact. The use of liquid biopsy offers numerous advantages in the clinical setting, including its non-invasive nature, a suitable sample source for longitudinal disease monitoring, a better screenshot of tumor heterogeneity, and lower costs compared to tissue biopsy. Increasing evidence indicates an important cellular function of exosomes and other extracellular vesicles (EV) particles in tumor biology and metastasis, presenting them as intriguing sources for biomarker discovery

and disease diagnosis. However, most current exosome/EV studies focus on their miRNAs, with few studies on functional proteins such as phosphorylated proteins. As phosphorylation is a major player in cancer and other disease progression, EV phosphoproteins are expected to become actively pursued targets for *in vitro* disease diagnosis. In this project, we focused on non-invasive RCC detection by coupling CT scans. Recent equipment upgrades and mechanistic streamlining allowed for more rapid throughput of tissue and fluid analysis. Using tissue, plasma, and urine samples from 121 patients at different stages (low and high grades), 121 healthy controls, and 66 CKD as controls, we successfully validated a previously identified panel of EV phosphoprotein RCC biomarkers. We also successfully integrated imaging data with the biomarkers and derived a correlation between these biomarkers and the imaging results for RCC diagnosis. These results show that non-invasive molecular markers in EVs from isolated urine and plasma can be coupled with imaging scan results, potentially eliminating the need for surgery by more than 50%.

OA14.02 | Inter-cellular communication within a virus microenvironment drives the susceptibility of host cells to super-infections

Bokai Song[1] **(bokais@princeton.edu)**, Xinlei Sheng[1], Joshua L. Justice[1], Peter Metzger[1], James Kostas[1], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ*

Upon a virus infection, intracellular and intercellular communications drive the formation of a virus microenvironment (VME) that dictates the outcome and pathology of a given infection. The complex communication within a VME includes host defense responses, such as immune signaling. Additionally, evidence points to the ability of viruses to exploit multicellular communication to promote viral spread. However, given technological challenges, how an infected cell communicates with proximal or distal cells and the resulting balance of these rivaling anti- and pro-viral responses have remained unclear. Here, we performed a first-of-its-kind study to characterize a virus microenvironment. We discover a new mechanism through which an infected cell impacts its environment and demonstrate its effect on super-infections with multiple human viruses. To resolve molecular features and functional consequences of cell spatial address within the VME, we design a dual fluorescence labeling approach, integrated with FACS, quantitative mass spectrometry, microscopy, and virology assays. With an initial focus on the VME induced by human cytomegalovirus (HCMV) infection, we distinguish and characterize the proteome compositions of infected, proximal, and distal cells. We uncover signature changes in proximal (compared to distal) cells, including interferon-inducible, extracellular matrix, and cell cycle-related proteins. Indeed, subsequent cell cycle profiling demonstrates the striking dysregulation of mitosis in proximal cells. Parallel reaction monitoring (PRM) MS analysis of cell cycle-related proteins further reveals the checkpoints and regulation pathways responsible for the cell cycle disruption. Demonstrating the broad impact of these changes, we discover that proximal cells are less resistant to superinfections with a range of viruses, including HCMV, herpes simplex virus type 1 (HSV-1), and influenza A.

Overall, our study shows how a virus infection induces intra- and inter-cellular communications that reshape the surrounding environment in a proximity-dependent manner to facilitate the spread of an infection, as well as the susceptibility to new infections.

Poster Presentations

US HUPO 2023

Chicago, IL | March 4-8

P1: Advances in Technology

P1.01 | Streamlined and semi-automated proteomics pipeline from protein extraction to data analysis by coupling the BeatBox with the PreON automation platform

Measho Abreha[1] (Measho.Abreha@preomics.com), Katrin Hartinger[2], Jasmin Johansson[2], Zehan Hu[2], Jonathan R. Krieger[3], Katharina Limm[2], Nils A. Kulak[2]
[1]Preomics, [2]PreOmics GmbH, Planegg/Martinsried, Germany, [3]Bruker Canada, Toronto, ON

Introduction: Sample preparation is an often overlooked yet crucial component of successful LC-MS-based proteomics studies. Conventional manual workflows often have various limitations such as low throughput, insufficient reproducibility and time-consuming protocols. Here, we present a streamlined workflow from sample homogenization to data analysis by combining highly efficient protein extraction on the BeatBox platform with the robust iST sample preparation, fully automated on the PreON platform (PreOmics). Coupling this workflow to the newly developed PaSER technology (Bruker) provides improved and accelerated data processing by making the results available immediately after the run and eliminating the need for separate processing on high-performance computers.
Methods: The described sample preparation workflow was evaluated with human and bacterial cell lines as well as with various mammalian tissue types. Protein extraction was performed on the BeatBox using sample-specific settings, followed by automated digestion and peptide clean-up on the PreON applying the iST protocol. This workflow minimizes technical variability, sample loss and dramatically reduces hands-on time compared to manual processing. Data were processed in real time using PaSER, effectively generating a streamlined workflow from sample preparation to real-time results.
Results: From intact sample to finished data analysis (DDA on timsTOF Pro with PaSER data processing, Bruker) in less than one working day, we identified ~4900 protein groups from HEK cells and ~2500 protein groups from mouse liver tissue in just 1-hour total LC-MS time. At the same time, high sample quality with excellent digestion efficiency and low technical variability (median CV < 15 %) was obtained. Comparing the PaSER output with a traditional MaxQuant off-line search revealed similar results.
Conclusion: The presented workflow sets a new standard in speed, robustness and protein extraction efficiency in LC-MS based proteomic sample preparation. The use of PaSER for real-time data processing reduces the overall time required even further.

P1.02 | Comprehensive Sample Preparation for Proteins Enabled by AFA® Technology

Debadeep Bhattacharyya[1] (dbhattacharyya@covaris.com), Tom O'hare[1]
[1]Covaris Inc, Woburn, MA

Introduction: Robust, reliable sample preparation approaches usually offers the best possible chances to get the most from precious samples. Covaris' Adaptive Focused Acoustics® (AFA®) Technology can address a wide variety of sample types (FFPE, LCM, fresh tissue, cells, bacteria, yeast) for protein extraction, purification, digestion with higher speed, reproducibility, and reliability. AFA can further enhance the purification and digestion of proteins in conjunction with other well-used purification approaches like SP3 or S-Trap.
Methods: Samples obtained from different matrices were evaluated. For FFPE samples, an active deparaffinization (without any organic solvent/mineral oil) with AFA was used for protein extraction, depletion, digestion, and post-digestion clean-up. The workflow can be easily implemented in an 8-strip or 96-well plate depending on throughput requirements.
Results: Significant benefits were observed for proteomics studies with Tandem Mass Tag (TMT) based labeling. The TMT based labeling were found to be significantly faster (reduction from traditional workflows of ~48 hours to 1.5-3 hours) and simpler. A 50 step process was reduced to 24 step process. In addition, increased number of proteins and peptides were identified and quantified using for every bottom-up assay. In addition, enhanced binding and digestion efficiencies were observed with AFA for some of the assays.
Conclusion: The data in this study demonstrates the ability of AFA in achieving increased confidence in protein analysis by starting with complex biological matrices, regardless of the nature and type of proteins (phosphoproteins, membrane proteins, hydrophobic peptides, etc.).
References: Hughes et al, Nat Protoc. 2019 Jan;14(1):68-85; Hughes et al, Mol Syst Biol. 2014 Oct 30;10(10):757; Mueller et al, Mol Syst Biol. 2020, 16:e9111; Schweitzer et al, Covaris Application Note M020141; Covaris Application Note M020154; Staeber et al, Covaris Application Note M020155; Herrera et al, Clin. Proteom. 2020, 17:24; Mueller et al, Mol. Syst. Biol. 2020,16:e9111; Herrera et al.

P1.03 | Fast, High-Yield, and Universal Proteomics Sample Preparation Using the Reversible Protein Tag ProMTag

Svitlana Yablonska[1] (svitlana@impactproteomics.com), Stephanie Biedka[1], Jonathan S. Minden[1]
[1]*Impact Proteomics, Pittsburgh, PA*

High quality protein extraction and cleanup during sample preparation is critical for achieving the coverage, yields, and reproducibility required for a successful proteomics experiment. Many current technologies for proteomics sample preparation rely on precipitation or filtration-based technologies, which suffer from sample loss, long processing times, incompatibility with certain lysis buffers, and limitations for automation. At Impact Proteomics, our Universal Protein Extraction and Cleanup Kits (UPECK) have overcome these limitations. Our novel protein sample preparation method utilizes a protein-specific tag (ProMTag) where one end of the tag forms a reversible, covalent link to proteins and the other end irreversibly anchors the protein to a capture resin using a fast-acting, click chemistry reaction. The ProMTag reaction is not affected by high detergent, denaturant, or salt conditions accommodating harsh cell lysis conditions. Once ProMTagged proteins are covalently linked to the capture resin, contaminating salts, detergents and other molecules that interfere with MS analysis are removed by rapid washing steps. Following washing, the ProMTag is reversed, releasing the protein in its original, unmodified form. To yield peptides ready for MS analysis, modified MT-Trypsin, which also binds to the capture resin, is added and digestion is complete in one hour. The UPECK workflow ensures fast, high yield, and streamlined sample preparation that yields highly reproducible results and requires about four hours to complete with less than an hour of hands-on time. Here, we demonstrate compatibility of the UPECK Peptide workflow with a wide variety of sample types, including cultured cells, various tissues from brain to muscle, and bodily fluids, all using a universal lysis and solubilization buffer.

P1.04 | MetaDrive: An Instrument-Agnostic Mass Spectrometry Control Program to Improve Top-Down Proteoform Identification

Nicholas Bollis[1] (nbollis@wisc.edu), Austin Carr[1], Mark Scalf[1], Lloyd M. Smith[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Top-down proteomics has gained significant interest in biomedical research for structural elucidation and analysis of proteoforms. Data-dependent acquisition (DDA) of intact proteoforms is the leading acquisition technique. However, it has several limitations including proteoform intensity being spread across multiply charged product ions and their respective isotopologues, different charge states often having distinct fragmentation spectra, and fragment coverage being insufficient to confidently localize and/or identify a post-translational modification leading to an ambiguous identification. The ability to construct real-time customized data acquisition workflows could vastly improve top-down proteomics. Presented here is MetaDrive, an instrument-agnostic program that grants users full control of their data acquisition workflow. Previously developed programs are limited in scope as they aim to improve only DDA and lack in extendibility as the code is not publicly available and is written to operate a single instrument. MetaDrive workflows are constructed by the user through a unique combination of predefined tasks in a graphical user interface, removing the need for substantial expertise in computer science to create novel workflows. Two such workflows are whole charge state envelope isolation and proteoform selected reaction monitoring (SRM). Isolating and fragmenting entire charge states, as opposed to traditional DDA selecting the most abundant species in a spectrum, increases the number of unique fragmentation products, in turn improving the confidence of proteoform identifications. To perform proteoform SRM, a traditional DDA experiment is performed and an inclusion list of ambiguous proteoforms is constructed for subsequent SRM analysis of the same sample. MetaDrive is easily extendable with well-defined interfaces for adding unsupported instruments and new tasks, allowing the program to grow with the proteomic community.

P1.05 | An In-Source Fragmentation Data Acquisition Strategy Increases Proteoform Identifications in Complex Samples

Austin Carr[1] (avcarr2@wisc.edu), Mark Scalf[1], John Pavek[1], Michael R. Shortreed[1], Lloyd M. Smith[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Top-down proteomics most commonly employs a data dependent acquisition (DDA) strategy. Because DDA methods select the most abundant peaks in a precursor scan for MS2 fragmentation, large proteins (> 30 kDa), which have low signal-to-noise ratios due to isotopic and charge state dilution, are frequently not selected for fragmentation. Additionally, due to charge state-dependent differences in fragmentation efficiency, DDA methods may require multiple MS2 scans to identify a single proteoform. To address shortcomings in current DDA-based approaches, we have developed MS/in-source fragmentation MS (MS/isfMS), a data-independent fragmentation strategy for top-down proteomics. We apply MS/isfMS to a series of increasingly complex protein mixtures, demonstrating that MS/isfMS outperforms DDA methods in protein and proteoform identifications. Using single standard

proteins and simple mixtures, we demonstrate the advantages of using MS/isfMS over current, comparable DDA methods for proteoform identifications. We further analyzed size-based fractionated yeast lysates with MS/isfMS. MS/isfMS methods identified 165% more protein spectral matches, 25% more protein groups and 114% more proteoforms than a top-5 DDA method. Compared to DDA, MS/isfMS was able to identify more proteins above 30 kDa, suggesting the new method may be a viable approach for the analysis of larger proteins. Finally, we show that the faster duty cycle of the MS/isfMS methods improves label-free MS1 quantitation due to a higher number of sampled points across a chromatographic peak.

P1.06 | Evaluation of an unbiased and scalable multi nanoparticle-based deep proteomics workflow for limited-volume plasma analysis in small model organisms

Shao-Yung Chen[1] **(schen@seer.bio)**, Lucy Williamson[1], Margaret Donovan[1], Purvi Tandel[1], Gabriel Castro[1], Taher Elgierari[1], Aaron S. Gajadhar[1]
[1]*Seer Inc., Redwood City, CA*

Blood plasma and serum are rich, readily available sources of protein that are commonly used in clinical research studies. However, the large dynamic range in the plasma/serum proteome has hindered large scale plasma/serum proteomic research. The recently introduced Proteograph™ Product Suite (Seer Inc.) enables high-throughput in-depth plasma proteome quantification, employing a panel of engineered nanoparticles (NPs) with distinct physicochemical properties. This collective panel of NPs provides optimized protein identifications in terms of depth, breadth, and analysis reproducibility. Model organisms like mice, rats, pigs, and monkeys, are utilized to unveil insights to human biology research. However, due to the nature of these smaller organisms the available blood sample volume is often reduced, which also makes repeated sampling over longitudinal studies more challenging compared to sampling from human donors.

Here we evaluate the current performance of Proteograph Product Suite coupled with Thermo Fisher Scientific Exploris 480 Mass Spectrometer when a limited sample volume is utilized. We evaluated 250, 125, 50, 25, and 10 ul of starting volume for (1) human plasma, and (2) mouse serum using Proteograph Product Suite. We investigated the LC-MS/MS performance using 30-minute Liquid Chromatography (LC) methods using data-independent-acquisition strategies (DIA) and analyzed the data via Proteograph™ Analysis Suite 2.0, evaluating depth of proteome coverage, dynamic range, assay yield, and reproducibility of Proteograph proteome profiling. Notably, with 5-fold lower than standard Proteograph input volume (50 uL vs 250 uL), the Proteograph obtained ~3.5-fold increase in protein group IDs (2129 vs 627, n=3) compared with traditional neat digestion. This reduced volume workflow improved coverage of low abundance proteins compared to neat digestion while maintaining levels of quantification precision shown by the standard Proteograph protocol. In summary, we demonstrate the current performance and trade-off when limited starting materials are utilized, while balancing minimum injection mass for standard Nanoflow LC-MS/MS workflows.

P1.07 | High-resolution Photobleach Labeling Method for Imaging MS in Mature and Embryonic Zebrafish Models

Julia Ekiert[1] **(jekiert2@uic.edu)**, Xinghang Jiang[1], Beatriz Ibarra[1], Angela Mahaffey[1], Ankur Saxena[1], Yu Gao[1]
[1]*University of Illinois at Chicago, Chicago, IL*

Studying the proteome of in vitro and ex vivo tissue models bring us closer to physiological reality and are crucial in elucidating molecular mechanisms and biological functions that participate in various pathologies. A long-standing challenge in the field of proteomics has been the isolation of proteins from specific cells with minimal morphological alterations and high precision. Currently, laser capture microdissection (LCM) paired with mass spectrometry-based proteomics is used for mapping protein expression in tissue sections and studying spatial relationships. Although effective, there are still limitations to LCM such as hindered optical resolution, loss of fluorescence signal, and staining and tissue fixation methods that are incompatible with downstream applications including proteomics. Additionally, current imaging MS techniques such as MALDI-MSI face the challenge of mainly detecting peptides and metabolites, but not proteins. Here, we developed a non-invasive, high-resolution cell and tissue labeling method with photobleach encoding in both mature and embryonic zebrafish models. When coupled with flow cytometry and LC-MS/MS, we are able to separate cell layers of interest, while maintaining their integrity and specificity, and perform quantitative proteomics. This new protocol uses a combination of ATTO NHS-ester dyes (425 - 647 nm) to stain protein in tissue slices (14 - 60 µm thickness). Layers of tissue are then labeled by photobleaching using a fluorescent microscope. Using this method, we were able to successfully stain and photobleach sections of tissue in the mature zebrafish model with a resolution as low as 3 µm. This further allows us to isolate cells of interest with high spatial resolution for proteomic analysis with LC-MS/MS. Lastly, we aim to refine our protocol in the embryonic zebrafish model, allowing us to omit tissue fixation (avoiding any errors that affect staining quality) and then photobleach freshly stained tissue for accurate downstream analysis of spatially resolved protein expression.

P1.08 | Unbiased human biofluids analysis using a scalable, deep, automated, multi-nanoparticle-based proteomics workflow

Aaron S. Gajadhar[1] (agajadhar@seer.bio), Shao-Yung Chen[1], Lucy Williamson[1], Taher Elgierari[1]
[1]Seer Inc., Redwood City, CA

Introduction: Biospecimen proteomics including characterization of non-blood biofluids like urine, cerebral spinal fluid (CSF), and cell line conditioned media (CM) has the potential to reveal new insights to human health and disease, while serving as a less-invasive and longitudinal sampling source. While recent advances in sample collection and mass spectrometry have deepened our understanding of the proteomes for these samples, the field is plagued by non-standardized sample preparation and analytically complex workflows to characterize proteomes at acceptable depth. In this work, we evaluated the utility of the Proteograph™ workflow, a standardized, automated multi-nanoparticle-based deep plasma proteomics approach, to interrogate a variety of conditions and sample types spanning urine, CSF, and CM.
Methods: To evaluate the performance of the Proteograph workflow, biospecimens were sourced from commercial biobanks to create a representative panel of healthy and diseased samples. For human CSF, we profiled samples from normal donors as well as donors with Alzheimer's disease, Amyotropic Lateral Sclerosis, and Parkinson disease. Human urine samples were profiled from normal donors and donors with bladder cancer, T-cell lymphoma, and prostate cancer. To assess in vitro secretome models, CM samples from breast, colon, cervical, and prostate cancer cell lines were analyzed. Samples were processed directly using the Proteograph™ Product Suite (Seer Inc.), and in parallel with conventional sample prep techniques as a control. Tryptic peptides were analyzed by DIA LC-MS analysis using an Orbitrap Fusion Lumos MS, and data processing was performed using Proteograph™ Analysis Suite.
Preliminary Results: We demonstrate the feasibility of performing biofluid and CM analysis with the Proteograph workflow in a rapid, deep, and unbiased fashion with improved depth of coverage compared to conventional sample preparation.
Conclusions: The Proteograph workflow simplifies analysis of biofluids without the need for complex sample manipulation and fractionation, offering a robust solution for proteomic and biological insights.

P1.09 | Nano-omics: nanotechnology-enabled harvesting of blood-circulating biomarkers

Marilena Hadjidemetriou (marilena.hadjidemetriou@manchester.ac.uk), University of Manchester, Manchester, United Kingdom

The quest for novel blood biomarkers has led to the development of nanotechnology-based blood analysis solutions. We have recently introduced the *'Nano-Omics'* paradigm, to describe the nanotechnology-enabled enrichment and analysis of blood-circulating molecular biomarkers. Nano-Omics utilizes nanoparticles (NPs) as scavenging platforms to capture, enrich and isolate disease-associated analytes from biological fluids for downstream omics analyses. Specifically for protein biomarker discovery, Nano-Omics exploits the spontaneous and untargeted adsorption of proteins onto the NP surface once in contact with biological fluids, known as *'protein corona'* formation. Recovery and purification of corona-coated NPs from unbound proteins and subsequent analysis by LC:MS/MS addresses the issue of albumin masking and offers substantial 'broadening' of the blood proteome coverage. This results in the identification of low molecular weight and low abundance proteins that cannot be directly detected by conventional proteomic analysis of blood. Comprehensive comparison between *'healthy'* and *'diseased'* coronas enables the identification of multiple previously unrecognized candidate biomarkers. While the *ex vivo* corona formation (upon incubation of NPs with biofluids in a tube) has been exploited for the analysis of human clinical samples, the molecularly richer *in vivo* corona (forming upon intravenous administration of NPs and their subsequent recovery from blood) has been shown to enable the discovery of biomarkers in preclinical models. More recently, the NP protein corona formation has conceptually morphed into the multi-molecular self-assembly of layers composed of proteins, lipids, polysaccharides and nucleic acids, termed the *'biomolecule corona'*. For example, we demonstrated the interaction of cfDNA with lipid-based NPs upon their incubation with human plasma samples. The discovery of this additional omics dimension paves the way for further investigations of the potential exploitation of the NP biomolecule corona to enrich proteogenomic biomarkers in blood. The Nano-Omics platform technology could be deployable across a range of biomarker applications and pressing clinical challenges.

P1.10 | Fast deep-plasma proteomic platform by combining Fluent automated liquid handling and diaPASEF mass spectrometry

John Laycock[1] (john.laycock@tecan.com), Claudia Martelli[2], Nils A. Kulak[3], Manuel Bauer[4], Fabian Wendt[4], Gary Kruppa[5]
[1]Tecan SP, Baldwin Park, [2]Bruker Daltonics, Fällanden, Switzerland, [3]PreOmics GmbH, Planegg/Martinsried, Germany, [4]Tecan, Männedorf, Switzerland, [5]Bruker Daltonics, Billerica, MA

Introduction: Plasma is the most available clinical specimen with a minimal collection invasiveness and represents a rich source of

disease-biomarkers. Discovering protein biomarker candidates in human plasma is exceptionally challenging due to the high complexity and the wide dynamic range of. Especially for large cohort studies, reproducible sample preparation poses a challenge. We have therefore developed a pipeline consisting of automated liquid handling and next generation diaPASEF mass spectrometry designed for the comprehensive characterization of the plasma proteome at high-throughput and high-sensitivity.
Methods: The plasma sample preparation was automated using theFluent automation workstation (Tecan) in combination with the Resolvex® A200 positive pressure workstation (Tecan).The Fluent was equipped with the eight-channel Flexible Channel Arm™ mounting Disposable Tips (DiTis) and a Robotic Gripper Arm™, as well as on-deck modules enablingsample/reagent storage, magnetic separation, and heating/shaking. The sample preparation workflow covers the protein enrichment, the iST-based protein denaturation, digestion and peptide clean-up for a total of 96 samples/run(PreOmics). Peptide samples were analyzed with a 45-min gradient on the nanoElute system (Bruker) coupled to thetimsTOF Pro 2 (Bruker). Data acquisition was performed with optimized diaPASEF schemes and the results were elaboratedusing the DIA-NN algorithm implemented in the PaSERsoftware (Bruker).
Results: The performance of the presented pipeline was evaluated by processing healthy donor plasma samples in a 96-well plate format. With the optimized comprehensive automatedproteomic platform, we quantified a wide range of protein groups from 300 ng of loaded sample with excellent reproducibility, robustness and minimal user intervention.
Conclusions: By combining the quality of the iST technology with the throughput of the Fluent automation workstation, and the quantitation power of diaPASEF we established a fast, simpleand standardized workflow for deep plasma proteomics characterization.

P1.11 | A high-throughput and robust multi nanoparticle-based label-free mass spectrometry workflow for deep plasma proteomics at scale

Biao Li[1] (bli@seer.bio), Veder Garcia[1], Tianyu Wang[1], Fredric Murolo[1], Alex Kessler[1], Susan Dang[1], Kevin Quach[1], Ray Schmidt[1], Lucy Williamson[1], Kate Zhao[1], Purvi Tandel[1], Evangelina Bahu[1], Gabriel Castro[1], Rebecca Kiss[1], Taher Elgierari[1], Bryn Levitan[1], Mrittika Bhattacharya[1], Xiaoyan Zhao[1], Khatereh Motamedchaboki[1], Ryan W. Benz[1]
[1]Seer, Redwood City, CA

There is great interest in analyzing deep proteome data generated from human blood plasma to assess the health status of individuals. However, the large dynamic range of circulating proteins combined with the diversity of proteoforms present in plasma have limited the comprehensive characterization of the plasma proteome in a high-throughput manner. To help address this challenge, Seer has developed a multi-nanoparticle-based platform that facilitates deep and broad plasma proteomic measurement at scale. This approach enables the quantification of thousands of proteins from plasma without compromising depth, throughput, or reproducibility, creating a unique opportunity to detect protein biomarkers for complex diseases in an unbiased and robust manner. Here, we evaluate the performance of this workflow with a set of control plasma samples and highlight the reproducibility and depth of proteomic coverage provided by the platform. Control pooled human plasma samples were processed with three Proteograph™ instruments, across two days, resulting in a total of 6 batches. The samples were enriched with multiple nanoparticles to produce tryptic digested peptides for downstream LC-MS analysis using Data Independent Acquisition (DIA) on an Orbitrap Exploris 480 MS with a 30-minute LC gradient. Data visualization and statistical assessment of the performance and reproducibility of the assay were performed including quantified peptide mass, peptide intensity CVs, peptide and protein group counts, and Jaccard Index (JI) overlap across samples.

Based on preliminary data, the Proteograph workflow resulted in >15,000 unique peptides and >2,500 unique protein groups covering proteins broadly and deeply across the Human Plasma Proteome Project (HPPP) database. Protein group JI overlap between replicate pairs were typically >0.8 for both intra and inter-batch comparisons, and peptide intensity CVs were <25% and <30% for intra and inter-batch comparisons, respectively. These results demonstrate advancement in enabling large scale plasma proteomics without compromising depth, throughput, or reproducibility.

P1.12 | Single-Molecule Characterization of PARP1 Proteoforms Reveals New Biomarkers of Response to PARP1 Inhibition

Philip L. Lorenzi[1] (PLLorenzi@mdanderson.org), Jibin Ding[1], Deborah Park[2], Wai Kin Chan[1], James Joly[2], Rosemary Wang[2], Parag Mallick[2], Rehan Akbani[1], John N. Weinstein[1]
[1]The University of Texas MD Anderson Cancer Center, Houston, TX, [2]Nautilus Biotechnology, San Carlos, CA

Poly(ADP-ribose) polymerase 1 (PARP1) modulates DNA repair and genome stability and is associated with a number of cell death pathways. PARP1 inhibitors (PARPi) have been approved for the treatment of cancer, but resistance is a major challenge clinically. Here we aimed to elucidate novel associations between PARP1 proteoforms and the anticancer activity of PARP1i *in vitro*. We first analyzed public data on the NCI-60 cell line panel for PARP1 expression and response to PARPi. Using the MCF7 and OVCAR-8/ADR lines as models of sensitivity and resistance to PARPi, the Bio-Techne Peggy Sue nanoimmunoassay platform, and *PARP1* siRNA to validate antibody specificity we characterized 25 antibodies for the ability to detect PARP1 proteoforms. To generate new

hypotheses about mechanisms of sensitivity and resistance to PARPi, we optimized a protocol and analytical conditions to examine PARP1 proteoform response to treatment with vehicle (negative control), BMN-673 (PARP1 inhibitor), methylnitronitrosoguanidine (MNNG; positive control for the induction of parthanatos), or staurosporine (positive control for induction of PARP1 cleavage). We identified several antibodies that, in aggregate, were capable of detecting a range of PARP1 proteoforms—9 proteoforms under denaturing conditions and 36 proteoforms under non-denaturing conditions. Notably, a 124 kDa proteoform of PARP1, which may represent ubiquitinated PARP1 as supported by our antibody characterization data, was ablated by BMN-673 in MCF7 cells but not OVCAR-8/ADR cells. These findings prompt the hypothesis that activity of ubiquitin-proteasome system may be a predictive biomarker of sensitivity/resistance to PARPi. Further studies are underway to elucidate PARP1 proteoform expression at the single-molecule level to gain further insight into possible mechanisms of sensitivity and resistance to PARPi. Our results suggest that specific PARP1 proteoforms or patterns of modification may serve as valuable biomarkers of response to therapy with PARP1 inhibitors.

P1.13 | Lysate to Chip: Preparation of high-density and scalable protein arrays for single-molecule proteomic studies

Parag Mallick (paragm@nautilus.bio), Nautilus Biotechnology, San Carlos, CA

Introduction: We have created a novel, scalable system with minimal sample input for single-molecule protein interrogation with wide dynamic range. This system has two components: a mono-disperse DNA nanoparticle with a single protein attachment site and a nanoscale-patterned surface with millions of homogeneous binding sites. We have created the first large scale, super-Poisson single protein arrays.
Methods: Lysate is processed through a denaturing workflow employing two high efficiency chemistries (a cysteine targeted cross-linking and a two-reaction sequence with lysines) to attach single proteins to a DNA nanoparticle. These nanoparticles are deposited onto a nanostructured array. The nanoarray, created using standard lithography techniques, is comprised of DNA binding sites surrounded by HMDS. To demonstrate single-molecule occupancy of the nanostructured array, the nanoparticle structures are labeled with one of two different dyes, mixed 1:1, and loaded onto the array for imaging. By counting the features with only one wavelength (single loading) and both wavelengths (multiple loading), the single-molecule occupancy of the array could be calculated.
Results: From 1 ug lysate we obtain trillions of single-polypeptide-nanoparticle conjugates that overlap with proteins observed by mass spectrometry spanning a variety of protein classes. When these conjugates were loaded on a 25 mm x 25 mm array with ~75M features, 84% of the features were occupied, with fewer than 4% showing multiple occupancy.
Conclusions: Using DNA nanoparticles, bioconjugation schemes with diverse target reactivity, and scalable, standard nano-lithography techniques, we were able to create high-density, single-protein arrays from human lysate. These arrays have a wide range of potential uses, including iteratively probing with a variety of multi-affinity probes to both identify and quantify the composition of complex protein mixtures using Protein Identification by Short-epitope Mapping (PrISM).

P1.14 | Detection and Quantification of Proteins Using Protein Identification by Short-epitope Mapping (PrISM)

Parag Mallick (paragm@nautilus.bio), Nautilus Biotechnology, San Carlos, CA

Introduction: The field of proteomics is poised for a single-molecule revolution – enabling more comprehensive analysis of the proteins in a sample with increased sensitivity, reproducibility, and accessibility. Here we demonstrate the Protein Identification by Short-epitope Mapping (PrISM) methodology. PrISM is a single-molecule analysis method where intact proteins are immobilized and analyzed in parallel using multi-affinity probes to create a binding pattern used to identify each individual protein molecule.
Methods: PrISM uses non-traditional affinity reagents with high affinity and low specificity that bind to short epitopes in multiple proteins. Sample proteins are conjugated to DNA nanoparticles and deposited on a high-density patterned flow cell at optically resolvable locations. We acquired PrISM data on native biological and control samples using over 50 multi-affinity probes targeting trimer or tetramer sequences.
Results: We report single-molecule deposition of over 1 billion DNA nanoparticle complexes on a flow cell. We demonstrate protein how the PrISM methodology identifies individual protein molecules through iterative probing with our multi-affinity probes. Further, we provide an analytical assessment of the sensitivity and specificity of PrISM and demonstrate the ability to accurately estimate the false identification rate of these proteins using a target-decoy based statistical approach.
Conclusions: Combining single-molecule analysis, intact (non-digested) proteins, and iterative probing, PrISM provides a new tool for the quantitation of proteins. We demonstrate linear and reproducible quantification of proteins using PrISM, enabling application to biological samples and key biological questions.

P1.15 | Real-Time Dynamic Single-Molecule Protein Sequencing on an Integrated Semiconductor Device

Khanh (Kendrick) D. Nguyen (knguyen@quantum-si.com), Quantum-Si, San Diego, CA

Studies of the proteome would benefit greatly from methods to directly sequence and digitally quantify proteins and detect post-translational modifications with single-molecule sensitivity. Here, we demonstrate single-molecule protein sequencing using a dynamic approach in which single peptides are probed in real-time by a mixture of dye-labeled N-terminal amino acid recognizers and simultaneously cleaved by aminopeptidases. We annotate amino acids and identify the peptide sequence by measuring fluorescence intensity, lifetime, and binding kinetics on an integrated semiconductor chip. To demonstrate that this core methodology and its kinetic principles apply to a wide range of peptide sequences, we sequenced the synthetic peptides DQQIASSRLAASFAAQQYPDDD, RLAFSALGAADDD, and EFIAWLV—a segment of human Glucagon-like peptide-1 (GLP-1)—under the same sequencing conditions used for DQQRLIFAG. Each peptide generated a characteristic kinetic signature in accordance with its sequence. We obtained readouts as far as position 18 (the furthest recognizable amino acid) in the peptide DQQIASSRLAASFAAQQYPDDD, illustrating that the method is compatible with long peptides and capable of deep access to sequence information in peptides. This capability allows recognizers to identify multiple amino acids in an information-rich manner that enables discrimination of single amino acid substitutions and posttranslational modifications.

P1.16 | Whole-tissue Mapping of >5000 proteins by Micro-scaffold Assisted Spatial Proteomics (MASP)

Jun Qu[1] (junqu@buffalo.edu), Min Ma[1], Shihan Huo[1], Ming Zhang[1], Shuo Qian[1]
[1]University at Buffalo, Buffalo, NY

Quantitative mapping of proteins on ***whole-tissue levels*** provides comprehensive insights into the spatially organized regulatory processes/networks related to diseases/therapies, but remains a tremendous technical challenge. Previous works either fall short in low spatial resolution, insufficient depth, or are incapable of mapping whole tissue slices. To address this important need, here we describe a micro-scaffold assisted spatial proteomics(MASP) strategy, based on spatially resolved micro-compartmentalization of tissue using a 3D-printed micro-scaffold, capable of mapping thousands of proteins across a whole-tissue slice with excellent quantitative quality. MASP consists of three components, each was rigorously optimized to achieve accurate, quantitative protein mapping: ***first,*** robust/precise tissue micro-compartmentalization using a novel 3D-printed micro-scaffold and high-throughput procurement of location-specific micro-specimens; ***second,*** efficient/reproducible extraction, clean-up, and digestion of the micro-specimens followed by sensitive/reproducible LC-MS analyses; ***third,*** generation of protein distribution maps with a MAsP app following accurate protein quantification. The mapping accuracy was validated by corroborating the ***i)***spiked-in vs. observed patterns, ***ii)***expected vs. observed maps of brain-region markers, ***iii)***maps of the components of the same heterodimer. MASP was applied in mapping >5000 cerebral proteins in the mouse brain, encompassing numerous important brain markers, regulators, and transporters, where most of the proteins were *first-ever mapped* on the whole-tissue level. MASP also demonstrated the potential to reveal the landscapes of spatially organized signaling pathways and biological functions. For example, MASP successfully mapped >70% of the proteins involved in Alzheimer's disease pathway, and acquired maps of various blood-brain-barrier drug transporters, which provides valuable information for targeted drug delivery to specific brain regions. In summary, with MASP, for the first time, in-depth and accurate spatially-resolved proteomics mapping of whole tissue is achieved. The micro-compartmentalization approach is robust and versatile, and can be conjugated with various sample preparation and analytical platforms. Efforts directed to further improvement of resolution and analytical throughput are ongoing.

P1.17 | Reproducible and Quantitative Precipitation of Low Starting Protein Quantities in the ProTrap: Comparison to Conventional Precipitation

Jessica Nickerson[1] (jnickerson@allumiqs.com), Victoria Miller[1], Sara Lahsaee Little[1], Jean-François Noël[2], Angela Giraldo[2], Hugo Gagnon[2]
[1]Allumiqs, Halifax, NS, [2]PhenoSwitch Bioscience, Sherbrooke, QC

Reproducible, complete, and unbiased recovery of proteome samples is imperative to assure meaningful results for discovery-based and quantitative analyses. Maximized recovery is particularly important in the highly prevalent applications of low sample quantities and dilute biofluids. Precipitation of a protein sample is a simple and effective means to achieve that goal. Here, we evaluate the recovery and bottom-up quantitative precision following the preparation of minimal starting quantities within the ProTrap precipitation cartridge. Initial experiments examined the total protein recovery following precipitation of 1 to 10 µg of BSA using BCA and LC-UV assays. Assessment of residual protein in the supernatant fraction revealed approximately 90% of the initial sample was recovered in the pellet. The high-throughput precipitation approach was further evaluated in a bottom-up workflow against precipitation in a

conventional microcentrifuge tube for more complex samples including human cerebrospinal fluid and mononuclear cells (starting with 0.1-10 µg total protein). Following precipitation, pellets were re-suspended, subject to trypsin digestion, and analyzed by LC-MS/MS on a Sciex ZenoTOF 7600 and TripleTOF 6600 using both a standard-length gradient as well as a higher throughput gradient. The parallel processes were compared on the following criteria: total recovery, bottom-up identification efficiency and quantitative precision. The ProTrap-based precipitation method was found to be comparable to that of the conventional microcentrifuge tube method, while optimizing pellet recovery. The procedure described is beneficial in the analysis of very dilute samples, such as CSF and urine, as well as in the analysis of samples with limited availability.

P1.18 | LFQ HR-DIA Workflow for Plasma Proteomics on an Automated Platform

Santosh Renuse[1] (santosh.renuse@thermofisher.com), Woong Kim[1], Amirmansoor Hakimi[1], Steven Reeber[1], Kristan Bahten[1], Runsheng Zheng[2], David Horn[1], Daniel Lopez-Ferrer[1]
[1]Thermo Fisher Scientific, San Jose, CA, [2]Thermo Fisher Scientific, Gemering, Bayern, Germany

INTRODUCTION: Plasma proteome continues to be a good source of biological information to diagnose health and monitor health. Mass spectrometry-based proteomics has enabled characterization of plasma samples for biomarker discovery. One of the major challenges for handling plasma samples is processing of large cohort of clinical samples with high accuracy and precision. Here, we highlight a solution at the plasma samples preparation stage using AccelerOme, an automated sample preparation platform comprised of standardized and optimized software, liquid handler, and reagent kit in order to overcome challenges of plasma sample preparation. This study highlights metrics from 34 technical replicates of plasma from a single source processed using AccelerOme followed by Data-Independent Acquisition (DIA) mass spectrometry.
METHODS: Pooled plasma sample was purchased and was aliquoted in 34 technical replicates. AccelerOme LFQ kit was used with the AccelerOme system. Each sample was lysed using 50µl of lysis buffer and transferred to a 96 well plate. The 96 well plate was placed in AccelerOme where the samples were reduced/alkylated, digested, and desalted completely hands-free. Data acquisition was performed using an Orbitrap Exploris 480 Mass Spectrometer coupled to Vanquish™ Neo UHPLC system. Data analysis was done using Biognosys Spectronaut™ software.
RESULTS: AccelerOme system provides a hands-free solution for standardized, optimized, and verified hardware, software, and reagents to produce quality samples for plasma samples processing minimizing variability across different users. High quantitation accuracy of plasma proteome was achieved using LFQ HR-DIA workflow.

P1.19 | Leveraging Microsoft Suite of Tools to efficiently Organize and Operate an Academic Research Group

Faith M. Robison[1] (faithr@wustl.edu), Benjamin A. Garcia[1]
[1]Washington University School of Medicine in St. Louis, St. Louis, MO

Research inherently is a time-consuming endeavor and operation of a research group to facilitate research is a dramatically greater investment in resources; both time and personnel. Often, operational efficiency is placed to the wayside in order for the production of results. Here, we describe utilization of Microsoft's suite of tools, which are frequently readily available to academic institutions, in order to efficiently operate a research laboratory. The Microsoft platform includes modes of communication, collaboration, storage, and automation; all tasks critical to research group operation.

Planner can be used to organize projects, assign tasks, and provides a planning hub to track the progress of a project. Teams meetings allow for virtual trainings to be recorded and automatically stored for future training the research group. We also introduce a Powerapp developed in the Garcia lab used for asset management of all items from pipettes to large equipment. The Powerapp allows for asset tracking, issue management, and creation of a knowledgebase for tips and tricks for maintaining and running equipment.

P1.20 | Improved data-independent acquisition (DIA) and data-dependent acquisition (DDA) performance on low-level proteomic samples using a novel Zeno trap

Katherine Tran[1] (katherine.tran@sciex.com), Ihor Batruch[1], Stephen Tate[1], Nic Bloomfield[1], Anjali Chelur[1], Jose Castro-Perez[2]
[1]SCIEX, Concord, ON, [2]SCIEX, Framingham, MA

We used a ZenoTOF7600 system in-line with a Waters M-Class LC system to determine protein identifications across varying commercial K562 tryptic digest loads in either Zeno SWATH DIA or Zeno DDA modes. Using Zeno SWATH DIA at sample loads of

0.25, 0.5 and 1 ng loads, more than 900-1100, 1400-1500 and 2100-2300 protein groups were identified, respectively, and 45-55% of these identifications had a CV less than 20% when searched against a spectral library. At the peptide precursor level for the same loads, there were 2900-4100, 5000-5700 and 8700-12200 corresponding precursors for the 0.25, 0.5 and 1 ng loads. We tested higher loads and identified 4200, 5000 and 6100 protein groups for 5, 10 and 25 ng loads, respectively, and 6483% of these identifications satisfied the 20% CV cutoff. For a 50 ng load, more than 6300 protein groups were identified, of which 90% had less than 20% CV, and 56000 precursors were identified. When the data were searched against a FASTA library in library-free mode, the overall number of identifications and those at 20% CV cutoffs approach those achieved when processed using the spectral library approach. A 200 ng and 500 ng load of K562 tryptic digest was tested in Zeno DDA mode. From these experiments, we were able to identify 4600 and 5100 protein groups for the 200 and 400 ng loads, respectively, with 43000 and 56000 peptides for each load.

P1.21 | Accelerating DIA Studies with Fast Microflow LC and Zeno SWATH Acquisition

Todd Smith[1] **(todd.smith@sciex.com)**, Christie Hunter[2]
[1]*SCIEX, St. Paul, MN,* [2]*SCIEX, Redwood City, CA*

As the field of quantitative proteomics continues to evolve, larger biological cohorts are being studied, often using precious samples obtained from biobanks or other difficult-to-obtain sources. This creates 2 workflow requirements: the need to acquire quantitative data on the digested samples faster and the need to use smaller amounts of sample. For these types of studies, data-independent acquisition (DIA) continues to grow as the workflow of choice for reproducible quantitative analysis of large numbers of proteins from a proteomic sample. As such, new workflows and software tools have emerged. Previously, the combination of fast microflow chromatography and SWATH DIA enabled large numbers of proteins to be quantified from complex proteomics samples at very high rates, up to 100 samples per day. Zeno MS/MS on the ZenoTOF7600 system provides ~5-to 6-fold increase in peptide MS/MS sensitivity and can be used in MRMHR, DDA and DIA workflows.2,3,4 Also, multiple powerful algorithms have emerged that have enabled more proteins to be identified and quantified from DIA data, such as DIA-NN software. Here, the improvements in proteins identified and quantified using microflow SWATH DIA coupled with Zeno MS/MS is described. Four different gradient lengths (5, 10, 20 and 45 minutes) were tested to cover a range of application needs. The library-free approach to processing DIA data (using in silico generated spectral libraries) was also evaluated vs. the traditional shotgun proteomics approach with Zeno DDA. Other workflow comparisons were performed to benchmark the workflows. DIA data were processed with DIA-NN software and DDA data were processed with ProteinPilot app in OneOmics suite.

P1.22 | New skill sets are required with lab growth: Lessons learnt

Jennifer Van Eyk (Jennifer.VanEyk@cshs.org), Cedars-Sinai Medical Center, Los Angeles, CA

As a lab matures and grows, the lab structure, hierarchical organization and inner workings must also evolve. The role of the lab head or PI is to oversee lab finances (including the additional challenges for proteomics labs of paying for service contracts and the new mass spectrometers), personal and their career development while driving science forward. This requires skill development of the PI so they can be efficient and organized to allow for (maximal) sustainable lab growth. We hypothesized that working with a professional coach could provide missing skills that are often required for successful transition of small to midsize businesses. Results: With growth of a lab, the PI needs to ensure that the lab organization and administration hierarchical structure is defined and modified to ensure that it fits the number of scientists (staff and trainees) and scope of science and technology deployed. This required identifying the lab leadership, defining their roles, support systems and feedback mechanisms. Balancing and prioritization of finances between personal, personal growth and equipment needs can be challenging to deal with and to communicate. Different communication processes for the leadership and for the lab members were tried along with automation and process for common continuous tasks were invoked. For the PI, time management and calendar controls were assigned to single person removing from an open-door policy to structured time with flexible-meeting time embedded. Meeting agendas and goals were developed even for one-on-one meetings with next steps clearly stated. Level 10 meeting structure was evoked for larger meetings.
Conclusion: Scaleup requires developing a management structure so growth isn't bottlenecked by a leadership team needing to review every decision. Defining leaders and delegation of all major tasks although one does not have to change everything to succeed. Communicate early and often with lab members and act on feedback.

P1.23 | Toward a Routine, Robust, and Inexpensive Automated Proteomics Preparation Solution

Ahmed Warshanna[1] (awarsha1@jhmi.edu), Ben Orsburn[1]
[1]Johns Hopkins University School of Medicine, Baltimore, MD

Proteomics has unparalleled potential to study the many complexities of life but is historically limited by high cost and relatively low throughput methods, and a reputation for poor reproducibility. The aim of this project is to create a fully automated high throughput method for proteomics sample preparation for clinical and research applications. Using a 96-well S-Trap plate, our platform has a flexible sample input and allows for the high throughput form we seek to achieve in a format conducible to automation. When proteins are linearized on the S-Trap platform we hypothesized that trypsin quality might be less essential to achieve comprehensive digestion. As a proof of concept, we have completed a study of an extreme mouse aging model using 12 week old and 85 week old mice. Sixteen separate anatomical samples were removed for proteomic analysis from 4 mice of each cohort, for a total of 128 samples. Following preparation, samples were analyzed over the course of 7 days using an EvoSep One equipped TIMSTOF Flex mass spectrometer quantifying over 7,000 proteins. We aim to further automate this method using a low-cost benchtop liquid handler (Opentrons OT-2). The Opentrons will also allow for high reproducibility as the robot will supersede human error and speed. This method has the potential to allow proteomics analysis to be time and cost efficient, which will make its use tenable for clinical and research applications.

P1.24 | Multi-Affinity Probe Development for Short-Epitope Mapping

Sheri Wilcox[1] (sheriw@nautilus.bio), Ellis Anderson[1], Filip Bartnicki[1], Hunter Boyce[1], Kaitlyn Burke[1], Thomas Cunningham[1], Jacob Devine[1], Jarrett Egertson[1], Brooks Hoehn[1], Elvis Ikwa[1], Jonathan Leano[1], Cara Li[1], Jennifer McGinnis[1], Jessica Nicastro[1], Jayetha Panakkadan[1], Sudhakar Raburi[1], Tanvir Saini[1], Aimee Sanford[1], Patrick Schwientek[1], Noah Steiner[1], Sonal Tonapi[1], Jas Villanueva[1], Sophie Watts[1], Sujal Patel[1], Parag Mallick[1]
[1]Nautilus Biotechnology, San Carlos, CA

Protein Identification by Short-epitope Mapping (PrISM) depends on the ability to query proteins iteratively with a diverse set of multi-affinity probes that bind short amino acid sequences. Each individual multi-affinity probe is developed with high affinity and low specificity to bind short-epitopes in multiple proteins to overcome the challenges of single-reagent sensitivity and specificity. The platform is agnostic to the scaffold for each probe and therefore can employ diverse types of reagents. Selection targets typically consist of specific tripeptides, but we have demonstrated that the probes consistently recognize biosimilar patterns related to the selection targets. Not only do these reagents bind peptides that contain the epitope patterns, but they have also been demonstrated to recognize intact, human proteins containing the tripeptide sequences as well. The ability to develop these diverse reagents and use them to bind intact proteins provides a path forward for accessibility to essentially the entire human protein without obtaining probes against each of the 20,000 human proteins themselves. In fact, the ability of these reagents to iteratively recognize short-epitopes enables decoding of 95% of the human proteome with as few as 300 such probes. Furthermore, translation across species is similarly straightforward. There is no dependence of the species cross-reactivity of a multi-affinity probe when detecting short epitopes and decoding the counted proteins using PrISM. We demonstrate that diverse multi-affinity probes can be developed to bind short-epitopes. Characterization of the binding profiles and performance of these reagents with PrISM enables decoding and counting of far more proteins than previously possible.

P1.25 | Kitted universal MAM: automatable sample processing for all stages of biological medicines

John Wilson[1] (john@protifi.com), Alexandre Zougman[2]
[1]ProtiFi LLC, Fairport, NY, [2]University of Leeds, Leeds, West Yorkshire, United Kingdom

Introduction: Proteinaceous medicines continue to emerge as treatments for diverse ailments; indeed 27 of the 50 top-grossing therapeutics are biological drugs like antibodies. These medicines must be analyzed and QCed; Multiple Attribute Monitoring (MAM) is an LCMS technique designed to simultaneously and directly monitor critical quality attributes (CQAs) of biologics like impurities, PTMs and sequence. MAM's popularity continues to rise and presents challenges due to the highly varied nature of biotherapeutic manufacturing from bioreactors containing surfactants to final formulations. 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 biomedicine production.
Methods: The MAM kit was developed to meet the specific needs and requirements of MAM. The kit was stress tested using the NISTmAb RM 8671 monoclonal antibody spiked into various "sample preparation challenges" like PEG or surfactants. Samples were prepared by hand, in a semi-automated format on the Tecan A200 positive pressure workstation or fully automated on the

Tecan EVO fluid handler equipped with an A200. Samples were analyzed on an Agilent 6546, Q Exactive BioPharma and/or VelosPro Orbitrap/IT ETD.
Preliminary Data: The MAM-sample processing kit was effective without modification in protocol at removing all contaminant challenges including sweeteners, salts, surfactants, antioxidants, emulsifiers, soluble binders and bulking agents, lubricants, coatings, dyes and other small molecules. Excellent reproducibility was observed for both unmodified and PTM-modified peptides. Residual host cell proteins (HCPs) were detected, in line with previous results. Variation from sample preparation was similar to levels of variability observed in technical replicate injections. We anticipate the use of the same kit without change or optimization at all stages of biopharmaceutical manufacturing has potential to facilitate even broader uptake of MAM for process control, formulations, and QC and release testing.

P1.26 | Shredder: a new way to sequence

John Wilson[1] (john@protifi.com), Darryl J. Pappin[2], Alexandre Zougman[3]
[1]ProtiFi LLC, Fairport, NY, [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, [3]University of Leeds, Leeds, West Yorkshire, United Kingdom

Protein sequencing remains at the heart of proteomics, especially in cases of new species, variants or PTMs, and has traditionally been performed via fragmentation by enzymes, chemicals or, length allowing, by various techniques of fragmentation.

Each of these approaches has limitations: for full sequence coverage, often multiple enzymes and/or chemical fragmentations must be combined due to their specificity, lack thereof, or inappropriateness for a particular protein sequence of interest (e.g. submitting the basic tails of histones to tryptic digestion). Gas-phase fragmentation is limited not only by the size and length of peptide introduced, but also by the fragmentation behavior and specifically the lack of fragmentation, often in exactly the region of interest. To solve these problems, we developed the Shredder, a new approach to bottom-up proteomics sample preparation that randomly cleaves all along the peptide backbone using a combination of low-specificity active sites and activated residues in rapid reaction times.
We demonstrate the Shredder providing full sequence coverage, show its efficiency compared to multiple enzymatic digestions and suggest possible uses in both academic and industry settings.

P1.28 | A novel intelligent data acquisition Hybrid-DIA mass spectrometry strategy: enabling data-driven and hypothesis-driven approaches in one go

Yue Xuan[1] (yue.xuan@thermofisher.com), Min Huang[2], Xiangyun Yang[2], Kyle Fort[3], Jesse. D. Canterbury[4], Ana Martínez-Val[5], Jesper Olsen[5], Yu-Ju Chen[6], Yasushi Ishihama[7], Sandra Goetze[8], Bernd Wollscheid[8], Alexander Makarov[3], Thomas Moehring[1]
[1]Thermo Fisher Scientific, Bremen, Bremen, Germany, [2]Thermo Fisher Scientific (China) Co. Ltd, Shanghai, Shanghai, China, [3]Thermo Fisher Scientific (Bremen) GmbH, Bremen, Bremen, Germany, [4]Thermo Fisher Scientific, San Jose, CA, [5]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Copenhagen, Denmark, [6]Academia Sinica, Taipei, Taipei, Taiwan, [7]Kyoto University, Kyoto, Kyoto, Japan, [8]Institute of Translational Medicine (ITM), Dep. of Health Sciences and Technology (D-HEST), ETH Zurich, Zurich, Zurich, Switzerland

Translational scientists face the dilemma to choose between comprehensive profiling and sensitive targeted quantitation, especially with large sample cohorts. Proteomic profiling is commonly used to discover biomarkers, having a great potential for prognostic and predictive biomarkers; however, it still misses the sensitivity to quantify all the markers of interests. Therefore, targeted quantitation experiments of the potential markers are analyzed in the validation phase. This leads to high cost, time losses and more sample consumption. To address these challenges, we develop a novel intelligent data acquisition “Hybrid-DIA” MS strategy that enables comprehensive proteome profiling via high resolution MS1-based data-independent-acquisition MS and on-the-fly intelligently switching the acquisition mode to parallel reaction monitoring (PRM) for sensitive quantification of the markers, substantially increasing throughput and reducing sample consumption. The global profiling and quantitation performance of Hybrid-DIA MS have been investigated and benchmarked against the standard DIA MS methods by analyzing HELA cell lysate digest. Comparable number of proteins/peptides are identified and quantified with both methods, while Hybrid-DIA MS significantly improves the LOQ/LOD of the low abundant biomarkers down to attomole range. We then applied Hybrid-DIA to analysis the oncoprotein mutations within the lung cancer cell lines, known clinical markers and drug targets from the melanoma patient tissue samples, and major signaling pathways in the cancer cell phosphoproteomes, respectively. With high precision and reproducibility, Hybrid-DIA MS can quantify the EGFR and KRAS cancer mutations in lung cancer cell lines, capture clinically actionable markers in the melanoma patient tissue sample, as well as accurately quantify 100+ phosphorylation sites covering seven major signaling pathways in cancer cells (EGFR, RAS-MAPK, PI3K-AKT-mTOR, AMPK, apoptosis and stress response), while digitalizing their underlying proteotypes.

This novel Hybrid-DIA MS methodology presents a new capability to combine the data-driven and hypothesis-driven approaches in one go, accelerating the entire biomarker discovery and validation pipeline.

P1.30 | Ultrasensitive top-down proteomics analysis using spray-capillary-based capillary electrophoresis mass spectrometry

Si Wu (si.wu@ou.edu), University of Oklahoma, Norman

Traditional proteomics approaches typically started with bulk cells for protein expression profiling in biological samples. However, these approaches fail to account for cell-to-cell heterogeneity within the sample. Capillary electrophoresis (CE) is a promising platform for studying limited samples, such as single cells, due to its ultra-low sample consumption, high separation efficiency, and ultra-low detection limit. However, offline micro-sampling processes starting with large starting sample volumes (e.g., ul-level) are often involved in current single-cell CE-MS methods, which can decrease the sampling precision and accuracy. We have developed and implemented a spray-capillary CE-MS platform that performs online micro-sampling (*nL*-level), CE separation, and MS analysis for ultrasensitive top-down proteomics analysis such as a low number of *HeLa* cells (1-100). Our results demonstrated that the spray-capillary is capable of precise microsampling and high-throughput quantitative CE-MS proteomics analysis.

P2: Aging and Neurodegenerative Diseases

P2.01 | Investigating Regionally Regulated Control of Protein Synthesis and Degradation Within the Brain

Rebecca Burlett[1] (burlett.rebecca@gmail.com), Leena M. Patil[1], Yiran Liang[1], Kenneth Virgin[1], Ryan Kelly[1], John C. Price[1]
[1]Brigham Young University, Provo, UT

Regulation of metabolism is essential to health and lies at the center of many diseases within the body. Poor regulation of protein homeostasis within the brain leads to dysfunction and Alzheimer's Disease (AD). The risk of developing AD is strongly impacted by the Apolipoprotein E (ApoE) isoform that is produced by the body. However, it is poorly understood how ApoE (a lipid transport protein) affects brain proteostasis. We investigated control of protein synthesis and degradation using deuterated water metabolic labeling in AD model mice. Utilizing a cryostat, we collected 20 micron slices of brain tissue that were metabolically labeled with deuterium to allow analysis of the Hippocampus, Entorhinal, and Visual regions. Slices were then treated with a Nissl Stain and regions identified using Allen Brain Atlas. Next, Laser Capture Microdissection (LCM) was used to collect 200x200x20 micron samples within targeted areas. Nanofluidic sample preparation (AutoPOTS) methods were used to prepare samples for mass spectrometry. Trypsinized samples were analyzed using LC-MS to measure protein abundance and turnover rates. With our targeted microsampling technique we have reproducibly identified ~2500 proteins within our sampled regions. We analyze the protein abundances and turnover rates using analysis software, i.e. DeuteRater and Metaboanalyst. The outputs generated were then put through DAVID and produced ontologies for the proteins identified. We have been able to determine specific regulation trends within the proteins and pathways for each region. These trends allow us to determine whether pathways or regions are regulated by synthesis or degradation. With this research we are working towards a better understating of how the regions of the brain differ and are affected by ApoE's isoforms and AD progression overall.

P2.02 | Proteomics and Lipidomics Reveal Alterations Associated with Plasma Membrane Structure and Signaling in Niemann-Pick Type C

Stephanie M. Cologna[1] (cologna@uic.edu), Koralege Praneeth C. Pathmasiri[1], Thu T. Nguyen[1], Melissa R. Pergande[2], Daisy Edminson[1], Swetha Gowrishankar[1]
[1]University of Illinois Chicago, Chicago, IL, [2]University of Wisconsin-Madison, Madison, WI

Niemann-Pick Type C is a rare, fatal, progressive neurodegenerative disorder that arises from mutations in either the NPC1 or NPC2 cholesterol transporters. As a result of the genetic defect, endolysosomal cholesterol and sphingolipids accumulate as well as a number of downstream events take place such as oxidative stress, neuroinflammation, and calcium imbalance among others. To date there is no FDA-approved therapy and individuals typically succumb to complications of the disorder in the early adulthood years. Our laboratory utilizes mass spectrometry coupled with biochemical and molecular tools to gain a deeper understanding of NPC and to help pave the way for future therapeutic development. Using brain tissue from the null *Npc1* mouse model and mass spectrometry tools, we have identified candidate protein biomarkers, protein phosphorylation alterations, changes in lipid levels as well identify endogenous protein-protein interactions with NPC1. Example of our findings include alterations of proteins associated with synapse function including signaling, long-term potentiation and long-term depression. Evaluation of changes in phosphorylated proteins has revealed increased phosphorylation status of NMDA receptor subunit N2B as well as CaMKII – a critical kinase that regulates several membrane events. Lipidomics analysis revealed that the membrane phospholipid, PIP_2 is decreased in NPC brain tissue along with the precursors PI and PIP. Finally, we have identified that proteins that interact with the NPC1 protein differ by tissue type, and many play a role in the organization and fidelity of the cytoskeleton. Together, our data suggest that plasma membrane structure and function may be altered in NPC. Our current efforts are focused on plasma membrane specific changes and how those relate to the pathophysiology of the disease.

P2.03 | Chronic Levetiracetam Treatment Alters Synaptic Protein Turnover in Alzheimer's Mouse Model and Reduces Production of Aβ42

Olivia DeGulis[1] (oliviadegulis2023@u.northwestern.edu), Nalini Rao[1], Jeff Savas[1], Anthea Weng[1]
[1]Northwestern University, Chicago, IL

Alzheimer's disease (AD) is defined as dementia with the presence of intracellular tau tangles and extracellular amyloid beta plaques. Toxic $A\beta_{42}$ species, that are the foundation for plaques, are produced because of sequential proteolytic cleavage of the amyloid precursor protein (APP). When beta-secretase cleaves APP, beta-CTFs are released which, if then cleaved by gamma

secretase, release $A\beta_{42}$ peptides. These proteins accumulate throughout the brain and lead to cognitive decline. Levetiracetam (Lev) is an atypical antiepileptic drug. It has been generically prescribed for a decade for epilepsy treatment. It is understood that Lev works to stabilize neuronal networks in epileptic patients; therefore, there is a possibility the drug can help patients with dementia. While the explicit mechanism of action of Lev has not yet been discovered, it is known that Lev bins to SV2a, a synaptic vesicle protein. In AD, we previously discovered that Lev selectively normalized synaptic vesicle endocytosis machinery abundance and restored non-amyloidogenic processing of APP. We set out to investigate if chronic Lev treatment induces Aβ clearance or impedes the production of Aβ. To do this, we created a metabolic labelling paradigm which allowed for differentiation between N15 incorporated proteins while receiving Lev treatment and the older N14 labelled proteins. This for a high-level quantitative mass spectrometry-based analysis to determine how Lev affects $A\beta_{42}$ deposition. We found that chronic Lev treatment lowers beta-CTF levels while the levels of full-length APP remain unchanged, illuminating the possibility that Lev alters the cleavage process of APP. We also isolated amyloid fibrils to investigate if Lev alters production or clearance of Aβ. Using a targeted mass-spectrometry methodology, we determined a trend of lowered newly synthesized Aβ in the Lev-treated animals, suggesting Lev lowers Aβ production. Overall, Lev provides a potential mechanism for therapeutic target that could minimize Aβ AD pathology.

P2.04 | Profiling of Senescence-Associated Surfaceomes with Optimized Data-Independent Acquisition Methods

Amit K. Dey[1] (amit.dey2@nih.gov), Reema Banarjee[2], Dimitrios Tsitsipatis[2], Anjana Ram[2], Judith Campisi[3], Birgit Schilling[3], Allison Herman[1], Myriam K. Gorospe[1], Nathan Basisty[2]
[1]National Institute on Aging-NIH, Baltimore, MD, [2]National Institute on Aging, Baltimore, MD, [3]Buck Institute for Research on Aging, Novato, CA

Senescent cell accumulation contributes to aging and many pathological processes, including chronic inflammation, cancer, and neurodegeneration. Pre-clinical mouse studies have demonstrated that targeted removal of senescent cells in mice reverses these conditions and promotes health and longevity, suggesting that selective elimination of senescent cells is a promising therapeutic approach for mitigating several aging pathologies. However, in humans, removing senescent cells will require therapeutic targets. The *surfaceome* —the cell surface protein repertoire— is a promising source of therapeutic targets and biomarkers of senescence due to its extracellular accessibility, which enables the identification, quantification, and targeting of cells from tissues. Herein, we performed comprehensive and quantitative proteomic profiling of senescent cells with an optimized data-independent acquisition (DIA)-MS approach to identify *surfaceome* biomarkers of senescence burden in human tissues. Using an N-linked glycoprotein cell surface capture protocol, cell surface proteomes were oxidized, covalently biotinylated, and enriched from four human cell types: monocytes, fibroblast, vascular smooth muscle, and vascular endothelial cells. We optimized the DIA-MS workflows on a Q-Exactive HF mass spectrometer, comparing instrument parameters and precursor isolation window designs, ultimately using a 25-variable width window method with injection times and resolution optimal for low protein inputs. We identified 1444, 1335, 1753, and 1064 proteins in each respective cell type. Filtering for true surface glycoproteins, indicated by deamidation and N-X-S/T motif (Asn-X-Ser/Thr), resulted in 341, 420, 582, and 346 proteins in each cell type. We identified previously reported senescent surfaceome proteins, such as DPP4, and novel proteins, including CDCP1, VCAM1, Calreticulin, and others. Several top protein candidates were further validated with flow cytometry analysis.

A comprehensive analysis of the surfaceome landscape reveals novel cell surface targets for further validation. These results may provide clinically useful information for establishing therapeutic targets or biomarkers to translate senescence-targeted therapies to treat age-related declines.

P2.06 | Imaging Beyond the Slide: Novel Whole-Brain Serial Two-Photon Imaging and Spatial Proteomics in Pre-Clinical Alzheimer's Disease Animal Models

Michael L. Easterling[1] (mike@easterling.com), Mark Lim[2], Stefan Linehan[3], Gianna Ferron[4], Joshua Fischer[5]
[1]Bruker Daltonics, Billerica, MA, [2]AmberGen, Inc., Billerica, MA, [3]TissueVision, Inc., Newton, MA, [4]TissueVision Inc., Newton, MA, [5]Bruker Scientific, LLC, Billerica

Introduction: Alzheimer’s Disease (AD) has a strong spatial-temporal component to its progression, where different brain regions are affected by amyloid-beta (Aβ) plaque deposition at varying time points. Standard imaging and analysis platforms can neglect these details, as they lack the ability to pair high-yield whole-brain imaging with region-specific quantitation. Furthermore, many Aβ analyses require homogenization of tissue, prohibiting secondary analysis. To address this gap, we have developed a novel high-throughput whole-brain imaging pipeline for pre-clinical AD models to quantitatively track plaque progression as a function of brain region across time while producing indexed tissue sections for MALDI HiPLEX-IHC staining and imaging mass spectrometry (IMS) analysis.
Methods: Aβ plaques in the well characterized 5XFAD mouse model of AD were labeled with Methoxy-X04, an Aβ plaque-specific compound, prior to trans-cardial perfusion and whole-brain excision. The brains were processed on the TissueCyte Serial Two-Photon Plus (STP^2) imaging platform to produce fully aligned multi-channel volumetric datasets, yielding high resolution 3D models

of each brain. Registration to the Allen Mouse Brain Common Coordinate Framework (CCFv3) and region-specific plaque analysis were conducted to determine plaque size, density, and total number per animal brain. Resulting sections were evaluated with MALDI HiPLEX-IHC for Aβ42, pTau, MAP2, GAD67, GLUT1, NeuN, synapsin, myelin and neurogranin tags.
Results: The analysis of whole-brain plaque distribution in the 5XFAD mouse model revealed distinct spatial-temporal changes across the brains. STP[2] imaging, combined with CCFv3 mapping and MALDI-IHC analysis allowed for targeted evaluation of Aβ plaques.
Conclusions: This novel technology has great promise for quantifying the spatial-temporal Aβ plaque efficacy of AD animal models, and the production of translatable pre-clinical AD drug discovery data. The high sensitivity and precision of the STP[2] platform combined with the high multiplexity of MALDI HiPLEX-IHC and IMS imaging can benefit region-specific disease progression compared to standard laboratory approaches.

P2.07 | Deep Proteome Turnover in Human iPSC-derived Neurons

Ashley Frankenfield[1] (afrankenfield@gwu.edu), Jiawei Ni[1], Jamison Shih[1], Ling Hao[1]
[1]*The George Washington University, Washington, DC*

Human induced pluripotent stem cell (hiPSC)-derived neurons provide a valuable and relevant cellular model for studying live human neurons and neurodegenerative disorders. As post-mitotic cells, neurons are unable to replenish and must strictly regulate protein turnover to maintain cellular homeostasis. In this study, we coupled dynamic stable isotope labeling with amino acids in cell culture (dSILAC) of hiPSC-derived neurons with mass spectrometry-based proteomics to measure the global protein half-lives in huma neurons. Dynamic SILAC was performed by culturing the iPSC-derived neurons in a normal “light” amino acid-containing medium. After neuron maturation, neuron medium was switched to a “heavy” lysine and arginine- containing medium to allow new proteins to be synthesized in heavy lysine and arginine. Neurons are harvested at multiple time points after medium switch followed by trypsin/Lys-C digestion and desalting. To improve the proteome depth of protein half-live measurements, we conducted extensive high pH reverse phase fractionation of tryptic peptides and conducted LC-MS/MS analysis using an optimized data-independent acquisition (DIA-MS) method. Using the peptide heavy-to-light ratio, we were able to calculate the peptide and protein half-life with both multiple and single timepoint equations for over 7000 proteins and over 58000 peptides after removing contaminants and irreproducible measures. The average half-lives of human neuron proteins are between 4-5 days with extremely slow turnover of histone and nuclear proteins and fastest turnover of proteins related to neurosecretion and ubiquitin. Ongoing work focuses on performing a comprehensive comparison of protein turnover in different neuron subtypes and evaluate how cellular stresses influence global protein turnover in human neurons.

P2.08 | Chemical Imaging of Amyloid Peptide Aggregation and Plaque Formation Dynamics in Alzheimers Disease

Jorg Hanrieder (jh@gu.se), University of Gothenburg, Molndal, Vastra Gotalands lan, Sweden

As the Worlds population is getting older the prevalence of age associated diseases such as Alzheimer’s disease (AD) is steadily increasing accordingly, currently affecting 12% over the age of 65. As the underlying mechanism remain unclear there is still no curative treatment. It is therefore critical to understand how key pathological factors of AD including beta-amyloid (A?) plaque formation are interconnected and implicated in nerve cell death, clinical symptoms and disease progression. The main goal of our research is to elucidate the biochemical processes underlying early A? plaque formation in brain tissue.

We use advanced chemical imaging modalities including hyperspectral confocal microscopy, electron imaging and mass spectrometry imaging to delineate in vivo A? build up and deposition at cellular length scales. Together with stable isotope labelling, these tools allow us to visualize A? aggregation dynamics within single plaques across different brain regions and to follow the fate of aggregating A? species from before and throughout the earliest events of precipitating plaque pathology.
We further integrate these experiments with functional amyloid microscopy using structure sensitive fluorescent probes that allow delineate the chain of events underlying amyloid aggregation, deposition and maturation.
These data, provide a detailed picture of the earliest events of precipitating amyloid pathology at scales not previously possible. The results from these studies bring considerable novel information about the deposition mechanism of A? and its toxic interactions with the surrounding. This will open up for developing tailored strategies to affect AD pathology prior to any neurodegenerative mechanisms as well as to develop new biomarkers for AD.

P2.09 | Profiling substrate specificity to assay dysregulated CDK5 activity

Jason L. Heier[1] (jheier@umn.edu), Akaash Kannan[1], Laurie Parker[1]
[1]*University of Minnesota, Minneapolis, MN*

Growing evidence shows cyclin-dependent kinase 5 (CDK5) as a key mediator of Alzheimer's Disease (AD) and an attractive therapeutic target. CDK5 is a serine/threonine kinase which plays a crucial role in brain development and synaptic activity. However, upon exposure to neurotoxic signals, CDK5 becomes dysregulated, contributing to all three pathological hallmarks of AD: β-amyloid plaque deposition, formation of neurofibrillary tangles, and neuron death. Unlike other cyclin-dependent kinases, CDK5 activity is dependent on regulatory protein p35, which is anchored to the neuron membrane. In response to oxidative stress, a cytoplasmic influx of Ca^{2+} activates proteolytic calpains which cleave p35 into fully mobile p25. Whether aberrant CDK5 activity is the result of its delocalized activation by p25 or involves an altered substrate specificity when complexed by p25 rather than p35 is unknown. Although CDK5 is an attractive therapeutic target, there are no tools to measure its activity in cell-based assays. The development of a specific peptide substrate as a reporter for CDK5 is essential for quantification of CDK5 activity in a physiologically relevant environment and critical for the development of inhibitors targeting aberrant CDK5 activity. To determine substrate specificity of CDK5/p25 and CDK5/p35, in comparison to homologous CDK2/cyclin A, we are using an approach called SERIOHL-KILR (serine-oriented human library-kinase library reactions) developed in the Rinehart Lab. SERIOHL-KILR employs a genetically encoded peptide library based on >100,000 human phosphorylation sites transformed and expressed in *E. coli*. The peptide library is reacted *in vitro* with a kinase/regulatory protein of interest and the resulting phosphorylated peptides are phosphoenriched and identified via MS/MS. We then enter the sequences of phosphorylated peptides into our *in silico* pipeline KINATEST-ID to design CDK5 kinase-specific substrates. Selected sequences are synthesized and validated for specificity and will be applied in future assay and inhibitor development.

P2.10 | Thermal Proximity Coaggregation Reveals Dysregulated Protein Complex Dynamics in the Liver of Huntington's Disease Mouse Model

Josiah Hutton[1] (josiahh@princeton.edu), Tavis J. Reed[1], Todd Greco[1], Jeffrey Cantle[2], Jeffrey B. Carroll[2], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ,* [2]*Western Washington University, Bellingham, WA*

Huntington's disease (HD) is a neurodegenerative disease that results from the expansion of a poly-glutamine (polyQ) region at the N-terminus of the huntingtin (HTT) protein. In addition to brain-linked pathologies, HD patients suffer from serious liver disease, pointing to tissue-specific functions of HTT. Though its exact function remains unknown, HTT is thought to have both cytoplasmic and nuclear functions. Although protein-protein interaction (PPI) studies have identified HTT PPIs, it remains unknown how mutant HTT impacts globally PPI networks and intracellular communication and what drives HD progression. Here, we describe the first implementation of Thermal Proximity Coaggregation (TPCA) and quantitative proteomics for defining systems-level dysregulation of protein complexes in the liver HD mouse model and the dependence of these alterations on HTT. Hepatocyte perfusions were performed from mouse livers in which HTT is expressed as wild-type (WT), a 111 polyQ expansion (Q111), or knocked out (LKO). Early and later disease stages were assessed by testing 2- and 6-months of age. From this dataset, we monitored nearly 900 CORUM protein complexes and predicted *de novo* formed PPIs. Further computational analyses pointed to specific protein networks differentially regulated by age for the three genotypes. One functional cluster was linked to histone modification and gene regulation, which included the Mi2/NuRD complex. These complexes displayed increased associations in the LKO samples from 2-months to 6-months, while decreasing in the Q111 samples. We further validate and support the disease relevance of selected altered polyQ length- and HTT-dependent PPIs using behavioral assays in an HD fly model. Altogether, our study provides a broad view of the impact of wild-type or mutant HTT on the liver tissue in the context of HD progression, as well as a valuable resource for understanding altered intracellular communication and cellular pathways linked to HD pathologies.

P2.11 | APOE and Gender-Dependent Proteome Regulation

Benjamin Jones[1] (benjaminjones1212@gmail.com), Nate Zuniga[1], Noah Earls[1], Ethan Smith Smith[1], Noah Moran[1], Katie Brown[1], Kimberly Wagstaff[1], John C. Price[1]
[1]*Brigham Young University, Provo, UT*

The gene encoding apolipoprotein E (APOE) is highly correlated with the risk of developing Alzheimer's Disease (AD). APOE is a lipid transport protein involved in transporting lipids to and from the brain. Mutations in the baseline variant APOE3 have resulted in increased (APOE4) and decreased (APOE2) susceptibility to developing the disease. While this correlation is well-established, the mechanism behind this differential susceptibility remains unknown. These mutations also modify the protein's efficiency in delivering lipids. We seek to understand how APOE impacts the brain proteome in APOE model mice. We are further testing how risk due to APOE isoform interacts with risk due to biological sex. To elucidate this mechanism, we homogenized brain samples from APOE

transgenic mice with each of the relevant genotypes and each gender. We purified and digested the proteins preparatory to analysis through liquid chromatography mass spectrometry (LC-MS). We then processed this data to identify other concentration changes in the proteome because of APOE genotype.

Our preliminary results in male mice suggest APOE-dependent differential regulation of several pathways including oxidoreductase activity that may explain how APOE increases the risk for disease onset. Our female mouse results serve as an independent verification; some changes were replicated while we also saw unique patterns that may be due to the interaction between biological sex and APOE allele. We will discuss the data obtained from the female mice and its relation to our observations of the proteome changes in the male mice.

P2.12 | Emergence of heterogeneity in Amyloid beta aggregatopathy-relevant proteoform landscape in human AD and 5xFAD murine AD model

Soumya Kandi[1] **(soumyakandi2023@u.northwestern.edu)**, Steven Patrie[1], Neil L. Kelleher[1]
[1]*Northwestern University, Evanston, IL*

Alzheimer's Disease (AD) is a progressive neurodegenerative disease which presents extensive amyloidopathy within human brain. The emergence of aggregatopathy-relevant amyloid beta (Aβ) proteoform landscape over course of disease progression is less understood. Exclusive to human population, no animal model exists for AD, requiring generation of models that emulate certain aspect(s) of the pathology. The 5xFAD transgenic mouse model of AD is one such model that develops age-dependent Aβ-aggregatopathy.

Using label-free quantitation by Top-Down mass spectrometry, we have evaluated Aβ proteoform landscape in soluble and insoluble brain material in separate cross-sectional studies, comparing its evolution with aging in 5xFAD transgenic mouse model of AD to human AD manifestation. Unsupervised proteoform groupings from Aβ proteoform quantitation in soluble versus insoluble brain material has revealed shared physicochemical properties of proteoforms correlated with AD phenotype manifestation. Similarities observed in emergence of Aβ proteoform landscape include commonalities in physicochemical properties of grouped proteoforms (eg. truncation positions), accumulation of Aβ1-42 and pyroglutamate-Aβ (pE3-42), a therapeutics-relevant proteoform as well as proteoform identifications in 5xFAD compared to human AD brain, with 79% overlap. The mouse model is specifically designed to overexpress Aβ and primarily generates Aβ1-42, a contrast to AD brain proteoform landscape which is also occupied by high abundance proteoforms like Aβ4-42 and Aβ1-34.
Unlike human pathology where vast number of variables are at play, the mouse model provides a well-controlled system for evaluating flux in Aβ proteoform landscape with aggregatopathy progression. We have established a systematic approach to potentially identifying novel mechanisms underlying aggregatopathy progression on a molecular level without prior inferences. Such an approach can provide a means to understanding key Aβ proteoforms involved in driving Aβ aggregatopathy in the human brain to improve predictive and mechanistic models of AD, and equally importantly serve as a powerful tool to validating Aβ-based models of AD.

P2.13 | Using DIA and DDA to Quantify Proteome Changes in an AD Mouse Model

Noah Moran (noah.g.moran@gmail.com), Brigham Young University, Provo, UT

Loss of proteome homeostasis leads to disease, aging, and death. One prevalent disease in humans, Alzheimer's Disease (AD), affects over six million older Americans as of 2022. The major genetic risk factor for developing AD is the ApoE gene. The ApoE3 allele is risk neutral, ApoE4 increases risk, and ApoE2 decreases risk.

In our model of proteostasis, the risk associated with ApoE can be explained by changes to the proteome related to ApoE.
Using a transgenic mouse model and LC/MS, we study how ApoE4 and ApoE2 affect proteostasis in mice relative to ApoE3, and these we aim to identify potential mechanisms that explain how ApoE alleles modulate AD risk.
In addition to traditional proteomics workflow, we developed data-independent acquisition methods (DIA) to survey ApoE dependent changes to the proteome. Our DIA strategy incorporates aspects of data-dependent acquisition (DDA) and bioinformatics software to increase the robustness of our quantitative capabilities.
Similar to our DDA data, our DIA results identify ApoE allele specific changes to pathways that may explain how disease increases.

P2.14 | Utilizing isotope labeling and LC-MS to quantify discrepancies in lipid metabolism across APOE genotypes

Coleman O. Nielsen[1] (cniels21@byu.edu), Russell Denton[1], John C. Price[2]
[1]Brigham Young University- Chemistry and Biochemistry Department, Provo, UT, [2]Brigham Young University, Provo, UT

A person's apolipoprotein E (APOE) polymorphism is one of the most predictive genetic risk factors of Alzheimer's disease (AD). Three APOE polymorphisms are predominant. Of which, APOE2 decreases the likelihood of developing AD and APOE4 increases the likelihood, relative to APOE3. The modified lipid transport behavior of these ApoE variants further emphasizes that lipid metabolism is a shared variable in multiple facets of AD including long-term inflammation, reduced cellular repair, and modified energy availability. APOE facilitates lipid transport between organs through the blood and interstitial fluid and is prevalent in lipid motility in the brain. Given the nature of APOE, it is probable that each isoform biases lipid metabolism in the brain differently. This may alter the brain's physiological state to promote or resist AD. Here, using a transgenic mouse model and LC/MS, we utilize a shotgun lipidomic method to analyze lipid metabolism in the brain for each of the APOE polymorphisms. Utilizing deuterium label incorporation, we calculated percent *de novo* biosynthesis and turnover rates for a large number of lipids across the three APOE genotypes in mice. Of the three genotypes, APOE4 mice generally displayed the highest lipid turnover and lower lipid *de novo* synthesis. Mice with the APOE2 and APOE3 genotypes were not significantly different in either attribute. We will discuss further trends in the data.

P2.15 | Evaluation of Biomarker Candidates from Senescence-Associated Monocyte Proteomes

Thedoe Nyunt[1] (thedoe.nyunt@nih.gov), Reema Banarjee[1], Dimitrios Tsitsipatis[1], Anjana Ram[1], Ceereena Ubaida-Mohien[1], Bradley Olinger[1], Gulzar Daya[1], Zhongsheng Peng[1], Martina Rossi[1], Keenan Walker[1], Myriam Gorospe[1], Nathan Basisty[1]
[1]National Institute on Aging, Baltimore, MD

Cellular senescence is a complex stress response characterized by permanent cell-cycle arrest and 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 neurodegeneration. Selective elimination of senescent cells is a promising therapeutic for treatment of age-related pathologies in humans and requires development of biomarkers to assess senescent cell burden. Here we apply quantitative proteomics to identify biomarker candidates from the senescence-associated monocyte proteome and evaluate them functionally and clinically.

THP-1 senescence was induced with 7.5 Gy of ionizing radiation. After 7 days, proliferating and senescent cell pellets and culture media were collected for LC-MS/MS analysis and assessment of a panel of canonical senescence and viability markers. Analysis of samples was performed on the Q-Exactive HF Orbitrap mass spectrometer using Data-Independent Acquisition (DIA) and analyzed in Spectronaut. Stringent filtering and GO Enrichment Analysis identified promising biomarkers and recurring biological trends. Assessment of cytokines released in response to polyinosinic:polycytidylic acid (poly-I:C) challenges were performed for functional validation of interferon viral responses. Proteomic data were comparatively analyzed against clinical data to validate translatability of *in vitro* findings.
We report the first comprehensive, unbiased assessment of the senescent monocyte proteome, comprising 1715 significant protein changes, with both shared and unique features compared to other cell types. We highlight a downregulation of ribosomal, mitochondrial, and spliceosomal proteins, and increased interferon response. Cytokine panels reveal differential viral response in senescent monocytes in the presence and absence of poly-I:C stimulation. Comparison of senescent monocyte proteomes with existing biomarkers of aging and health outcomes in human blood reveals promising peripheral senescence biomarkers. 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.

P2.16 | Brain Myelin Proteomics in the Neurodegenerative Disorder Niemann-Pick type C1 Mouse Model Reveals Defects in ATP

Koralege Praneeth C. Pathmasiri[1] (kpathm2@uic.edu), Stephanie M. Cologna[1]
[1]University of Illinois Chicago, Chicago, IL

Niemann-Pick disease, type C1 (NPC1) is a fatal, lysosomal storage disorder, which results in progressive neurodegeneration. Complications are most severe in the central nervous system. Dysmyelination in NPC1 is an understudied phenotype and has

mostly been described using standard immunostaining markers. While proteins such as myelin basic protein and myelin-associated glycoprotein have been shown to be altered in NPC1, we sought to understand the myelin proteome defects in a more comprehensive manner. For example, proteins in myelin and myelin forming oligodendrocytes play a crucial role in metabolite transport, ion transport and energy supply to the high energy demanding axon. In addition, number of myelin proteins regulate the oligodendrocyte differentiation and myelin formation which has shown to be defective in NPC1 disease.

Myelin was isolated using sucrose gradient centrifugation from *Npc1*-null and control mice. Myelin lysates were digested using S-trap. Peptide separation was done using nano flow on a C18 column and detection via a Thermo Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer. Protein identification was obtained using Proteome Discoverer equipped with the SEQUEST search engine and false-discovery rate set to 1%. Protein identifications required minimum two unique peptides and differential levels was determined using the label free method. Proteomic data revealed decreased expression of electron transport chain proteins including ATP synthase. Quantification of ATP in the isolated myelin confirmed the reduced expression of these proteins results in depleted ATP levels in the mutant myelin. Furthermore, the effect of pharmacological inhibition of NPC1 function was studied using a human oligodendrocyte cell line. The cell culture model data agrees with the observations in mice brain myelin revealing low ATP production upon inhibition of NPC1 protein in human oligodendrocytes. This work suggests the loss of NPC1 function and lysosomal cholesterol accumulation leads to energy depleted oligodendrocytes and myelin that could affect axonal function.

P2.17 | Combining Proteomics and Lipidomics for a comprehensive analysis of homeostasis control in the brain

Leena M. Patil[1] (leena2411@gmail.com), Rebecca Burlett[1], Holden Kelly[1], Yiran Liang[1], Kenneth Virgin[1], Ryan Kelly[1], John C. Price[1]
[1]Brigham Young University, Provo, UT

Enzymes play a critical role in protein and lipid metabolism. Loss of metabolic regulation has been shown to contribute to neurodegeneration. Hence, understanding the dynamic relationship between metabolite flux and enzyme regulation may improve our ability to treat and prevent neurodegeneration. Metabolic pathways are dynamically regulated at the level of synthesis and degradation. However, measuring protein or lipid concentrations over time does not report the changes in synthesis and degradation that occur as the cell works to maintain homeostasis. These rates of synthesis and degradation can be measured in the form of turnover rates. Using a combination of mass spectrometry and stable isotope labeling allows us to measure turnover rates and capture both quantitative and kinetic information for detected species. From this data we can predict enzyme-lipid relationships. Using deuterated water to label proteins and lipids in mice, isotope incorporation was measured over multiple timepoints using mass spectrometry (MS). This data was then analyzed to yield turnover rates. We analyzed the entorhinal, visual, and hippocampal regions in tissue slices using Laser Capture Microdissection (LCM) to do targeted microsampling and AutoPOTs preparation for proteomic samples. We also formalized a quantitative 'sample addressing' protocol for establishing the location of samples within each region. For lipid metabolism, Desorption Electrospray Ionization (DESI) was optimized and applied.

Using these methods, we have compared regulation of ~2500 proteins and ~20 lipids with highly individualized patterns of protein expression throughout the brain. We have also been able to determine specific regulation, synthesis, and degradation trends for specific proteins and lipids in each region and compared this across regions. This has culminated in the comparison of concentration and turnover for lipids and associated enzymes to test for metabolic patterns connected to neurodegeneration.

P2.18 | Proteomic Changes associated with Race and Executive Function in Patients with Mild Cognitive Impairment

Khiry L. Patterson[1] (khiry.l.patterson@vanderbilt.edu), Julia B. Libby[2], Albert B. Arul[1], Min Ji Choi[1], Nekesa C Oliver[1], Logan Dumitrescu[3], Timothy J Hohman[3], Angela Jefferson[3], Renã A S Robinson[4]
[1]Department of Chemistry, Vanderbilt University, Nashville, TN, [2]Vanderbilt Memory and Alzheimer's Center, Vanderbilt University Medical Center, Nashville, TN, [3]Department of Neurology, Vanderbilt University Medical Center; Vanderbilt University Brain Institute, Vanderbilt University Medical Center; Vanderbilt Memory and Alzheimer's Center, Vanderbilt University Medical Center, Nashville, TN, [4]Department of Chemistry, Vanderbilt University; Department of Neurology, Vanderbilt University Medical Center; Vanderbilt Institute of Chemical Biology, Vanderbilt University; Vanderbilt Brain Institute, Vanderbilt University Medical Center; Vanderbilt Memory and Alzheimer's Center, Vanderbilt University Medical Center, Nashville, TN

Introduction: Mild cognitive impairment (MCI) predisposes conversion to Alzheimer's disease (AD); therefore, biomarkers of MCI may be useful in assessing future clinical progression. However accessible blood-based biomarkers for MCI remain largely broad-scale and often neglect diverse populations. Our laboratory employed a discovery-based, quantitative plasma proteomics analysis of diverse cognitively normal (control) and MCI patients from the Vanderbilt Memory and Aging Project (VMAP) to identify potential

biomarker candidates and discern the impact of self-identified race on the MCI proteome.
Methods: Plasma samples from 335 VMAP participants included non-Hispanic White/NHW (N=290, age=73±7, 38% male, MCI=39%) and African American/Black (N=35, age=71±8, 69% male, MCI=43%) individuals. Proteins from each sample were depleted (MARS-14) of abundant proteins and subjected to automated digestion and TMTpro-labelling using a robotic liquid handler. Labeled peptides were separated with high pH reverse-phase fractionation and quantified by liquid chromatography-tandem mass spectrometry. Multiple linear regression models of diagnosis, cross-sectional memory, and executive function (EF) were associated with protein abundance in race-stratified comparisons. Statistical significance was based on the False Discovery Rate-corrected p-value (<0.05).
Results: Overall, our proteomics analysis identified 3,784 plasma proteins at high confidence across 335 VMAP participants. Robust quality control, post-acquisition filtering, and data transformation analyses identified 489 plasma proteins detected across all participants, which were carried forward for regression analysis. Among NHW participants, 21 proteins significantly associated with an MCI diagnosis, 12 proteins with baseline memory, and two proteins with EF measures. Interestingly, among Black participants five proteins (tissue factor pathway inhibitor, malate dehydrogenase 1, semaphorin-7A, fructose-biphosphate aldolase B, and N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase) were significantly associated with reduced baseline EF. These proteins had a negative association in Black participants, and a positive association in NHW participants.
Discussion: This study identified 21 plasma proteins associated with MCI diagnosis and importantly, five proteins associated with cross-sectional executive function in African American/Black adults.

P2.19 | Proteome fidelity in the brain during aging

Nalini Rao[1] (nalinirao2018@u.northwestern.edu), Arun Upadhyay[1], Jeff Savas[1]
[1]Northwestern University, Chicago, IL

Aging represents a predominant risk factor for many diseases, most notably neurodegenerative diseases. Furthermore, neuronal proteostasis deficits have been implicated in age-related diseases and lead to protein accumulations. As proteostasis represents a heavily intertwined balance of synthesis and degradation, exploring protein turnover dynamics during aging would be especially important for uncovering changes in post-mitotic cells, such as neurons. In the brain, investigating aging proteome dynamics is especially crucial as maintaining protein homeostasis is delicate and complex. Developing a profound understanding of physiological changes associated with aging in the brain could provide insights into underlying mechanisms of disease. An understanding of these dynamics will offer insight into components of aging that may be susceptible or accelerated in age related diseases. Utilizing wild-type male and female mice at five different ages, we dynamically labelled animals with heavy nitrogen-15 (15N) for three months to create a pulse-step paradigm of protein turnover dynamics with aging in which 14N and 15N represent old and new proteins, respectively. After tissue collection, isobaric TMT labelling of cortical brain tissue was performed which then allowed for a quantification and analysis of proteome-wide changes in turnover that occur with aging. Next, to probe why select groups of proteins displayed slowed protein turnover with age, we isolated the SDS insoluble proteome with ultracentrifugation to probe misfolded proteins, and finally used MS-based identification of ubiquitination sites in the proteins identified to investigate ubiquitination changes. Our data demonstrates the dynamic nature of protein turnover with aging and how it varies between male and female animals. Using quantification of protein turnover dynamics of persisting proteins in the aged cortical proteome, we found that male mice experienced slowed protein turnover dynamics earlier when compared to female mice. In order to further investigate proteome-wide turnover dynamics, we utilized a clustering analysis

P2.20 | New approaches for identifying domains of chromatin decondensation in aging

Stephanie Stransky[1], Sarah Graff[1], Ronald Cutler[1], Jennifer Aguilan[1], **Simone Sidoli[1] (simone.sidoli@gmail.com)**
[1]Albert Einstein College of Medicine, Bronx, NY

Aging is a familiar phenotype also for non-scientists, characterized by hallmarks such as genomic instability, dysfunctional repair pathways and chronic cell function decay. Although, “what causes aging” remains a question without a definitive answer. DNA damage is the most quoted molecular mechanism as primary cause of senescence and aging. However, DNA mutations are random, different from cell-to-cell, and different between organisms. How a random phenomenon leads to these very familiar phenotypes we just discussed?
Our lab aims to characterize what cells share during their aging progression, which we identified to be the decondensation of silent/compacted heterochromatic domains. To do so, we have established an innovative 3D cell culture system to model anomalously decondensing chromatin, and we are the proteomics lab in charge of characterizing the chromatin proteome of the LonGenity program. LonGenity is a cohort of hundreds of donors, including centenarians and individuals who have centenarian parents, i.e. individuals with a projected exceptional longevity.
This poster will present data from three major projects from our group: (i) the identification of proteins that recognize and modulate decondensing chromatin domains, which we speculate play a critical role in delaying chromatin decay and spurious transcription

during aging. (ii) We have also identified histone succinylation to be a chromatin modification enriched on the chromatin of individuals with exceptional longevity, and we present data on why histone succinylation could be a protective marker for anomalously decondensing chromatin. (iii) We also present our efforts in the development of a single cell proteomics platform to study the heterogeneity of chromatin modifications in populations of aging cells. This last project is critical to resolve how heterogeneity in abundance of histone modifications could explain the increasing transcriptional noise during aging.

P2.21 | Mitotic inhibitors and 3D cell culture induce diverse pluripotent stem cell-derived neurons

Jessica Roberts[1] (jessica.roberts2@nih.gov), Erika Lara Flores[1], Daniel Ramos[1], Luke Reilly[1], Lirong Peng[1], Ziyi Li[2], Andrew Singleton[2], Michael Ward[2], Mark Cookson[1], Yue Andy Qi[3]
[1]National Institutes of Health, Bethesda, MD, [2]National Institute on Aging, Bethesda, [3]NIA/NIH, Bethesda, MD

The human pluripotent stem cell (iPSC)-derived neuron models have been widely applied to studies of neurological diseases. A thoroughly characterized iPSC line, KOLF2.1J, has been selected as the parental line for the iPSC Neurodegenerative Disease Initiative (iNDI) at the NIH. Through genetic engineering, Alzheimer's disease and related dementia-associated variants will be introduced into the KOLF2.1J line, and the effects of these variants will be investigated in KOLF2.1J-derived neurons. Here, we report a series of temporal proteomic and single-cell transcriptomic profiling of KOLF2.1J differentiation using six different commonly used neuronal differentiation protocols. Specifically, we compared cells grown in SMAD inhibitors, mitotic inhibitors, and 3D cultures using bioreactor systems and 2D cultures. Additionally, we evaluated neuronal differentiation induced by overexpression of transcription factor neurogenin 2 (NGN2) and co-expression of EMX1 (NGN2-EMX1). As expected, the single-cell transcriptomic data shows >80% cell populations are mature neurons, and ~10% are progenitor and pluripotent cells. The temporal proteomic analyses suggest similar proteome patterns of 6 different protocols from day 0 to 3; however, very diverse neuronal populations were observed from day 7 to 28. The integrated UMAP analyses indicate mitotic inhibitors and 3D cell culture protocols induce the most significant diversity compared to the NGN2 protocol. The 3D culture significantly increases synapse and microtubule markers, and the mitotic inhibitors increase the mature and cortical neuron markers, whereas SMAD inhibitors and NGN2-EMX1 protocols show relatively minor changes. Furthermore, we have generated an accessible and interactive web-based browser of the results from this study, which serves as reference and rich resource for the iNDI project. Overall, our findings demonstrate a diverse iPSC-derived neuron population induced by different differentiation conditions.

P2.22 | Investigating the Protein Conformational Landscape of Cognitive Decline in Aging in a Rodent Model Using Limited-Proteolysis Mass Spectrometry

Haley Tarbox[1] (htarbox1@jhu.edu), Audrey Branch[1], Michela Gallagher[1], Stephen D. Fried[1]
[1]Johns Hopkins University, Baltimore, MD

Cognitive decline in aging is becoming a major issue for families and healthcare providers as the population ages. Protein misfolding and aggregation are frequently implicated as driving many neurodegenerative diseases; however, this view is informed by a relatively small number of proteins, such as Aβ, α-synuclein, and tau. Limited proteolysis mass spectrometry (LiP-MS) is an emerging technique which can provide structural information about protein conformational states proteome-wide. In this approach, pulse proteolysis with a non-specific protease enables flexible and solvent-accessible regions within proteins to be recorded as cleavage sites, which are subsequently read out with mass spectrometry. This work leverages a genetically variable aged rat model, which has relevance as a preclinical model for Alzheimer's disease. In this colony, rats naturally experience distinct levels of cognitive decline during aging. We used LiP-MS to evaluate structural differences across the proteomes of neurons in three regions of the hippocampus and compared the conformational profile between aged-impaired and aged-unimpaired populations. We matched over 3000 proteins, with 10.7% of hippocampal CA1 proteins and 3.9% of hippocampal DG proteins identified as hits with potential conformational differences between aged-impaired and aged-unimpaired rats. Biophysical features that correlate with age-related conformational change also correlate with nonrefoldability, suggesting that proteins which possess spontaneous refoldability are less likely to misfold with age. Our study demonstrates that protein conformational states or other structural changes may be an important explanatory factor in assessing molecular determinants underpinning cognitive decline.

P2.23 | Semi-quantitative proteomics reveals novel therapeutic targets for Alzheimer's disease

Arun Upadhyay[1] (arun.upadhyay@northwestern.edu), Robert Vassar[2], Jeff Savas[1]
[1]Northwestern University, Chicago, IL, [2]Northwestern University Feinberg School of Medicine, Chicago, IL

Alzheimer's disease (AD) is the most prevalent neurodegenerative disorder affecting more than five million people in the United States. AD progresses at a slow rate and causes a gradual decline in cognitive abilities leading to dementia. The AD patients' brain tissues are loaded with deposits of amyloid beta (Aβ) plaques and neurofibrillary tangles of hyper-phosphorylated Tau protein. An increased resurgence in the amyloid cascade hypothesis has been observed, following the approval of aducanumab and the promising effects shown by Lecanemab, two antibodies targeting Aβ peptides and other higher-order structures, including fibrils. The importance of Aβ fibrils has long been recognized, yet, our understanding of plaque development, structure, and composition is subtle. The most intriguing among all is the magnitude of their diversity and their relation to the surrounding neuronal circuits. In past studies, a large number of amyloid-associated proteins were identified with plaques. However, the inefficacy in plaque/fibrils purification and lack of consistency in the proteomics dataset limit scope of many such studies. We have improvised the traditional biochemical isolation methods and achieved great purity with high reliability. We consistently recovered a consensus pool of proteins from Aβ-rich fibrils isolated from highly diverse and complex biological samples, including AD patients' postmortem brain samples and relevant mouse models. We hypothesized that some of these Aβ interaction partners could be active modifiers of amyloid formation and maturation. Therefore, by assessing *in vitro* aggregation kinetics of recombinant Aβ42 and assessing the neurotoxicity in Aβ42 overexpressing flies, we confirmed that multiple proteins could influence Aβ aggregation and maturation. These newly identified top protein candidates may pave way for the identification of multiple new drug targets.

P2.24 | Can Bacteria Turn α-Synuclein from Friend to Foe?

Tonya N. Zeczycki[1] (zeczyckit@ecu.edu), Whitney G. Bond[1], Kellyn S. Trowse[1], MIthil Patel[1]
[1]Brody School of Medicine at ECU, Greenville, NC

More than 10 million people worldwide are currently living with Parkinson's disease (PD), and within the US alone, the projected prevalence of PD in the next 15 years will exceed 1.6 million people. α-synuclein (αSyn), an intrinsically disordered protein implicated in PD pathogenesis, has a significant physiological role in the innate immune response. Importantly in the gut, αSyn amplifies immune response signals by promoting macrophage polarization. However, it is not known what changes occur which turns αSyn from a protein serving its normal function in the immune system to a protein associated with disease. Recent clinical and foundational research suggests that bacterial amyloid proteins produced by gut bacteria may be the missing link. Bacterial amyloids expressed on the surface of bacteria and biofilms in the gut are thought to interact with αSyn, promoting the formation of hybrid oligomers, capable of "hyperactivating" macrophages, however structural and functional studies providing cause-and-effect evidence for such claims are lacking. We have used structural and untargeted proteomic and metabolomics studies to identify how alterations in bacterial metabolism promote the oxidative environment promoting αSyn oligomerization. Further studies using *E. coli* curli oligomers and untargeted proteomics, metabolomics, and lipidomics show the impact αSyn and curli monomers, oligomers, and hybrid oligomers have on macrophage activation. We found that treatment of macrophages with curli oligomers in the presence of αSyn not only increased the expression proteins associated with a pro-inflammatory phenotype compared to αSyn monomers or oligomers alone, but also prevented macrophages from phagocytosing αSyn. Further, we observed distinct changes in macrophage metabolism after exposure to both αSyn and curli oligomers, indicative of a more oxidative environment. Taken together, these MS experiments begin to unravel the tangle of interactions between bacterial and human amyloid proteins and their potential role in promoting inflammatory response in neurodegenerative diseases. #TeamMassSpec #PIChallenge

P2.26 | Proteomic Characterization of a Novel Antibody for FTLD-TDP Pathologic TDP-43

Anika Wilen[1] (anika.wilen@northwestern.edu), Qinwen Mao[2], Jeff Savas[1]
[1]Northwestern University, Chicago, IL, [2]University of Utah, Salt Lake City, UT

Frontotemporal lobar dementia (FTLD) is the third most prevalent degenerative cause of dementia and can be characterized based on the contents of patients' pathological inclusions. One of the categories of FTLD, FTLD-TDP, is characterized by transactive response DNA binding protein with M_r 43 kD (TDP-43) abundant inclusions. A novel monoclonal antibody was produced that targets the c-terminal glycine-rich domain (GRD) of TDP-43 and seems to be strongly reactive to pathological TDP-43 inclusions with minimal reactivity to normal TDP-43. The aim of this study is to validate this novel antibody (MAb#9) using western blot, liquid chromatography with tandem mass spectrometry (LC-MS/MS), and immunoprecipitation using MAb#9 on the plasma and cerebrospinal fluid (CSF) of patients with a variety of neurodegenerative dementia diagnoses. By examining the MAb#9-detected pathological TDP-43 levels in the plasma and CSF of patients with Alzheimer's Disease, FTLD-TDP, FTLD-Tau, Pick's Disease, and

other dementias, we can validate MAb#9's ability to discern between dementias with pathological TDP-43 and without TDP-43. Potentially, MAb#9 can be used to create an assay to differentiate between FTLD-TDP and other dementias to diagnose patients faster and in a more cost-effective way.

P2.28 | Investigating Age- and Race-Associated Changes in the Mammalian Ovarian Proteome to Identify Mechanisms of Ovarian Cancer Pathogenesis

Shweta Dipali[1] (ShwetaDipali2023@u.northwestern.edu), Christina D. King[2], Jacob P. Rose[2], Judith Campisi[2], Joanna E. Burdette[3], Birgit Schilling[2], Francesca E. Duncan[1]
[1]Northwestern University, Chicago, IL, [2]Buck Institute for Research on Aging, Novato, CA, [3]University of Illinois at Chicago, Chicago, IL

High-grade serous ovarian cancer (HGSOC) is the most lethal cancer of the female reproductive system and is frequently diagnosed post-menopause, suggesting that the aging ovary may provide a niche for tumor progression. In fact, ovarian cancer cell adhesion to ex vivo cultured ovaries from reproductively old (10+ months) mice is increased two-fold compared to young (6-12 weeks) counterparts. Thus, we hypothesize that age-associated changes in the ovary enhance HGSOC cell adhesion and expansion. We performed data-independent acquisitions using an Orbitrap Exploris 480 to identify proteins that are significantly-altered in the ovary with age, which may impact HGSOC pathogenesis. To determine how aging alters the ovarian matrisome, we developed a protocol to enrich for the ECM prior to proteomic analysis. H&E staining of tissue sections revealed that 0.1% SDS effectively removes nuclei, while preserving ECM architecture in 12.5 hours for mouse ovaries and 24 hours for human ovarian tissue pieces. PicroSirius Red staining revealed that 0.1% SDS treatment maintains the collagen matrix and the age-dependent increase of collagen. Proteomic analysis of the native and ECM-enriched mouse ovary identified 4,721 proteins, of which 383 were significantly-altered with age. 71 of the significantly-altered proteins were either core matrisome or matrisome-associated. Gene ontology analysis of significantly-altered proteins revealed that pathways associated with immune function and ECM organization are upregulated with age and pathways maintaining genomic stability were downregulated. Preliminary proteomic analysis of human ovarian tissue pieces identified 1,065 proteins, 181 of which were significantly-altered in different sub-compartments of the ovary. These results identify a unique proteomic signature of the aging ovary and candidate proteins that can be further interrogated in mechanistic studies. HGSOC disproportionately affects non-Hispanic Black (NHB) women, so proteomic analysis of ovarian tissue from age-matched NHB and non-Hispanic white participants is ongoing to identify pathways that may contribute to this racial disparity.

P3: Biomarkers and Targeted MS Assays

P3.01 | A Nested Active Learning Model for Clinical LC-MS/MS Data Analysis

Duran Bao[1] (dbao@tulane.edu), Qingbo Shu, Micheal Tang[2], Christopher Lyon, Tony Hu, Jia Fan[1]
[1]Tulane University, New Orleans, LA, [2]University of Michigan, Ann Arbor, MI

Targeted LC-MS/MS has become a promising analytical technique for biomarker identification and quantification in disease diagnosis. However, its clinical application is limited because technicians with extensive MS experience are required for data analysis. A fully automated MS data analysis tool for clinical diagnosis will accelerate MS application in healthcare. Machine learning (ML) has already made great achievements in the healthcare field but needs to catch up in clinical MS applications. Therefore, we built ML-based novel pipelines by 1) Developing a comprehensive 80-feature extraction by combining the statistical and morphological features, which were produced by absolute and normalized MS signals. 2) Enabling small data sets for model development by evaluating different ML algorithms for pipeline construction. 3) Applying a nested active learning (nActL) algorithm to control the quality of the training set (TrainS). Among the three ML algorithms, the random forest (RF) model showed an area under curve (AUC) value of 0.98±0.01, which was statistically higher than the other two, logistic regression (LR) and gradient boosting (GB), which showed AUC of 0.89±0.02 and 0.87±0.02, respectively. We further assessed the models' reproducibility and generalizability in three independent data sets. The RF-model consistently provided the best performance compared to LR-model and GB-model in this external validation (average AUC: RF, 0.94±0.02; LR, 0.84±0.02; GB, 0.77±0.03). We also hypothesized that data quality matters more than data quantity. Therefore, a nActL algorithm was further applied to control the quality of the TrainS. Even though using a small and unbalanced TrainS, the nActL algorithm provided higher and more stable sensitivities of 0.92±0.03 or 0.86±0.08 than the model without applying quality control which showed a sensitivity of 0.73±0.12 or 0.66±0.08 when the positive proportion of the TrainS was below 25% or above 75%, respectively. This study constructed a robust, high-performance ML classifier to enable automatic clinical MS-based diagnosis.

P3.02 | Multiomic analysis of clear cell renal carcinoma

Aron Bartha[1,2] (aron.bartha@nyulangone.org), Gyongyi Munkacsy[2], Zsuzsanna Darula[3], Eva Klement[3], Peter Nyirady[4], David Fenyo[1], Balazs Gyorffy[2]
[1]Institute for Systems Genetics, NYU School of Medicine, New York, [2]Department of Bioinformatics Semmelweis University, Budapest, Hungary, [3]Laboratory of Proteomics Research, BRC, Szeged Hungary, Szeged, Hungary, [4]Department of Urology Semmelweis Unveristy, Budapest, Hungary

Introduction: Clear cell renal carcinoma is the most frequent form of renal malignancy, with an increasing incidence rate worldwide. In this study we used a multiomic approach to differentiate normal and tumor tissues in clear cell Renal Cell Carcinoma (ccRCC).
Methods: Using transcriptomic data of patients with malignant and adjacent normal tissue samples from gene chip and RNA-Seq cohorts, we identified the top genes over-expressed in ccRCC. We collected surgically resected ccRCC specimens to further investigate the transcriptomic results on the proteome level. The differential protein abundance was evaluated using targeted mass spectrometry (MS).
Results: We assembled a database of more than 600 renal tissue samples from NCBI GEO and TCGA and used these to uncover the top genes with higher expression in ccRCC. For protein level analysis 162 malignant and normal kidney tissue samples have been acquired. The most consistently upregulated genes were IGFBP3, PLIN2, PLOD2, PFKP, VEGFA, and CCND1 ($p<1E-05$ for each gene). Mass spectrometry further validated the differential protein abundance of these genes (IGFBP3, $p = 7.53E-18$; PLIN2, $p = 3.9E-39$; PLOD2, $p = 6.51E-36$; PFKP, $p = 1.01E-47$; VEGFA, $p = 1.40E-22$; CCND1, $p= 1.04E-24$). We also identified proteins correlating with overall survival.
Conclusions: We used transcriptomic and proteomic data to identify a minimal panel of proteins highly specific for clear cell renal carcinoma tissues. The introduced gene panel could be used as a promising tool in the clinical setting.

P3.03 | A multiplexed quantitative analysis of naturally occurring single amino acid variants by targeted proteomics in non-depleted human plasma

Panshak P. Dakup[1] (panshak.dakup@pnnl.gov), Tai-Du Lin[1], Adam C. Swensen[1], Athena M. Schepmoes[1], Thomas L. Fillmore[1], Tujin Shi[1], Wei-Jun Qian[1], Jon M. Jacobs[1], A2CPS Consortium[2]
[1]Pacific Northwest National Laboratory, Richland, WA, [2]NIH Commonfund, Richland, WA

Single amino acid variants (SAAVs) in protein sequences are often a result of single nucleotide polymorphisms (SNPs). There are several naturally occurring SNPs in the human population that are associated with diseases; they either increase the risk of

development or act as protective factors. SAAVs of Complement Factor B (R32Q and R32W), Clusterin (N317H), Fetuin B (K360R), and Kininogen (L212P) have shown biological, and in some cases clinical, relevance in macular degeneration, chronic hepatitis B, Alzheimer's disease, leukemia, and William's trait conditions respectively. Therefore, the precise identification and quantification of these SAAVs is critical for predicting and diagnosing their associated diseases. We have developed a targeted assay that monitors the variant-wildtype peptide pairs using heavy internal standards of 9 peptide targets. By screening 70 healthy human plasma samples, we identified these variants at a slightly higher frequency relative to the reported global allele frequencies. We also measured significantly higher concentrations of Complement Factor B and Fetuin B variants and a significantly lower concentration of the Clusterin variant relative to their wildtype counterparts. Furthermore, our results distinguished phenotypes of homozygosity and heterozygosity of SAAV presence of Complement Factor B through direct concentration level characterization. Ongoing work includes 1) establishing the assay's figures of merit to assess detection and quantification limits, repeatability, and reproducibility, and 2) validating our proteome level findings with genomic data by SNP genotyping. These findings will provide additional detection methodologies which will increase throughput for personalized precision medicine applications.

P3.04 | dia-PASEF for targeted proteomics: development of large-scale assay for quantitation of more than 500 proteins in human plasma sample

Stephanie Kaspar-Schoenefeld[1], Sebastian Mueller[2], Tejas Gandhi[2], **Sira Echevarria-Zomeno[2] (sira.echevarria@biognosys.com)**, Markus Lubeck[3], Gary Kruppa[4]
[1]Bruker Daltonik GmbH & Co. KG, Bremen, Bremen, Germany, [2]Biognosys, Schlieren, Zurich, Switzerland, [3]Bruker Daltonik GmbH, Bremen, Bremen, Germany, [4]Bruker Daltonics, Billerica, MA

Introduction: dia-PASEF merges the benefits of DIA with the advantages of ion mobility in proteomics experiments making it an advantageous method to be integrated into a platform for large-scale biomarker studies without the need for in-depth method optimization. Here, we use dia-PASEF in combination with the PQ500 kit to develop a targeted quantitation assay for peptides in human plasma samples.
Methods: Plasma samples were digested using the iST kit from PreOmics. The PQ500™ kit (Biognosys) was spiked into the prepared digests. Tryptic peptides were separated on a nanoElute system coupled to a timsTOF HT mass spectrometer via a CaptiveSpray source using a 30-min ACN gradient. For the dia-PASEF acquisition, a window placement scheme consisting of 6 TIMS ramps with 3 mass ranges per ramp was applied. Data was processed in Spectronaut (v16, Biognosys) using an ion mobility annotated PQ500 library.
Results: Here, we developed a targeted quantitation assay for human plasma proteins using dia-PASEF. The major advantage of the approach is that there is no need for tedious method development as is typically required for targeted approaches like SRM and MRM. The assay was applied to a proof-of-concept study of non-depleted plasma samples from patients diagnosed with lung cancer. In total, 663 peptides and 463 protein groups were identified and quantified, covering around 80% of the PQ500 panel. Of those, 55 proteins were found to be significantly regulated with three of them (Fibronectin, Immunoglobulin lambda-like polypeptide 1, Immunoglobulin lambda-like 1 light chain) detected to be higher abundant in healthy donors.
Conclusions: We developed a workflow for targeted quantitation in non-depleted human plasma using dia-PASEF and applied it to a biologically relevant lung cancer study. Significantly regulated peptides are known to be associated with cancer, which confirms the potential of applying the presented approach to clinical research studies.

P3.05 | High-flow plasma profiling with LC-HRAM-MS

Richard Gibson[1] (Richard.Gibson@thermofisher.com), Stephanie Samra[1], Cristina Jacob[1], Philip Remes[1], Yvonne Song[1]
[1]Thermo Fisher Scientific, San Jose, CA

Introduction: Human proteome analysis through liquid chromatography mass spectrometry (LC-MS) can play an important role in biomarker discovery in matrices such as plasma, serum or urine. Analyzing digested human plasma allows the detection of 1,000s of peptides within one run. Bioinformatic analysis then allows the identification of signature peptides that may be used to identify protein biomarkers of a disease. Data from high-flow LC-MS has been shown to be able to identify 100s of proteins within one run that may then be used as targets for future studies. MS2 scans can be data-dependent (DDA), data-independent (DIA) or targeted, such as with parallel reaction monitoring (PRM). Each of these scan types is accessible through high resolution accurate mass (HRAM) mass spectrometry (MS) and have been demonstrated to efficiently screen for proteins in plasma.
Results: Protein and peptide IDs were determined from DDA, DIA and PRM scans (Table 1). Precursor transitions for the PRM scans were chosen using seven small window gas phase fractionation (GPF) DIA scans. Selectively reducing the database size was demonstrated to increase the resolving power of the identification method, significantly increasing the number of protein and peptide IDs for DDA and PRM scans.

Table 1: Protein (and peptide) IDs from 45 ug plasma digests.

	Protein (peptide) IDs
DIA	186 (672)
DDA	268 (1803)
DIA GPF	270 (1513)
PRM	166 (610)

Conclusions: Protein and peptide IDs were calculated for DDA, DIA and PRM data, highlighting the versatility of HRAM mass spectrometry. Successful collection of DDA, DIA and PRM data highlights orbitrap technology being suitable for acquiring discovery data and translating it to a targeted method on the same instrument.

P3.06 | In-depth identification and accurate quantification of mitochondrial and lysosomal crosstalk proteins

Young Ah Goo[1] (ygoo@wustl.edu), Byoung-Kyu Cho[1]
[1]Washington University in St. Louis, Saint Louis, MO

Mitochondria and lysosomes coordinate a pivotal role in cellular metabolism by physical and functional interaction. The alterations in their crosstalk are known to be involved in muscular and neurodegenerative disorders such as Parkinson's, Alzheimer's, Huntington's, and amyotrophic lateral sclerosis. We established a method for organelle-specific in-depth proteome profiling without the use of enrichment, followed by targeted quantification of crosstalk proteins for mitochondria and lysosomes by dia-PASEF and prm-PASEF. Three different human cell lines, HeLa, K562, and HEK293, were analyzed by dia-PASEF and prm-PASEF methods using timsTOF Pro2 with nanoElute system. The dia-PASEF data was queried against Bruker's built-in spectral library. The experimental collision cross section (CCS) and the normalized retention time (RT) values of identified peptides were applied to the prm-PASEF analysis. The prm-PASEF data was processed using the Skyline software. By employing the dia-PASEF analysis using whole cell lysate without the organelle-specific enrichment, over 1200 and 500 mitochondria- and lysosome-specific proteins, respectively, were identified from three human cell lines. The in-house spectral library was generated for all identified peptides based on the CCS and the normalized RT values. The prm-PASEF analysis was developed to target 103 mitochondrial and 84 lysosomal proteins based on the three molecular functions, amino acid metabolism, autophagy, and calcium homeostasis, known for their functional crosstalk between organelles. To evaluate the analytical performance, we analyzed 15 independent prm-PASEF datasets generated by three biological/five technical replicates and confirmed that the majority of the target peptides showed less than 15% of the coefficient of variation across all the data. Our unbiased approach demonstrated that the use of dia-PASEF and prm-PASEF provides in-depth organelle-specific profiling and accurate quantification for tracking key organelle proteins involved in disease-specific crosstalk.

P3.07 | Trigger-based mass spectrometry for proteogenomic discovery and expanded detection of the alternative proteome

Erin Jeffery[1] (edf4n@virginia.edu), Saikat Bandyopadhyay[2], Qing Yu[3], Ben T. Jordan[4], Madison M. Mehlferber[2], Leon Sheynkman[2], Jamie Saquing[2], Aidan Fenix[5], Charles Murry[5], Gloria M. Sheynkman[2]
[1]University of Virginia, Charlottesville, VA, [2]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA, [3]Department of Cell Biology, Harvard Medical School, Boston, MA, [4]Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Frederick, MD, [5]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA

Alternative splicing (AS) is a major contributor of transcriptomic complexity. Confirming protein expression of alternative isoforms is essential for understanding their role in health and disease. However, due to the high degree of sequence similarity between alternative protein isoform sequences, most peptides detected in traditional data-dependent acquisition (DDA) mass spectrometry (MS) are shared and thus uninformative for isoform detection. Despite only 9% tryptic peptides (9-30 AA) being unique to isoforms, they can inform on over 46% of isoforms. Therefore, increasing their sampling rates could significantly improve isoform characterization. Advanced targeted acquisition MS could be applied to address this problem, such as the Tomahto strategy, which enables real-time decision making and is capable of targeting peptide lists greater than 1,000 with TMT-based sample multiplexing. Here, we integrated Tomahto with a "long read proteogenomics" workflow to target isoform-specific peptides in a human cell line (WTC-11). As a proof-of-concept, we selected 192 isoform-distinguishing peptides from 55 genes, including peptides that distinguish the major and minor (alternative) isoforms. Analyzing the same 8 high pH fractions, we compared detection rates between Tomahto and DDA analysis, and found that Tomahto out-performed DDA, doubling the number of isoform-specific peptides identified (61 versus 31). The Tomahto approach yielded 61 out of 192 (31.8%) isoform-distinguishing peptides: 35 out of 82 (42.7 %) from major and 26 out of 110 (23.6 %) from minor isoforms. Furthermore, we found that Tomahto results from the 8 fractions still outperformed a deep proteome DDA analysis of 30 fractions (Tomahto: 61 peptides, Deep DDA: 45 peptides). We observed seven major-minor isoform pairs identified exclusively by our method, including an isoform of C1orf52, which is involved in RNA binding. These data

demonstrate the power of integrating advanced targeted discovery approaches with a long-read proteogenomics pipeline for proteomic confirmation of known and novel isoforms.

P3.08 | On-Column Digest Followed By SPE Desalting For Rapid And Simplified Sample Processing without Sample Transfer

John Laycock[1] (john.laycock@tecan.com), Heather Eastwood[2], Fabian Wendt[3], Manuel Bauer[3]
[1]Tecan SP, Baldwin Park, CA, [2]Tecan, Baldwin Park, CA, [3]Tecan, Männedorf, Switzerland

Bottom-up proteomics is widely used for protein identification and post-translational modification analysis. In this approach, proteins are first digested with an enzyme into peptides, and then analyzed by mass spectrometry. However, prior to analysis, samples must be purified of impurities and contaminants, such as salts and buffers that are used in the digestion. In this presentation, we demonstrate the simple and rapid analysis of trypsin-digested bovine serum albumin (BSA) by coupling on-column digestion and peptide clean-up by Solid Phase Extraction (SPE) utilizing Atlas® sorbent in the same column using Tecan's Narrow Bore Extraction (NBE) technology. NBE columns offer potential workflow advantages due to their unique anti-flow airgap between the upper reaction vessel chamber and bottom SPE sorbent chamber. This airgap prevents liquid from entering the lower sorbent chamber until positive pressure is applied. The empty upper chamber of the NBE column can be used independently of the lower chamber as a vessel for sample preparation. This upper chamber is widely used for assays utilizing enzymatic hydrolysis prior to SPE. When coupled with the Tecan Resolvex® A200 for solvent dispensing and positive pressure application, the workflow is further streamlined. The aim of the experiment is to demonstrate the benefit of using NBE columns as the reaction vessel as well as the sample clean-up column for proteomics workflows, resulting in a simplified and fast sample processing procedure.

P3.09 | Progranulin alterations in the Niemann-Pick type C1 I1061T mouse model

Wenping Li[1] (jklwp007@uic.edu), Thu T. Nguyen[2], Ying Yan[2], Stephanie M. Cologna[2]
[1]University of Illinois at Chicago, Chicago, IL, [2]University of Illinois Chicago, Chicago, IL

Niemann-Pick Type C1 (NPC1) is a fatal, progressive cerebellar neurodegenerative disease with a hallmark of accumulated unesterified cholesterol in the lysosomal. To investigate the differential proteome in an NPC1 mouse model containing the most common pathological variant I1061T (*Npc1*$^{I1061T/\ I1061T}$), cerebella from 9-week-old mice were used. We observed 141 significantly upregulated proteins and 123 downregulated proteins in *Npc1*$^{I1061T/\ I1061T}$ mice compared to controls. Particularly, progranulin (PGRN), which was upregulated, was chosen to be further studied due to its multiple roles in the brain, including regulating lysosomal biogenesis and functions as a neurotrophic factor in neuronal cells, as well as participating in (anti-)inflammation in microglia. In this study, the increased PGRN in the whole cerebellum lysate was validated by western blot on 7-, 9-, and 15-week-old *Npc1*$^{I1061T/\ I1061T}$ mice. Furthermore, the increased PGRN is also observed in the peripheral liver tissue in 4-, 9-, and 15-week-old *Npc1*$^{I1061T/\ I1061T}$ mice. To test PGRN expression levels in the cerebellum at the cellular level, immunohistochemistry (IHC) was performed on 4-, 9-, and 15-week-old mouse brain slices. The number of IBA1-positive microglia, representing activated microglia, was increased in 9- and 15-week-old mutant mice, consistent with prior studies. Since PGRN is a secreted protein mainly from microglia and is taken up by neurons for various functions, we sought to interrogate the level of secreted PGRN using a mouse microglial cell line treated with or without U18666A, an NPC1 inhibitor. We observed that PGRN was increased in U18666A treated microglia, and decreased in the media compared to control, indicating an impairment of PGRN secretion in absence of functional NPC1. Our studies suggest that in the brain, less PGRN is available for neurons which then contributes to neurodegeneration. Finally, to translate this observation in patients, CSF or blood samples will be used to measure PGRN expression level.

P3.10 | Isobaric labeling-based relative quantification for discovery of serum metabolite biomarkers associated with growing early colorectal adenomas

Yuan Liu[1] (liu788@wisc.edu), Lingjun Li[1], Hua Zhang[1], Zicong Wang[1], William Dove[1], Mark Reichelderfer[1], Perry Pickhardt[1], Zhijun Zhu[1]
[1]University of Wisconsin-Madison, Madison, WI

A major challenge for reducing the death rate of colorectal cancer (CRC) is to screen patients using low-invasive testing. Compared to current methods, blood test is likely to have higher compliance rate with reduced invasiveness. The purpose of this work is to take advantage of multiplex isobaric tag labeling strategy coupled with mass spectrometry to relatively quantify primary and secondary amine-containing metabolites in serum for the discovery of CRC-relevant biomarkers. Blood drawn from individuals carrying adenomas classified with different risk status and growth status were studied. Herein, customized 4-plex N,N-dimethylleucine (DiLeu) isobaric tags were used for labeling amine-containing metabolites. Labeled metabolites were measured using HPLC ESI-MS/MS. Relative quantification of metabolite standards and pooled serum samples (1:1:1:1 mixed after labeling) using DiLeu tags

showed less than 10% coefficient variance, indicating a satisfactory accuracy for quantification. Metabolite identification was based on accurate mass matching and/or retention time of labeled metabolite standards. We detected 40 metabolites across all the serum samples, including 18 metabolites validated with standards. We found that a down-regulation of threonine and cystine levels in the growing adenomas bearing patients. In contrast, elevated levels of kynurenine and ornithine were found. Threonine, cystine and kynurenine were not significantly related to the total adenoma volume, while ornithine showed significant association ($p < 0.05$) with the total volume of adenomas in Pearson's test. Postpolypectomy blood samples demonstrated that the abnormal levels of threonine, kynurenine and ornithine renormalized after polypectomy. While reduced cystine levels persisted after polypectomy. Furthermore, uniform manifold approximation and projection (UMAP) was used to visualize the metabolic of these samples, showing that normal control individuals and colonic adenoma bearing patients could be clustered into different groups. Based on these results, we concluded that metabolite biomarkers identified here could be used to distinguish colon cancer patients from normal individuals.

P3.11 | Leveraging FAIMS and dynamic retention time for multiplexed, high sensitivity PRM assay design

Camille Lombard-Banek[1] (camille.lombard-banek@astrazeneca.com), Zifeng Song[1], Steve Sweet[1], Yeoun Jin Kim[1]
[1]Astrazeneca, Gaithersburg, MD

In the clinical setting, a sensitive, multiplexed assay to monitor cancer-related protein targets is helpful to answer clinical trial questions. Sample amounts are typically limited, e.g., 5 mm^2 of tumor from a single 10-micron thick FFPE section of a core-needle biopsy. In this context increasing multiplexity while maintaining throughput (~45 min gradient) and sensitivity requires balance to maintain the quality of the data (e.g., points across the chromatographic peak). Our group has recently demonstrated the benefit of FAIMS in reducing interferences for PRM in complex matrices, however it was until now incompatible with dynamic retention time (RT). Here, we tested the benefits of dynamic RT in combination with FAIMS using our ~80 targets assay (~160 precursors) for increased multiplexing without compromising sensitivity and throughput. Our preliminary data showed that the dynamic RT helped in adjusting the target acquisition time window on the fly, decreasing the instances of targets being missed. We also demonstrated that thanks to the dynamic RT implementation, the acquisition window can be reduced from 2 min to 1 min, allowing more points per peaks (up to ~70%) or higher multiplexing (300 additional targets). In parallel, we also implemented in-sample calibration (ISC) with multiple isotopologues PRM (MIPRM) for our canonical cancer targets (HER2, EGFR, and ESR1). ISC-MIPRM is used to evaluate the change in sensitivity of our assay while evaluating different acquisition parameters involving dynamic RT: acquisition windows between 30 s and 2 min; resolution between 30,000 and 240,000; ion injection time between <<50 and >200 ms.
The implementation of an improved FAIMS dynamic RT PRM assay promises to increase the robustness and target number of our PRM assay. We also demonstrate the benefit of ISC-MIPRM to evaluate sensitivity of the assay on-the-fly following method changes.

P3.12 | Parallel Reaction Monitoring of Proximity Labeling Enables Sensitive and Quantitative Characterization of Stress-Dependent Protein Mistargeting

Ziqi Lyu[1] (zlyu002@ucr.edu), Joseph Genereux[1]
[1]University of California, Riverside, Riverside, CA

During ER stress, protein import into the endoplasmic reticulum (ER) is attenuated to protect the ER from misfolded protein accumulation. This process, termed ER pre-emptive quality control (ER pre-QC), has not been extensively characterized. We recently reported a peroxidase proximity labeling-based assay to identify mistrafficking of secretory proteins in living cells during ER pre-QC following ER stress. This assay depends on protein quantification by immunoblotting and densitometry, which is only semi-quantitative, laborious and suffers from poor sensitivity. Here, we integrate and optimize parallel reaction monitoring (PRM) mass spectrometry to enable a more sensitive and quantitative platform for ER import. The PRM-assayed peptide peak area correlates well with immunoblotting densitometry (Pearson's $R^2 > 0.99$) across distinct drug treatment conditions. Sensitivity by PRM is at least 5 times higher than near-IR immunoblotting. By assaying multiple peptides per protein, we avoid artifacts associated with epitope damage during labeling. We applied our PRM assay in HEK293T cells to evaluate mistargeting of the secretory protein transthyretin (TTR) in response to a panel of chemical ER stressors that act through distinct mechanisms. Surprisingly, only the stressors that deplete ER calcium through SERCA inhibition (thapsigargin and cyclopiazonic acid) induce TTR cytosolic accumulation through pre-QC. These results establish that proximity labeling in conjunction with PRM is effective for characterizing stress-dependent protein mistargeting, and suggest that ER stress alone is not sufficient to induce pre-QC.

P3.13 | Identification of differential biomarkers in subjects with TBI, with and without neurological deterioration from intracranial hematoma expansion

Heather S. McCoy[2] (heather.mccoy@pnnl.gov), Marina Gritsenko[2], Holly Hinson[1], Jon M. Jacobs[2]
[1]Oregon Health and Sciences University, Portland, OR, [2]Pacific Northwest National Laboratory, Richland, WA

Introduction: Traumatic brain injury (TBI) is a form of acquired brain injury, producing significant morbidity. Efforts to identify subtypes of TBI using fluid-based biomarkers has potential to improve outcome and survival. Preliminary data suggests a clear link between specific immunoregulatory biomarkers released after TBI and progression of intracranial hemorrhage (PICH). Herein we utilized mass spectrometry (MS) techniques to identify differential plasma protein expression in patients with TBI.
Methods: A representative subset (n=30) of baseline plasma samples of antecedent trauma were analyzed via global proteomics which included IgY14 depletion, TMT labeling, off-line HPLC fractionation, and liquid chromatography (LC)-MS/MS. Downstream data analysis included quantification for statistical comparisons between patient subgroups with PICH and intra-axial (IA)-PICH.
Results: Blood plasma obtained an average of 1.36 (±0.65) hours from the antecedent trauma showed 19 subjects (63%) with PICH, 15 (50%) IA-PICH, and 12 (40%) who deteriorated. The PICH subgroup showed significant differences in sex (14M, 5F), nominal differences in age (8 ≤ 50, 7 > 50). The IA-PICH subgroup demonstrated significant differences in age (10 ≤ 50, 5 >50) and sex (11M, 4F). A total 3,147 proteins were characterized in the cohort, generating 187 and 145 proteins significant to PICH and IA-PICH respectively ($p < 0.05$). Pathway analysis of the differential protein markers populated core biological processes for coagulation cascades and complement pathways. We also observed differential protein patterns based upon age, 492 proteins significant, resulting in an overwhelming mapping towards cell adhesion markers, implying a role of cellular responses in influencing TBI.
Conclusions: Global proteomic analysis by MS of plasma samples from TBI patients with and without PICH provided prospective markers of occurrence based upon core circulatory pathways. The mapping of differential proteins markers based upon age covariates provides an opportunity to investigate the mechanistic effect of age in the context of TBI.

P3.14 | Discovery and Quantification of Novel Celiac Disease-related Gluten Immunogenic Peptides and Their Deamidated Forms in Human Urine

Jie Pu[1] (jiepu@buffalo.edu), Qingqing Shen[2], Chao Xue[1], Timothy W. Sikorski[3], Bo An[3], Zhuo Chen[3], Thomas Angel[3], John Mehl[3], Huaping Tang[3], Jun Qu
[1]SUNY at Buffalo, Buffalo, NY, [2]University at Buffalo, Buffalo, NY, [3]GSK, Collegeville, PA

Celiac disease is caused by the ingestion of gluten proteins (*e.g.*, α2-gliadin) in certain populations. Recognition of gluten immunogenic peptides (GIP) by HLA-DQ2/DQ8 is strongly enhanced through deamidation of glutamine residues, causing severe immune responses. Selective identification and quantification of native and deamidated forms of GIPs is critical for therapeutic efforts but remains challenging owing to the very close m/z, the existence of many Q-residues in close proximity, and interference by multiple isotope peaks. To address these challenges, here we developed a suite of LC-MS based strategies which unambiguously identified and quantified various novel GIPs and their deamidated forms in urine samples. *First*, we developed an extensive nano-LC separation for GIP identification, achieved complete baseline separation among native, mono-, di-deamidated forms of GIPs. *Second*, urine samples from healthy subjects after gluten-rich diet were surveyed to identify GIPs by the developed methods. Multiple antibodies have been assessed for immunoprecipitation to improve identification rate while avoiding capture bias by any antibody. To ensure a reliable identification, especially the discovery of novel species, we used urine samples from volunteers with gluten-free diet as negative control to minimize false positive discovery. We discovered that the widely reported 33-mer (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF), which has been described as the main CD-immunodominant peptide, was not detected in any of the urine samples. Instead, we identified multiple novel species containing CD-patient-specific T-cell epitopes with high confidence. The discovery of 15 major forms of GIPs were confirmed using synthesized peptides. *Finally*, to enable high-throughput and robust clinical quantification of GIPs, we developed a MAX-based strategy for enrichment, and a weakly acidic-pH LC-FAIMS/SRM-MS method was employed to selectively quantifying multiple native vs. deamidated GIPs within 15 min/sample and with LOQs of 5-10 pg/mL. Using this multiplexed quantitative method, we measured time courses of GIPs in subjects after gluten-rich meals, and gluten-related GIP profiles were captured.

P3.15 | Antibody-free, Accurate quantification of Low-abundance Obesity Biomarkers by using a novel LC-FAIMS-dCV-SRM-MS Method

Qingqing Shen[1] (qshen2@buffalo.edu), Jie Pu[2], Chao Xue[2], Xiaoyu Zhu[2], Boeser Cornelia[3], Claudia Martins[3], Peterman Scott[3], Tai-Du Lin[4], Wei-Jun Qian[4], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY, [2]SUNY at Buffalo, Buffalo, NY, [3]Thermo Fisher Scientific, San Jose, CA, [4]Pacific Northwest National Laboratory, Richland, WA

The insufficient sensitivity in LC-MS-based targeted quantification of most circulating biomarkers, the majority of which are in the low-ng/mL range in plasma, severely impedes the application of LC-MS in clinical analysis. Although immunocapture-LC/MS could substantially improve sensitivity, accuracy/precision is heavily dependent on the quality of the antibody, and the development/selection of suitable antibody is often time-consuming and/or costly. Here we devised a novel antibody-free strategy for ultra-sensitive, accurate and robust quantification of low-abundance circulating biomarkers, which consists i)high-throughput removal of albumin/IgGs with spin cartridges and ii)a LC-FAIMS-dCV-SRM-MS method to substantially decrease the chemical noise and thereby greatly improving S/N of biomarkers. Plasma with the targets selectively depleted was employed for calibration/QC samples, which enabled accurate quantification. We applied the strategy in quantification of six key obesity-related biomarkers, including leptin, PAI-1, MMP2, MMP9, NGAL, and TIMP1. The targets in pooled human plasma were depleted to generate a target-specific blank matrix. Albumin and IgGs were removed from by Agilent HAS/IgG MARS spin cartridges and processed by the SEPOD strategy we described previously. A trapping-micro-LC system, coupled to a FAIMS-Pro interface and a TSQ Altis™ triple-quad MS were used for quantification. The absolute recoveries of the targets after albumin/IgG removal ranged from 50%-102%, with excellent reproducibility(inter-batch CV%=2-11.5%, N=5). With the optimized LC-FAIMS-dCV-SRM-MS approach, the validated LOQs for six proteins were 1.2-18 ng/mL in plasma without any immunocapture. Cohorts of clinical samples from healthy volunteers and obese patients were analyzed, which revealed significant difference in leptin and PAI-1 levels between the two groups (p value<<0.01). The strategy developed here permits high-throughput, ultra-sensitive and multiplexed quantification of low-abundance circulating biomarkers, with a straightforward/robust protocol, high accuracy/precision, and without the need of immunocapture. Therefore, we believe it represents a promising solution to the current dilemma faced by LC-MS-based biomarker quantification in the clinical domain.

P3.16 | Monitoring Therapeutic Monoclonal Antibodies in Human Serum using HRAM-MS for Clinical Research

Yvonne Song[1] (yvonne.song@thermofisher.com), Stephanie Samra[1]
[1]Thermo Fisher Scientific, San Jose, CA

Each year, more monoclonal antibodies (mAbs) are approved by regulatory agencies to treat a wide range of diseases with an expected market value of over $200 billion by 2026. In clinical testing, the presence of endogenous immunoglobulins with almost identical structures from patients' samples adds another challenge to the accurate quantitation of therapeutic mAbs. Accordingly, mass spectrometry has gained substantial popularity for therapeutic mAb monitoring in clinical laboratories due to its great versatility to detect both tryptic peptides and intact light and heavy chains quantitatively. Here we present the intact light chain quantitation approach for measuring concentrations of therapeutic mAbs in human serum using Orbitrap Exploris 240 MS for clinical research. One of the benefits of intact light chain quantitation is a simplified workflow with faster sample preparation compared to peptide quantitation. Protein L magnetic beads were implemented to purify mAbs containing kappa light chains. After IdeS digestion and reduction steps, three subunits were generated, and Orbitrap Exploris 240 MS was operated at a resolution of 120k with intact protein application, low-pressure mode to detect the mAb intact light chains. For the targeted protein quantitation, TraceFinder was used from data acquisition to processing/reporting.

The workflow was successfully optimized as the middle-up approach, only requiring a low-volume sample of 10 uL per analysis within < 1.5 hours of sample preparation. Also, on-beads sample treatment was incorporated, offering seamless sample treatment with minimal sample loss. The LOQs were determined to be between 1 to 5 ug/mL of the mAb concentration in serum with great linearity of R^2 > 0.99. Through a quick column reproducibility evaluation, reproducible data were generated over two different column lots showing a 0.2-minute shift with less than 20% peak area differences. This approach can potentially be applied to other modality antibodies or Fab fragments containing kappa light chains.

P3.17 | IS-PRM for NRF2 Signaling and Prognostic Biomarkers in HNSCC

Nathan Wamsley[1] (n.t.wamsley@wustl.edu), Petra Erdmann Gilmore[1], Emily Wilkerson Zarbock[1], Kyle LaPak[1], Travis Schrank[2], Robert Sprung[1], Sophie Gerndt[1], Jose Zevallos[1], Reid Townsend[1], Li Guan[3], Dennis Goldfarb[1], Michael B. Major[1], Quynh-Thu Le[1]
[1]Washington University in St. Louis, St. Louis, [2]University of North Carolina School of Medicine, Oxford, NC, [3]Stanford University, Stanford, CA

Mass spectrometry-based (MS) targeted proteomics has enabled selective and reproducible protein quantitation suitable for clinical assays. This work presents a targeted data acquisition algorithm, which compared to a commercial implementation, SureQuant™, reduced median cycle times and improved quantitative robustness for a panel of over seven-hundred kinase-derived peptides. We applied the method to study pathways and proteins of clinical significance in head and neck cancer. MS revealed a handful of proteins sufficient to quantify NRF2 activation by analysis of a cancer cell line cohort and genetically engineered models. Then, in a cohort of thirty formalin-fixed paraffin-embedded oropharyngeal squamous cell carcinomas, NRF2 signaling intensity varied over a wide range and correlated with SOX2 expression. Moreover, programmed cell death 1 ligand 2 and markers for T-cell infiltration correlated positively with one another and with HPV infection status. MS quantitation of p16 protein discriminated between HPV-positive and negative cases, and for five tumors, detection of the oncogenic viral protein, E7, confirmed translationally active virus. This work validates a protein biomarker assay designed for a population with unmet clinical need.

P3.18 | Investigation of systemic inflammation induced FcRn dysregulation and related mAb disposition, with sensitive, multiplex trapping-µLC-FAIMS-dCV-MRM

Chao Xue[1] (chaoxue@buffalo.edu), Qingqing Shen[1], Jie Pu[1], Ming Zhang[1], Yang Liu[1], Jun Qu[1]
[1]University at buffalo, Buffalo, NY

By mediating the recycling and transcytosis of IgG, FcRn plays a highly critical role in maintaining the homeostasis and prolonging the circulation half-life of IgG. Recent studies suggested that FcRn levels may be significantly elevated in the presence of inflammation. Therefore, it is important to extensively investigate the inflammation-induced FcRn dysregulation and its effect on IgG disposition, which requires accurate, absolute quantification of FcRn in tissue. However, accurate, absolute quantification of FcRn is challenging because of the poor extraction recovery of FcRn, a transmembrane, hydrophobic protein with relative low abundance in tissues. To address these issues, we developed and validated a protocol consisting highly efficient/reproducible recovery of FcRn from tissues, and a trapping-µLC-FAIMS-MRM strategy, which enabled simultaneous, ultra-sensitive, accurate and robust quantification of FcRn and IgG in one analysis. We optimized a previously developed SEPOD procedure, including the use a detergent cocktail for extraction/denaturation and 2-step sonication, which exhaustively recovered FcRn and its signature peptides from tissues. By the FAIMS-dCV strategy we developed, substantial S/N improvement was achieved, with LOQs of 200 ng/g in tissue for both FcRn and IgG. To our knowledge, this is the first method that can directly quantify FcRn.

We successfully applied the strategy in analysis of LPS-induced systemic inflammation model, and the time courses of the absolute levels of FcRn in multiple tissues were accurately quantified. We also conducted a full IgG PK study in systemic-inflammation vs. healthy control groups and observed a 1.5-fold decrease of IgG half-life in inflammation group. Besides, elevated tissue to plasma ratio (Kp value) in multiple tissues was found at later time points during the course of treatment. In summary, the quantified FcRn expression levels at different inflammatory conditions provides novel insights into the inflammation-induced FcRn dysregulation, which provide highly valuable information for mAb PK under inflammation conditions.

P3.20 | Extensive Investigation on Antibody-Drug Conjugates Induced Ocular Toxicity by Integrating Global Proteomics and Targeted Drug Disposition Analysis

Xiaoyu Zhu[1] (xzhu27@buffalo.edu), Min Ma[1], Ming Zhang[1], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY

ADCs-induced ocular toxicities have been widely reported in clinical investigations. A clear payload association with ocular toxicity has been observed. Typically, ocular toxicities were induced by ADCs with DM4 or MMAF payloads but rarely described for ADCs with MMAE, regardless of the antibody. The underlying mechanisms are largely unknown. To provide insights into ADCs-induced ocular toxicities, a comprehensive investigation of ADC ocular tissue disposition and elucidation of the toxicity-related mechanisms are highly desirable to direct therapeutic efforts. Here we developed a novel pipelined LC/MS-based strategy combining ultra-sensitive drug disposition analysis and global proteomics using a unique IonStar quantitative proteomics strategy, which allows accurate ADC disposition characterization and comprehensive toxic mechanism exploration in an animal model that reproducing clinically-observed ADC-induced ocular toxicities. Following intravenous administration of clinically-equivalent doses of

ADCs with DM4 or MMAE payload, ocular toxicities were only observed in the cornea and lens of DM4-ADC group. Ocular tissues were collected for drug disposition and proteomics analysis using the novel LC/MS-based pipeline. We found the two ADCs showed distinct dispositions in both cornea and lens: the free toxin accumulation of DM4-ADC is higher than that of MMAE-ADC. In the same set of samples, using IonStar quantitative proteomics strategy, we quantified 4,372 proteins in cornea and 949 proteins in lens with low missing-data. The proteomics data revealed:1)DM4-ADC-associated corneal toxicity was related to activation of inflammatory response, disruption of corneal cell survival, proliferation and migration, and disrupted corneal function; 2)DM4-ADC-associated lens toxicity was related to disturbance of lens transparency maintenance, disruption of microtubule organization and induction of cell apoptosis; 3)By comparison, MMAE-ADC-related biological processes involved in maintenance of regular corneal and lens functions. Finally, a set of potential ocular toxicity-related biomarkers were identified. In conclusion, in a clinically-relevant animal model and using a novel LC/MS strategy, we found the ocular tissue disposition difference between two ADCs with or without ocular toxicity, and elucidated the possible toxicity-related mechanisms.

P3.22 | Systematic Study of DIA and DDA for Protein Identification and Label-Free Quantitation

Charlie Yang[1] (charlie.yang@rosalindfranklin.edu), Keith D. Philibert[1], Xiaomeng Shao[1], Marc Glucksman[1]
[1]*Rosalind Franklin University (RFUMS), North Chicago, IL*

Data-dependent acquisition (DDA) has been the longtime standard for protein identification/quantification in proteomics. However, Data Independent Acquisition (DIA) is a reproducible, precise analytic tool, particularly for large sample sets. Herein, we focus on popular Orbitrap-based mass spectrometers comparing DIA and DDA based on topN for serial MS2. We systematically examined scan rates (normal/rapid/enhanced/turbo scan), protein identification and coverage, peptide mapping, and label-free quantitation with HeLa cell digests in triplicate. Optimizing DIA settings and 30'/60'/120' LC-MS/MS yielded 5728, 6184, and 6470 proteins were identified respectively, from 500ng of HeLa digest. This compares to DDA at 30'/60'/120', identifying 1792, 2703, and 3142 proteins, respectively. ~3x more proteins were identified using DIA. Data were filtered in terms of -10lgP, FDR, unique peptides, and de novo score by PEAKS X PRO and Spectronaut software. Proteome sequence coverage (FDR <1%) was significantly improved by use of rapid scans compared with normal/rapid/enhanced/turbo scans. The number of protein identifications using HeLa cell digests with rapid scan mode were compared to normal/rapid/enhanced/turbo scan and increased 7%, 16%, and 25%, respectively and identified peptides increased 8%, 22%, 36% utilizing the same scan paradigm. Label-free quantification increased accuracy by decreasing errors: 9%, 36%, and 43% comparing rapid scan to normal/rapid/enhanced/turbo scans. All settings included: dynamic exclusion, signal thresholds triggering MS/MS, mass resolution, ion injection time, MS/MS number, monoisotopic precursor selection, FTMS preview mode, normalized collision energy, AGC target value, microscan number, LC conditions and bioinformatics analyses; all controlled herein.

Utilizing DIA or DDA, and scan mode significantly affects protein, peptide identification rate, identified coverage, and label-free quantitation. Herein, we compared DIA and DDA in proteinID LFQ with different scan rates, resolutions (15-240K) and collision modes (CID/HCD/ETD) towards protein identification and label free quantitation with HeLa digests. DIA not only surpasses DDA ~3-fold but is more reproducible and precise.

P4: Cancer Proteomics

P4.01 | MS proteomics of tumor interstitial fluid extravascular vesicles is a new class of biomarkers for immunotherapy monitoring

Dlal Baljoon[1] (dlal.baljoon@gmail.com), Marissa A. Howard[1], Lance Liotta[1], Weidong Zhou[1], Alessandra Luchini[1]
[1]George Mason University, Manassas, VA

Extracellular Vesicles (EVs) resident within Interstitial fluid (IF) of the tumor tissue in vivo microenvironment plays an essential role in cell-cell communication processes driving tumor progression. Tumor IF EVs are a major source of information that drains to the local lymph node for immune inspection. The previously unknown EV proteomic biomarker information contained within the interstitial space is a new class of information reflecting the tissue's physiologic and disease state. Understanding the EV constituents surrounding or bathing a tumor could provide insights into mechanisms that aid tumor immune escape and metastasis. We have developed a method to isolate the IF from an excised tumor without damaging the underlying tissue cellular morphology. We aim to better understand the global proteomic constituents within the tumor IF EVs compares to the tumor tissue parenchyma using untargeted mass spectrometry (MS). First, we isolated the IF from untreated and drug-treated (anti-PD-L1) tumors. We verified the IF EVs were from the tumor by using GFP labelled tumor cells which transmit their GFP tag into EVs found within the tumor IF. The GFP expression CV between tumors was less than 10%, therefore this isolation method is highly reproducible. Next, the IF and tissue were prepared for MS analysis. We used the following bioinformatic tools during our post-analysis: Proteome Discoverer, InteractiVenn, and WebGestalt pathway overexpression. The tumor IF EVs are rich source of proteins with over 3000 unique markers returning per tumor sample. Moreover, for the drug treated tumors the overexpressed IF EV proteomics revealed high levels of neutrophil degranulation and T-cell activation indicative of the drug treatment's effect on the tumor. In conclusion, we have developed an MS analytical workflow which can eavesdrop on the tumor pre or post-treatment state using IF EVs as a completely novel form of in-vivo tissue biomarker.

P4.02 | Data-Independent Acquisition and Quantification for the Stratification of Breast Cancer Subtypes from Human FFPE Tissues

Jordan Burton[1] (jburton@buckinstitute.org), Rosemary Bai[2], Joanna Bons[1], Deng Pan[2], Christie Hunter[3], Thea D. Tlsty[2], Birgit Schilling[1], Philippe Gascarde[2]
[1]Buck Institute for Research on Aging, Novato, CA, [2]University of California San Francisco, San Francisco, CA, [3]SCIEX, Redwood City, CA

Breast cancer is a highly complex and heterogenous disease with multiple subtypes of breast cancer that influence patient outcomes and direct strategies for treatment and therapies. It is instrumental to comprehend the 'molecular landscape' of human breast cancer subtypes by determining what dynamic changes occur within the tumor in comparison to disease-free tissue to improve patient outcomes by developing new biomarkers and potential therapies. We demonstrate the use of a robust proteomic pipeline to gain insights into molecular tissue remodeling during breast cancer, analyzing human formalin-fixed, paraffin-embedded (FFPE) breast tissue. A novel, high-resolution ZenoTOF 7600 mass spectrometer with microflow chromatography is used to acquire proteomic data from FFPE human tissues with a label-free quantification method – Data-Independent Acquisition (DIA). Spectronaut 16 was used for protein group identification and stringent statistical processing of DIA data to determine molecular changes that are occurring in ER+PR+HER2- (ERPR, n=7, ER-PR-HER2+ (HER2, n=7), Triple Negative (TN, n=7), Triple Positive (TP, n=7), and Metaplastic (Meta, n=7) primary breast tumor tissues compared to disease-free tissues (DF = control, n=7) to ultimately discover novel biomarkers. Over 5,700 protein groups were identified and quantified with ? 2 unique peptides, of which over 35% were annotated with the gene-ontology term "Extracellular". Molecular changes confirm dramatic tissue remodeling using very strict statistically significance thresholds (|log2(FC)| ? 0.58 & q-value ? 0.001). Overall, over 1,800 significantly altered unique protein groups were determined in a comparison between ERPR vs. control, HER2 vs. control, and TN vs. control. Over 1,500 significantly altered protein groups were quantified in a comparison between TP vs. control, and over 1,600 significantly altered protein groups were quantified between Meta vs. control. We are using the significantly altered protein groups from each breast cancer subtype to find biomarkers conserved between breast cancer subtypes and biomarkers specific to each breast cancer

P4.03 | Role of Olink Explore 3072 in integrative multi-omics analysis in infectious diseases

Joe Caruso (joseph.caruso@olink.com), Olink, Detroit, MI

The clinical outcome and disease severity in infectious diseases are heterogeneous, and the progression or fatality of the disease cannot be explained by a single factor such as age, ethnicity, or co-morbidities. Using network-based system biology methodology,

we aim to stratify patient groups based solely on omics data signatures and to identify the mechanism associated with disease severity in the acute phase of infection and dysregulated immune function and its impact on aging in chronic infectious diseases at the personalized and group level. We used Similarity Network Fusion (SNF) by integrating transcriptomics, Olink Proteomics, and metabolomics data to stratify cohorts of COVID-19 patients as well as patients that had successfully been treated for HIV-1. Weighted co-expression network analysis was conducted to identify severity-specific gene co-expression and other associated mechanisms. We generated a context-specific genome-scale metabolic model for all individuals and patient groups, and performed flux balance analysis to characterize metabolic reactions altered upon the disease state. Topology analysis of the metabolic network was performed to investigate metabolic disruptions at the metabolite level and link this to infection-associated proteins identified by proteomics analysis to confirm their use as biomarkers for disease severity.

Our results demonstrate how integrative omics provide better patient clusters and disease classification. Furthermore, Olink® Explore 3072 provides biomarkers of disease severity that can easily be translated to define the targeted panel.

P4.04 | Turnover and thermal stability profiling in isogeneic lung cells of chromosome 3 aneuploidy identify heterogenous proteome buffering mechanisms

Yi Di[1] (yi.di@yale.edu), Wenxue Li[1], Joan J. Castellano[2], Qian Ba[1], Alison M. Taylor[2], Yansheng Liu[1]
[1]Yale Cancer Biology Institute, West Haven, CT, [2]Columbia University Irving Medical Center, New York, NY

Aneuploidy, large-scale structural chromosomal alteration, is a near-universal feature in cancers. Chromosome 3p arm deletion (3p loss) occurs in 78% of lung squamous cell carcinomas (SCC), and amplification of the 3q arm (3q gain) is in the next most frequent aneuploidy event in lung cancer. Protein turnover control was discovered to be crucial in compensating gene dosage effects in chromosomal “gain-type” aneuploidy. However, it is not clear whether protein turnover is universally regulated for all aneuploidy types. Herein, we addressed this fundamental question using an isogenic model of immortalized lung epithelial cells discordant for chromosome 3 aneuploidy. Firstly, based on high-performance DIA-MS, 9,518 ± 24 proteins were identified and measured among “3p loss”, “3q gain”, and wild-type cells, in which 8,800 proteins were also profiled with RNA expression values. As expected, we found that protein dosage compensation is particularly strong for protein complex subunits. Next, we combined pulsed SILAC with DIA-MS which successfully quantified protein turnover rates for 8,470 proteins. Interestingly, protein turnover regulation was confirmed for protein complex subunits in chromosome “3q gain” genotype (71 proteins, p-value < 0.0001), but not in “3p loss” cells, indicating protein turnover control are not universal for all aneuploid cells. To determine alternative proteomic buffering mechanisms, we utilized a proteome-wide thermal proteome profiling (TPP) assay. TPP-DIA determined melting temperature (Tm) for 4,424 ± 366 proteins with high reproducibility. Intriguingly, Tm of protein complex subunits encoded by 3p in the “3p loss” genotype increased significantly (p-value < 0.01), uncovering a novel mechanism of protein dosage adaptation against “loss-type” aneuploidy. Bioinformatic analysis of TPP-DIA indicated perturbed protein-protein interactions in “3p loss” scenario. We further validated our findings using biochemical approaches. Overall, our data demonstrates how “gain-type” and “loss-type” aneuploidy differentially impact protein abundance, turnover, and structural thermal stability, providing implications for future cancer therapeutics.

P4.05 | Understanding Phosphoproteome of Colorectal Cancer Spheroids after Inhibition of Fatty Acid Synthase

Brian D. Fries[1] (fries.94@osu.edu), Amanda Hummon[1]
[1]The Ohio State University, Columbus, OH

Cancer cells will undergo *de novo* synthesis of lipids to maintain their rapid growth and proliferation rather than obtain them exogenously. Fatty acid synthase (FASN) is responsible for synthesizing these lipids and has been shown to be upregulated in several human cancers. There has been keen interest in inhibiting FASN to curb the growth and proliferation of tumors. The first generation of FASN inhibitors stemmed from the discovery of cerulenin in 1994 from the fungus *Cephalosporium caerulens*. The popularity of inhibiting FASN waned after it was observed that cerulenin and its derivatives caused dramatic weight loss in mice. TVB-2640, a second-generation FASN inhibitor synthesized by Sagimet Biosciences, has gained immense popularity as being highly specific for the inhibition of FASN in human patients and has entered a phase 1 clinical trial for the treatment of solid colon tumors. Though it is known that FASN inhibition can alter the growth and proliferation of tumors, the distinct molecular differences that occur upon inhibition of FASN with a generation one or two inhibitor have yet to be determined. Here, we demonstrate the differences amongst the proteome and phosphoproteome of generation one and two FASN inhibitors within HCT 116 and HT-29 colon cancer spheroids. Proteomics and phosphoproteomics data were collected using data-independent acquisition (DIA). Changes were observed across all conditions within both the proteomes and phosphoproteomes, indicating that inhibition of FASN causes changes in cellular signaling events. Heat shock protein 105 was amongst the most upregulated proteins within HCT 116 and HT-29 spheroids treated with cerulenin, but not with TVB-2640 treatments. The proline/alanine-rich protein kinase STK39 was

found to be downregulated within HT-29 spheroids treated with TVB-2640. The results obtained may indicate that TVB-2640 being more specific for FASN may also have lower off-target effects, resulting in a more subtle inhibition of FASN.

P4.06 | Elucidation of Src Family Kinase Substrates via a Streamlined Endogenous Library Preparation Method

Samson J. Joseph[1] (josep462@umn.edu), Laurie Parker[1]
[1]University of Minnesota, Minneapolis, MN

Kinases represent a major target for drug development, due to their implication in many cancers and autoimmune diseases. However, only 25 kinases have a total of 68 different FDA approved kinase inhibitors, in a kinome containing over 500 kinases. These understudied “gap” kinases are potential pharmaceutical targets, and with suitable assays, kinase inhibitor development can be facilitated. The design of these assays is the focus of the Parker Lab. The assay development workflow involves the KALIP protocol (Kinase Assay LInked with Phosphoproteomics), where substrates are acquired by lysing cells, separating out proteins, digesting them, and treating them with phosphatase in order to produce a phosphate-free, endogenous peptide library, which is used in kinase assays in order to identify substrates. The substrate data is fed through KINATEST-ID, an *in silico* workflow for scoring substrate preference and specificity, which is used to develop synthetic reporter peptides. Unfortunately, KALIP has a low throughput. It requires multiple rounds of solvent removal and desalting with Sep-Pak C18 columns, which have a limited binding capacity and are time and user-intensive to use. The new workflow utilizes Protifi S-Trap™ columns, which have a much higher throughput (nearly 20 times that of the C18 columns). The S-Trap™ technology significantly reduces sample processing time, and allows for enzymatic treatment in a one-pot telescopic reaction process. This protocol is currently being used to prepare substrate libraries to study the Src family kinases, but is applicable to many more kinases. The reduced time required and higher throughput of the S-Trap™ enables a single library batch to be used across multiple kinases, reducing variation, and increasing the practicality of kinase substrate profiling through KALIP. As more substrate information for gap kinases is filled in, KINATEST-ID will possess higher specificity and accuracy, allowing for the development of better reporter peptides and inhibitors.

P4.07 | Proteomics Interrogation of an Uncharacterized Zinc Finger Transcription Factor Associated with Lung Adenocarcinoma (LUAD)

Xingyu Liu[1] (xingyu@wustl.edu), Yixuan (Axe) Xie[1], Zongtao Lin[1], Benjamin A. Garcia[1]
[1]Washington University School of Medicine in St. Louis, Saint Louis, MO

Zinc fingers are small protein structural domains that bind and are stabilized by zinc ions, and frequently appear in proteins as repeating domains and bind DNA in a sequence specific manner. Many zinc finger repeat containing proteins have been found to regulate gene expression patterns and comprise a large family of transcription factors (ZnTFs). Some of these ZnTFs have been associated with various types of human diseases including cancers. However, due to the large number of known members, many ZnTFs are yet to be further characterized. One vastly understudied ZnTF, ZNF687, was identified as potential marker for lung adenocarcinoma (LUAD) using a bioinformatic analysis of ‘The Cancer Genome Atlas’(TCGA) database. To date, no further investigations into the role of ZNF687 in LUAD have been reported. Here, we aim to characterize how ZNF687 regulates transcription as well as influences epigenetic mechanisms to promote LUAD. We interrogated the function of ZNF687 using different liquid chromatography–mass spectrometry (LC/MS) based proteomics approaches to interrogate proteome expression, protein-protein interactions and histone modifications (epi-proteomics) in LUAD cancer cell lines. The resulting proteomics and epi-proteomics data can be combined with genomics data such as RNA-seq and ChIP-seq to provide a comprehensive view of the functions of ZNF687 in LUAD cells. Ultimately, the outcome from this study will provide valuable information to further develop new diagnostic methods and therapies for LUAD.

P4.08 | Chemotherapy-Treated Organoid Models of Ovarian Cancer Drive Alterations in Protein Abundances and Interactions

Krystal K. Lum[1] (klum@princeton.edu), Oscar Pundel[2], Bokai Song[1], John Muroski[1], Benjamin Neel[2], Ileana Cristea[1]
[1]Princeton University, Princeton, NJ, [2]NYU Langone Health, New York, NY

High-grade serous tubo-ovarian cancer (HGSC) is the most prevalent and lethal type of epithelial ovarian cancer. While HGSC patients initially respond to standard of care chemotherapies, drug resistance and cancer recurrence is pervasive. HGSC tumors are comprised of genotypically distinct mutations in oncogenes and tumor suppressor genes, and studying HGSC pathogenesis has been hindered by a lack of accurate model systems that recapitulate key HGSC features. Recently, human genomics-informed organoid models of HGSC were developed that address this gap, providing a new opportunity to determine how genotype-defined

tumors remodel cell-autonomous signaling pathways during chemotherapy treatments. NF1-deficient HGSC organoids revealed a significant sensitivity to the widely used, FDA-approved anti-cancer drug, Paclitaxel, when compared to other HGSC genotypic organoids. However, mechanisms underlying this hypersensitivity are unknown. Understanding mechanisms governing chemotherapy sensitivity are crucial to the rational design of more effective combination therapies. Here, we integrate proteome analyses with targeted mass spectrometry (MS) and thermal proximity co-aggregation (TPCA)-MS to define protein changes across multiple genotype-defined organoid models of HGSC upon Paclitaxel treatment. These include mutational combinations representing the most frequent and aggressive HGSC subtypes, such as *Trp53*$^{-/-}$;*Brca1*$^{-/-}$;*Myc*^{OE} and *Trp53*$^{-/-}$;*Pten*$^{-/-}$;*Nf1*$^{-/-}$. We compare three-dimensional organoids to two-dimensional monolayers, identifying altered metabolic and cytoskeletal proteins. These observations prompted us to perform targeted MS based on parallel reaction monitoring, which uncovered protein abundance changes in organelle-organelle contacts and energy metabolism, as well as microtubule dynamics. Using TPCA-MS, we further capture system-wide dynamics of protein complexes upon Paclitaxel treatment. Follow-up analyses using live-cell, super-resolution microscopy and functional biomolecular assays identify differences in mitochondrial bioenergetics and Paclitaxel-induced microtubule stabilization. Altogether, our findings reveal altered protein networks of drug targets and signatures in therapeutic response within ovarian cancer organoids. Our results have broad implications in establishing key features of ovarian cancer pathogenesis and organelle structure-function relationships critical for tumorigenesis.

P4.09 | Persistence of Targetable Lesions, Therapy Sensitivity and Proteomes Through Disease Evolution in Pediatric Acute Lymphoblastic Leukemia

Amanda C. Lorentzian[1,2] (amanda.lorentzian@bcchr.ca), Jenna Rever[1,2], Enes K. Ergin[2,3], Meiyun Guo[1,2], Neha M. Akella[1,2], Nina Rolf[1,2], C. James Lim[1,2], Gregor S.D. Reid[1,2], Christopher A. Maxwell[1,2,4], Philipp F. Lange[2,3,4]

[1]Department of Pediatrics, University of British Columbia, Vancouver, Canada, [2]Michael Cuccione Childhood Cancer Research Program at the BC Children's Hospital Research Institute, Vancouver, Canada, [3]Department of Pathology, University of British Columbia, Vancouver, Canada

Precision oncology programs have the potential to improve outcomes for relapse or high-risk pediatric cancers. However, it remains challenging to identify and test personalized treatment protocols for each child in a timeframe that enables effective clinical intervention. Large-scale genomic analyses have recently indicated that relapses often arise from sub clonal outgrowths. However, the impact of clonal evolution on the actionable proteome and response to targeted therapy is not known. We present a comprehensive retrospective analysis of paired ALL diagnosis and relapsed specimens to reveal how variants change or persist through disease progression. To first investigate persistence of pediatric cancer driving mutations, we performed targeted next generation sequencing and observed 9 of 11 paired samples having at least 50% retention of mutations. Selected samples were then treated with variant-targeted therapies, and although the inhibitors showed poor selectivity, paired samples showed high correlation of drug response. To next determine if proteomes are similarly stable, we employed a comprehensive data independent acquisition (DIA) analysis to robustly quantify 8153 unique proteins and 146 cancer associated proteins (CAPs). We observed stable abundance of CAPs and stable proteomes through disease progression, especially in cancer related processes. Finally, discovery-driven whole proteome analysis prioritized PARP1 as a new pan-ALL target candidate which was confirmed by robust sensitivity of ALL primary cells to treatment with PARP1/2 inhibitors. Together our findings suggest that initiation of a precision medicine workflow when pediatric ALL is first diagnosed has the potential to identify disease sensitivities that are likely to persist at disease relapse.

P4.10 | Multi-omics and Imaging Approaches Reveal New Insights in Combined Small Cell Lung Cancer

Michael B. Major (bmajor@wustl.edu), Washington University, St. Louis, MO

Small Cell Lung Cancer (SCLC) accounts for 15% of all lung cancers and is an aggressive neuroendocrine carcinoma with a 5-year survival rate of 7%. SCLCs are classified histologically as pure-SCLC (P-SCLC) or combined-SCLC (C-SCLC). More than 85% of P-SCLC lack P53 and RB1 and present with uniform histology. In contrast, C-SCLC comprise 15% of all SCLC and are characterized as mixed histology of P-SCLC with lung adenocarcinoma, lung squamous cell carcinoma or large cell neuroendocrine carcinoma (LCNEC). The molecular mechanisms driving C-SCLC evolution and histological plasticity remain unclear, in part due to a lack of experimental models. We developed the first GEMM that resembles the histology and genetics of human C-SCLC/LCNEC. In mice carrying an activating NRF2 mutation with loss of TP53 and P16, we observed small P-SCLC lesions and large invasive C-SCLC with mixed histology of SCLC and LCNEC. Here we used H&E staining and the Bruker timsTOF fleX MALDI-2 system to spatially characterize the lipidome and metabolome of C-SCLC arising in the TP53;P16;NRF2active GEMM. Over 300 putative identifications of metabolites and lipids were visualized, including PE 36:1 which showed high abundance in tumor regions while PC 32:0 was predominately absent from tumor regions and seen in high abundance in the normal adjacent regions. TIMS separation provided resolution for many isobaric species that were not separated by mass spectrometry alone. To study the proteome, we used the Bruker timsTOF SCP system to study liquid capture microdissected peptide from tumor cells, large airways and normal adjacent lung tissue. With a DDA-educated DIA library and PASEF, we observed 6,747 protein groups and 81,811 precursors from 20ng of

peptide in a 1 hour gradient. Collectively, our data nominate new pathways involved in the histological plasticity of C-SCLC and strengthen the use of spatially-aware and ultrasensitive MS platforms for cancer study.

P4.11 | Proteomic Signatures from Fresh Surgical Resections and Formalin Fixed Paraffin Embedded Tissues from Lung Squamous Cell Carcinoma and Adenocarcinoma

Jacob Rose[1] (Jacob.rose714@gmail.com), Joanna Bons[1], Deng Pan[3], Samah Shah[1], Rosemary Bai[3], Chira Chen-Tanyolac[5], Xianhong Wang[5], Nathan Basisty[4], Veena Sangwan[2], Jordan Burton[1], Philippe Gascarde[3], Lorenzo Ferri[2], Thea D. Tlsty[3], Birgit Schilling[1]
[4]Buck Institute for Research on Aging, Novato, CA, [2]McGill University, Montreal, [3]University of California San Francisco, San Francisco, CA, [4]NIA, Baltimore, MD, [5]UCSF Parnassus, San Francisco, CA

Lung squamous cell carcinoma (LSCC) is categorized as a chronic inflammation-associated cancer (CIAC). Here, we present a combined approach of analyzing fresh-frozen and formalin fixed paraffin embedded (FFPE) preserved samples, leveraging a widely available resource to gain greater insights into CIAC as well as lung adenocarcinoma (non-CIAC) cancers. We investigated 21 fresh surgical resections of LSCC patients and their matched normal tissues, followed by an enrichment of the extracellular matrix (ECM), ECM protein tryptic digestion, and subsequent peptide deglycosylation. In parallel, a separate comparison of widely available FFPE samples, including 5 LSCC, 5 lung adenocarcinoma, and 5 normal lung tissues, was performed. The latter samples were deparaffinized, were not enriched for ECM, but were otherwise subjected to similar sample preparation for digestion as described above. All samples were analyzed by mass spectrometry on a TripleTOF 6600 platform (SCIEX) operated in data-independent acquisition (DIA) mode. DIA data were processed with Spectronaut (Biognosys) using a pan-human spectral library. These acquisitions resulted in 3,249 identified protein groups, with ≥2 unique peptides in the ECM enriched, fresh samples and 1,982 identified proteins groups from the FFPE cohort. We present an ECM specific signature of CIAC lung cancer that are also robustly conserved in FFPE samples including an increase in Tenascin, Periostin, and SerpinH1. Interestingly, some of the most dramatic ECM alterations that appear to be specific to CIAC cancers did not reach significance in the non-CIAC samples, identifying CIAC-specific alterations. Additionally, FFPE samples provided access to integral intracellular protein changes. Interrogating FFPE as a whole tissue uncovered catalase as a major change in both CIAC and Non-CIAC FFPE cohorts, which was reported to be a protective agent in cancer progression. It is encouraging that widely accessible and rigorously characterized FFPE samples still maintain our discovered ECM signatures changing with cancer, even without ECM enrichment.

P4.12 | Spatial Multi-Omic Profiling of ALK-Rearranged Lung Cancer

Beata Szeitz[1] (szeitz.beata@phd.semmelweis.hu), Tibor Glasz[1], Gabor Toth[2], Mirjam Balbisi[2], Janos Fillinger[3], Szabolcs Horvath[3], Reka Mohacsi[1], HoJeong Kwon[4], Judit Moldvay[3], Lilla Turiak[2], Attila Marcell Szasz[1]
[1]Semmelweis University, Budapest, Hungary, [2]Research Centre for Natural Sciences, Budapest, Hungary, [3]National Koranyi Institute of Pulmonology, Budapest, Hungary, [4]Yonsei University, Seoul, South Korea

ALK rearrangements are the oncogenic drivers of merely 3-6% of lung adenocarcinoma cases. Due to its rarity, investigations of both inter- and intratumoral heterogeneity are needed to better understand the molecular characteristics of this distinct lung cancer type. We examined 7 treatment-naïve, primary lung cancers with ALK-rearrangement using GeoMX digital spatial profiling and label-free shotgun proteomics. On each FFPE tumor slide, 12 smaller transcriptomic and 2-6 larger proteomic regions (containing 3.1 RNA regions on average) were analyzed. Both tumor regions and adjacent tissue were selected. The tumor regions were further annotated for various histopathological characteristics, including lymphocytic infiltration and morphology (solid, tubular, papillary). A total of 1,812 genes and 2,318 protein groups were quantified across the regions, and only 306 genes were shared between the two molecular layers. Transforming the data to single-sample gene set scores enabled a direct comparison between the proteome and transcriptome, demonstrating commonly dysregulated pathways in tumors compared to adjacent tissue. On the other hand, pathway activities derived from the transcriptome changed more prominently across high vs low immune infiltrated regions. The extent of intratumoral heterogeneity varied substantially across the 7 examined cases, both at molecular and pathway-level. Together, the findings underscore the importance of multi-omic approaches in mapping the spatial heterogeneity of ALK-rearranged lung cancers, which we plan to foster in the future. Supported by grants KNN121510, OTKA FK 131603, EFOP-3.6.3-VEKOP-16-2017-00009, and ÚNKP-22-3-II.

P4.13 | Proteomic Phenotyping of a Metabolically Defined Putative Mammary Stem Cell Population

Matthew Waas (waasmatthew@gmail.com), Princess Margaret Cancer Centre, Toronto

Breast cancer, a heterogeneous disease of multiple origins, is the most common cancer in women. The deadliest breast cancers are

thought to arise from stem/progenitors in the mammary epithelium. 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-based digestion using commercially available consumables and, in combination with optimization of sample preparation, instrument acquisition, and data analysis, we reproducibly obtain >3,000 proteins from 500 sorted cells. We have performed a repeatability study to characterize the 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 profiled the proteomic phenotype of the metabolically defined putative epithelial stem/progenitor cells from six individual mice (~2000 cells/mouse). We detect 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 proteome 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.

P4.14 | Proteome Analysis Reveals a Mitochondrial Aconitase (ACO2)-Iron Axis Driving Aggressive Lung Cancers

Leanne Wybenga-Groot[1] (leanne.wybenga-groot@sickkids.ca), Shideh Mirhadi[1], Wen Zhang[1], Ming-Sound Tsao[2], Michael Moran[1]
[1]The Hospital for Sick Children, Toronto, ON, [2]University Health Network, Toronto

The ability of a patient tumour to engraft an immunodeficient mouse is the strongest known independent indicator of poor prognosis in early-stage non-small cell lung cancer (NSCLC). Analysis of primary NSCLC proteomes revealed low-level expression of mitochondrial aconitase (ACO2) in the more aggressive, engrafting tumours. Knockdown of ACO2 protein expression transformed immortalized lung epithelial cells, whereas up-regulation of ACO2 in transformed NSCLC cells inhibited cell proliferation *in vitro* and tumour growth *in vivo*. High level ACO2 increased iron response element binding protein 1 (IRP1) and the intracellular labile iron pool. Impaired cellular proliferation associated with high level ACO2 was reversed by treatment of cells with an iron chelator, whereas increased cell proliferation associated with low level ACO2 was suppressed by treatment of cells with iron. Expression of CDGSH iron-sulfur domain-containing protein 1 (CISD1; also known as mitoNEET) was modulated by ACO2 expression level and inhibition of mitoNEET by RNA interference or by treatment of cells with pioglitazone also increased iron and cell death. In conclusion: (1) ACO2 is identified as a regulator of iron homeostasis; (2) mitoNEET is implicated as a target in aggressive NSCLC; and (3) Iron-sulfur-cluster-associated proteins including mitochondrial aconitase ACO2, mitoNEET, and iron response element binding protein 1 (IRP1; encoded by ACO1) are part of an "ACO2-Iron Axis" that regulates iron homeostasis and is a determinant of a particularly aggressive subset of NSCLC.

P4.16 | Use of Deep Quantitative Proteomics to Predict Drug Response in Acute Myeloid Leukemia

Karin D. Rodland[1] (karin.rodland@pnnl.gov), Sara Gosline[1], James Pino[1], Sunil Joshi[2]
[1]Pacific Northwest National Laboratory, Richland, WA, [2]Oregon Health & Science University, Portland, OR

Acute myeloid leukemia is a poor prognosis cancer commonly stratified by genetic aberrations, both innate and acquired through treatment, which enable the tumor to adapt and evade treatment. Our multiomic approach, combining genomic, transcriptomic, proteomic and phosphoproteomic datasets along with *ex vivo* drug sensitivity data, offers a multidimensional view to help resolve the underlying pathophysiology of AML beyond mutations alone. We measured the proteome and phosphoproteome of a 210-patient cohort and explored how addition of this data could further stratify patient drug response and provide insight into drug resistance mechanisms. Combining proteomic with transcriptomic data helped to identify four distinct proteogenomic subtypes. A classifier, built on protein level data, enabled the subtype classification to extend across external datasets. Leveraging the *ex vivo* drug sensitivity data from these samples we identified subtype-specific drug response patterns and used these to identify drugs that showed greater efficacy in combination. We further leveraged the ability to map molecular signature to drug response to evaluate drugs that could target cell lines representing various stages of quizartinib resistance. Our results show how integration of multi-omics data, together with drug sensitivity data, can better inform therapy stratification and drug combination therapy in AML.

P5: Cardiovascular Proteomics

P5.01 | Proteomic Comparison of Novel iPSC-derived Vascular Endothelial Cells with an Immortalized Human Endothelial Cell Line

Nethika R. Ariyasinghe[1] (nethika.ariyasinghe@cshs.org), Roberta De Souza Santos[1], Andrew Gross[1], Arwin Aghamaleky-Sarvestany[1], Simion Kreimer[1], Dhruv Sareen[2], Sarah J. Parker[1]
[1]Cedars-Sinai Medical Center, Los Angeles, CA

The vascular endothelium constitutes the inner lining of the blood vessel and is directly involved in the control of vascular tone, platelet aggregation, inflammation, and angiogenesis. Malfunction and injuries of the endothelium can cause cardiovascular diseases, which are the leading causes of death globally. The impairment of the endothelium is a hallmark in other diseases such as sepsis, stroke, tumor growth, insulin resistance, and venous thrombosis. Generation of effective sources to replace injured endothelial cells (ECs) could have significant clinical impact and somatic cell sources like peripheral or cord blood cannot credibly supply enough endothelial cell progenitors for multitude of treatments. Pluripotent stem cells are a promising source for a reliable EC supply that have the potential to repair, reconstruct, replace damaged cells, and revascularize ischemic areas in order to restore tissue function and treat vascular diseases. We have developed methods to differentiate induced pluripotent stem cells (iPSCs) efficiently and robustly across multiple iPSC lines into non-tissue-specific pan vascular ECs (iECs) with high purity. These iECs present with canonical endothelial cell markers and exhibit measures of endothelial cell functionality with uptake of acetylated low-density lipoprotein (Dil-Ac-LDL) and tube formation. Using proteomic analysis, we revealed the iECs are more proteomically similar to established umbilical vein ECs (HUVECs) than to iPSCs. Post-translational modifications (PTMs) were most shared between HUVECs and iECs, and potential targets for increasing the proteomic similarity of iECs to HUVECs were identified. PTM analysis can be used in the future to glean additional insights and generate novel hypotheses. Here we demonstrate an efficient robust method to differentiate iPSCs into functional ECs, and for the first time provide a comprehensive protein expression profile of iECs, which indicates their similarities with a widely used immortalized HUVECs, allowing for further mechanistic studies of EC development, signaling and metabolism for future regenerative applications.

P5.02 | Proteomic characterization of human acellular amniotic fluid that provides cardio protection

Ryan Bia[1] (ryan.bia@utah.edu), Grace Mitchell[2], Hadi Javan[2], Jan Pierce[2], Craig Selzman[2], Sarah Franklin[2]
[1]CVRTI, University of Utah, Salt Lake City, UT, [2]University of Utah, Salt Lake City, UT

Multiple studies have demonstrated the potential in treating myocardial ischemia-reperfusion, few have translated into clinical practice. Amniotic fluid derived products are considered a promising resource for cell therapy and tissue engineering due to its diverse biologic properties that include anti-inflammatory, angiogenesis, and anti-fibrosis activities. In this study, we performed LC-MS/MS to understand the influence and mechanisms of hAF on cardiac remodeling in a diseased rat model. hAF was collected from patients at full term elective caesarean section. hAF was subsequently sterile-filtered and processed using previously published methods. Adult rats were intramyocardial injected with hAF prior to left anterior descending (LAD) coronary artery ligation for 60 minutes and subsequent release. Interestingly, rats that were administered hAF prior to LAD ligation compared to the controls maintained their cardiac function and did not experience adverse cardiac remodeling. hAF treatment was also associated with lower myocardial infarct area and tissue fibrosis. Proteomic characterization of hAF was performed via tandem mass spectroscopy. Raw files were searched against the Uniprot human database and analyzed using the INFERYS rescoring of Sequest HT search engine results in Proteome Discoverer 2.5. Analysis of hAF (n=6) identified 540 unique proteins and INFERYS workflow improved peptide identification by 28%. The top biological processes were linked to immunity, inflammatory response, and cell adhesion. KEGG pathways comprised of complement and coagulation cascades, anti-viral and anti-microbial, and neutrophil extracellular trap formation. Molecular function assessment classified the majority of proteins as protease inhibitors, hydrolases, proteases, and oxidoreductase. Overall, we are seeing that hAF preserves cardiac function following ischemia-reperfusion injury. Rather than targeting a single factor or pathway, proteomic analysis suggests that hAF provides a plethora of immune-modulating proteins that work in concert to provide this robust cardio protection. hAF is a ubiquitous, non-antigenic fluid that potentially offers a valuable biologic adjunct for treating reperfusion injury.

P5.03 | Elucidating Extracellular Matrix Dysregulation in Ischemic Cardiac Disease using a Photocleavable Surfactant and diaPASEF

Kevin Buck[1] (kmbuck2@wisc.edu), Holden Rogers[1], Zhan Gao[1], Timothy Aballo[1], Morgan W. Mann[1], Ying Ge[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Ischemic cardiomyopathy (ICM) resulting from coronary artery disease myocardial infarction (MI) is a leading cause of heart failure (HF). A characteristic of ICM is early-phase, reparative left ventricular (LV) remodeling characterized by sarcomere reorganization. Degradation and structural alteration of extracellular matrix (ECM) in the infarct zone are hallmarks of this early healing phase, but can also extend to long term inflammatory processes, adverse remodeling of non-infarcted LV tissue, and fibrosis. The ECM is critically implicated in ICM progression towards HF, but a comprehensive molecular characterization of these proteins has not been performed. Historically, ECM proteomics has been difficult due to protein insolubility and the high dynamic range of the matrisome. Herein, we applied a photocleavable surfactant-based strategy developed in-house, coupled with high-resolution mass spectrometry (MS), to enable the characterization of ECM proteins in ischemic heart remodeling. In this study, we examine ECM proteins from human cardiac apex tissue obtained from left ventricular assist device surgeries. Tissue was decellularized in high-molarity lithium chloride, and proteins were extracted by sonication in 0.5% hexylphenylazosulfonate. Digested peptides were analyzed on a timsTOF Pro mass spectrometer operating in Data Independent Analysis – Parallel Accumulation Serial Fragmentation (diaPASEF) mode. We quantify ~4000 proteins and identify ~120 statistically significant changes to ECM proteins. Aside from upregulation of periostin and collagen-VI isoforms in the ICM tissues, the most statistically significant difference we observed was in thrombospondin 3, a matricellular protein that worsens cardiac injury in mice but has not been majorly studied in human samples. Our high-salt decellularization allows us to deplete cytosolic and myofilament proteins, and the photocleavable surfactant allows deep coverage of insoluble ECM. In combination with the diaPASEF acquisition, this resulted in improved label-free quantitation of matrix proteins. This data provides new insights into the mechanism of post-ischemic remodeling in cardiovascular tissue.

P5.04 | Distinct Transcriptomic and Proteomic Profile Specifies Heart Failure Patients with Potential of Myocardial Recovery upon Mechanical Unloading and Circulatory Support

Sarah Franklin (sarah.franklin@utah.edu), University of Utah, Salt Lake City, UT

Background: Extensive evidence from single center studies indicates that a subset of chronic advanced heart failure (HF) patients undergoing left ventricular assist device (LVAD) support show significantly improved heart function and reverse structural remodeling (i.e. termed “responders”). Furthermore, we recently published a multi-center prospective study, RESTAGE-HF, demonstrating that LVAD support combined with standard HF medications induced remarkable cardiac structural and functional improvement, leading to high rates of LVAD weaning and excellent long-term outcomes. This intriguing phenomenon provides great translational and clinical promise although the underlying molecular mechanisms driving this recovery are largely unknown. Methods: To identify changes in signaling pathways operative in the normal and failing human heart and to molecularly characterize patients who respond favorably to LVAD unloading, we performed global RNA-sequencing and phosphopeptide profiling of left ventricular tissue from 93 HF patients undergoing LVAD implantation (25 responders and 68 non-responders) and 12 non-failing donor hearts. Patients were prospectively monitored via echocardiography to characterize their myocardial structure and function and identify responders and non-responders. Results: These analyses identified 1,341 transcripts and 288 phosphopeptides which are differentially regulated in cardiac tissue from non-failing control samples and HF patients. In addition, these unbiased molecular profiles identified a unique signature of 29 transcripts and 93 phosphopeptides in HF patients which distinguished responders after LVAD unloading. Further analyses of these macromolecules highlighted differential regulation in two key pathways: cell cycle regulation and extracellular matrix/focal adhesions. Conclusions: This is the first study to characterize changes in the non-failing and failing human heart by integrating multiple -omics platforms to identify molecular indices defining patients capable of myocardial recovery. These findings may guide patient selection for advanced HF therapies and identify new HF therapeutic targets.

P5.05 | Differential Identification of Ubiquitin and Ubiquitin-like Modified Proteins Using a Targeted Triple Enrichment Assay

Oleg Karpov[1] (oleg.karpov@cshs.org), Christopher Murray[1], Aleksandr Stotlan[1], Simion Kreimer[1], Jennifer Van Eyk[1]
[1]*Cedars-Sinai Medical Center, Los Angeles, CA*

Ubiquitin-like post-translational modifications (Ub-like PTMs) are a vital part of the cell’s ability to regulate internal cell processes, such as intracellular signal transduction, DNA synthesis, and cell cycle regulation to name a few. This has been canonically carried out through protein degradation and turnover, however recent research suggests that Ub-like PTMs may play an alternate modifying role, in certain cases modulating the protein’s activity. These PTMs include the well-known ubiquitin (Ub) protein, small ubiquitin-like modifier (SUMO) proteins, NEDD8, and ISG15, among several others. Moreover, it has been a challenge to identify proteins

modified with these PTMs due to their fast-degrading nature, as well as due to technical challenges within the spectrometric identification of such PTMs. Specifically, most of the Ub-like PTMs leave a conserved diGly (GG) motif on respective peptides following tryptic digestion, which then becomes a challenge to decipher which specific PTM the diGly motif pertains to. Presented here is a novel approach to distinguish proteins modified with a ubiquitin, NEDD8, and/or ISG15 PTMs, and identify them using DDA-based proteomic mass spectrometry. A transiently expressed construct was created using orthogonal tags fused with Ub, NEDD8, or ISG15, allowing for the differential enrichment of Ub-like PTM-modified proteins. Following incubation with FCCP, a stress-inducing drug, a total of 653 ubiquitinated proteins, 283 Neddylated proteins, and 437 ISG15-modified proteins were recovered from cardiac-specific AC16 cells. These results portray the ability to study the specific dynamics and interplay of specific Ub-like PTMs, their influence on intracellular protein turnover during times of stress, and ultimately their respective intracellular networks.

P5.06 | Deciphering Posttranslational Arginylation Sites in the Human Proteome Using Deep Proteomic Profiling

Zongtao Lin[1] (zongtao@wustl.edu), Joanna Gongora[1], Yixuan (Axe) Xie[1], Samaneh G. Kondalaji[1], Xingyu Liu[1], Faith M. Robison[1], Chenfeng Zhao[1], Xian Han[2], Anna Kashina[3], Benjamin A. Garcia[1]
[1]Washington University School of Medicine in St. Louis, St. Louis, MO, [2]St. Jude Children's Research Hospital, Memphis, TN, [3]University of Pennsylvania, Philadelphia, PA

Posttranslational arginylation on proteins installed by arginyltransferase ATE1 is a critical modification for mammalian cellular homeostasis and development. Absence of this modification in ATE1 knockout mice was embryonically lethal with various signs of cardiovascular defects, emphasizing the importance of arginylation in physiological processes. Existing methods have proposed plausible arginylation sites on a handful of proteins, most of which have not been further validated or functionally investigated. Therefore, protein arginylation remains an understudied field. In this work, we have developed a deep proteomic profiling approach using high-pH fractionation and mass spectrometry. The acquired data was searched and further processed by an in-house developed script "Rdiscover" to yield the endogenous arginylation sites from human cells. We firstly applied this method to HEK293T cells and extracted histones, and have identified promising peptide candidates which were then cross-validated by heavy and/or light cells using stable isotope labeling by amino acids in cell culture (SILAC) approach. By reconstituting the *in vitro* ATE1 assay, we have successfully added the isotopic *L*-arginines (Arg^0 and Arg^{10}) to the proteome-wide ATE1 substrates in the whole lysates from ATE1 knockout cells. The MS1 isotopic pattern (ratio of Arg^0:Arg^{10} = 1:1) of candidate peptides allowed us to reveal the identities and sites unbiasedly and computationally from human arginylome (*e.g.*, D18, E24, E17, and E107 in PDIA, ERO1, SSBP, and CH60, respectively). Further, we were able to validate the newly discovered sites using recombinant proteins (*e.g.*, histones) and ATE1 assay by a combination of top-down, middle-down, and bottom-up proteomics. The developed platform for arginylome discovery could be applied to any biological samples, and will pave the way for functional studies of arginylation biology in the human proteome.

P5.07 | Profiling the proteome and phosphoproteome of the heart after transient pressure overload in a porcine model

Sailee Rasam[1] (saileesu@buffalo.edu), Brian R. Weil[1], John M. Canty[1], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY

While myocyte injury after myocardial ischemia is a common phenomenon, this injury in the absence of ischemia has been difficult to characterize. One trigger for such myocyte injury(MI) can be elevated left ventricular(LV) filling pressure. Toward understanding how LV pressure overload(PO) can trigger myocyte injury, we developed a novel large animal model in the absence of ischemia. Swine heart tissues(LV,n=5) were collected from normal controls, and swine subjected to 1-hour phenylephrine(PE) infusions(18mg/hour). To understand the molecular changes underlying the observed MI, we performed proteomics and phospho-proteomics analyses on LV tissue from these animals. Protein was digested using our SEPOD strategy which ensures an exhaustive extraction of cytosolic and membrane proteins. The peptides were injected on a nano-LC setup combined with Orbitrap Fusion Lumos MS followed by data processing using our well-established and robust IonStar method which offers high-quality and accurate quantitative data. These samples were also enriched for phosphoproteins using TiO2 beads and analyzed using SPS-MS3 method on Orbitrap Lumos. We quantified >4000 proteins across 10 samples. We first looked for apoptosis-related proteins to confirm the myocyte injury observed with the IHC staining and increased cTnI as shown in our previous publication. The phosphoproteomics data showed 19 altered proteins as well as altered phosphopeptides from ~80 proteins responsible for both positive and negative regulation of apoptosis. This regulation of apoptosis in both directions was expected because of the upregulation of various cardioprotective proteins which are activated to minimize cardiac injury after a damaging stimulus like PO. Around 28 proteins belonging to mitochondria were altered which include metabolism-related proteins, as well as those, belonging to regulation of ROS including RESTAT and FXN. Various novel phosphorylation sites were found in our dataset, especially on known regulator proteins that have been implicated in cardiovascular disorders and will be explored further.

P5.08 | nanoCSC reveals novel insight into the surfaceome of primary human cardiomyocytes in heart failure

Roneldine Mesidor[1] (manepolite.mesidor@unmc.edu), Melinda Wojtkiewicz[1], Rebekah L. Gundry[1]
[1]*University of Nebraska Medical Center, Omaha, NE*

Advanced heart failure is associated with structural remodeling of cardiomyocytes and cardiac fibroblasts, leading to impaired cardiac contractility and fibrosis. Cell surface glycoproteins play crucial roles in maintenance of cardiac cell structure and function, but we do not yet fully understand how cell surface glycoproteins in human heart cells change in response to or contribute to advanced heart failure. Due to limited availability of human heart tissue, cell surface proteomic analysis of human heart cells is challenging. Here, we developed nanoscale Cell Surface Capture (nanoCSC) for surfaceome mapping of sample limited cell types. nanoCSC is a miniaturized, automated workflow that combines centrifugation-free preparation and biotinylation of cell surface glycans, automated protein digestion optimized for membrane proteins, and our novel LEVITATE system for automated magnetic bead-based surface peptide enrichment and cleanup. nanoCSC was applied to human B cells for performance testing and subsequently to primary cardiomyocytes isolated from non-failing and failing human hearts. Peptides were analyzed with an Exploris 480 MS, and data were analyzed using Proteome Discoverer, Spectronaut, R, and annotated with Veneer. From just 1000-4000 ng total peptide (0.5-5 million cells), nanoCSC quantified 400-600 cell surface *N*-glycoproteins / experiment with CV <20%. Applying nanoCSC to primary cardiomyocytes isolated from failing and non-failing human hearts revealed proteins associated with ischemia-reperfusion injury and cardiovascular diseases. Hierarchical clustering and principal component analysis group samples by condition and demonstrate that nanoCSC is sensitive enough to distinguish protein differences among cells isolated from failing and non-failing hearts. We also identified previously unreported changes in the surfaceome of cardiomyocytes between failing and non-failing hearts and identified >20 proteins associated with cell proliferation and GPCR cell signaling. In summary, nanoCSC enables surface proteomic studies of small sample sizes and reveals cell-type specific targets for elucidating pathophysiological mechanisms and future drug development for heart failure.

P5.10 | Using Covaris AFA technology to investigate cardiac perturbations in a HTP cell-based assay system

Saeed Seyedmohammad (saeed.seyedmohammad@cshs.org), Cedars Sinai Medical Institute, Los Angeles, CA

Introduction: Precision medicine via molecular phenotyping enables the quantification of an individual's proteome, where a disease risk can be assessed. Having an established pipeline such as a High Throughput (HTP) screening workflow can allow the accurate quantification for discovery, mechanism, and clinical data generation, resulting in the identification of potential biomarkers for morbidities. We developed a bottom-up proteomics approach based on HTP cell-based assay. Methods: AC16 human cardiomyocytes (Sigma) were grown in 96-well plates and optimized density of cells was lysed in a detergent-free buffer on an LE220 Plus sonication system (COVARIS). Proteins were reduced, alkylated, and digested on a Biomek i7 (Beckman) automated workstation, in a temperature regulated on-deck incubator. Protein digestion was optimized, and desalted peptides were loaded on an Ultimate 3000 (Thermo) liquid chromatography system coupled to an Exploris 480 Orbitrap (Thermo) mass spectrometer. Data-Independent Acquisition (DIA) was performed over a 45-minute gradient and searched using the Pan human library (PXD000954). Results: Data analysis allowed the quantification of 2,225 proteins over a dynamic range of 6 orders of magnitude (median over 4 replicates) for the optimized classic digestion. Addition of AFA sonication during digestion significantly increased the number of differentially detected proteins, where 450 of the 546 proteins were associated with AFA-assisted digestion and displayed an increase in intensity over the entire proteome dynamic range that also allowed the detection of lower abundant proteins. Protein interaction analysis revealed the detection of 82 additional membrane-associated proteins, of which 63 of the differentially detected proteins were associated with mitochondrial organization and transport. We then performed our optimized sonication-assisted sample preparation to highlight protein regulations involved in the response of cardiac cell perturbations. Conclusions: We identified important cardiac modulators in ischemic AC16 cells from hypoxic assays and hope to apply our method to the analysis of proteomes using stable isotope-labeling.

P5.12 | Lactate-based Purification of Human iPSC-Derived Cardiomyocytes Does Not Alter Proteome or Function of Engineered Cardiac Tissue

Kalina Rossler[1] (kjreese@wisc.edu), Jake Melby[1], Willem deLange[1], Timothy Aballo[1], Morgan W. Mann[1], Tim Kamp[1], Carter Ralphe[1], Ying Ge[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Human induced pluripotent stem cell–derived cardiomyocytes (hiPSC-CMs) are commonly utilized in *in vitro* models for heart disease. Three-dimensional engineered cardiac tissue (ECT) constructs using purified hiPSC-CMs have emerged as appealing

model systems due to their closer representation of heart tissue complexity. Recent studies suggest that widely-used metabolic (lactate) purification of 2D hiPSC-CM cultures results in an “ischemic-like” phenotype. Herein, our objective was to investigate if these initial responses to lactate purification persist when the cells are subsequently established into ECTs.
hiPSC-CMs were purified using either a lactate-based media or magnetic antibody based (MACS) depletion of the non-myocyte cell populations. Both purification methods resulted in over 98% cTnT+ hiPSC-CMs measured by flow cytometry. After purification, hiPSC-CMs were combined with hiPSC-cardiac fibroblasts to create 3D ECT constructs and maintained in culture for four weeks. Isometric twitch force and Ca^{2+} transient measurements were performed to characterize contractile performance (n=5 per condition). Using the same ECTs, proteins were extracted and analyzed using integrated liquid chromatography-mass spectrometry (LC-MS)-based approaches to both quantify changes in sarcomere sub-proteome post-translational modifications and profile global protein expression.
Functionally, the lactate ECTs displayed similar magnitude of twitch force, overall contraction times, and overall relaxation times to the MACS ECTs. Time to peak and decay in calcium transient measurements were also the same. Global proteomics revealed little differences between treatments with 99.45% overlap in protein expression (33 differentially expressed proteins out of 5500 proteins quantified). Pathway analyses revealed no significant pathway trends with differentially expressed proteins. Total phosphorylation of α-tropomyosin, myosin light chain 2v, and cardiac troponin T was consistent between conditions.
Our results reveal that lactate selection of hiPSC-CMs generate 3D ECTs constructs with phenotypes similar to MACS hiPSC-CMs. This study suggests that lactate purified hiPSC-CMs may recover phenotypically in three-dimensional constructs with time in culture.

P6: Cellular Signaling and Systems Biology

P6.01 | Coupling auxin-inducible degradation with quantitative phosphoproteomics reveals a new role for PP2ARts1 in stabilizing eisosomes during mitosis.

Andrew DeMarco[1] (demarco0@purdue.edu), Mark C. Hall[1]
[1]Purdue University, West Lafayette, IN

Protein functional characterization typically involves gene deletions, transcript knockdowns, or conditional mutations that can result in indirect physiological changes from prolonged target protein loss. This is problematic for many proteomic applications, like identification of kinase or phosphatase substrates. Chemical inhibitors are ideal, providing rapid and specific target inactivation. Unfortunately, specific chemical inhibitors are not available for most proteins. Inducible protein degradation technologies like the auxin-inducible degradation (AID) system are an attractive alternative, providing rapid depletion of virtually any target protein under diverse physiological conditions. AID acts as a surrogate for specific chemical inhibition, which minimizes non-specific effects associated with long-term target perturbation. Our lab has been interested in combining AID with proteomic methods to study signaling pathways and characterize direct kinase and phosphatase substrates. We established and validated a workflow for AID-coupled phosphoproteomics in budding yeast by targeting subunits of the abundant protein phosphatase 2A (PP2A). The AID system reduces endogenous levels of individual PP2A components by >85% within 10 minutes, leading to quantifiable phosphoproteomic perturbations by 20 minutes. Using this system, we demonstrate-PP2A in complex with its B-subunit Rox Three Suppressor 1 (PP2ARts1) contributes to the phosphoregulation of a conserved fungal-specific membrane protein complex called the eisosome. Eisosomes are attractive targets for novel antifungal therapeutics and play multiple physiological roles within the cell, including maintaining polarized growth. Rapid degradation of PP2ARts1 leads to significant increases in phosphorylation of multiple eisosome proteins, including the eisosome core protein Pil1. Fluorescence microscopy analysis of PP2ARts1-depleted cells showed a significant difference in Pil1-EGFP signal at the plasma membrane, consistent with evidence that eisosome phosphorylation leads to disassembly. We are currently testing the model that PP2ARts1 is required to maintain functional, hypophosphorylated eisosomes during mitosis, when cellular kinase activity peaks and many phosphatases are suppressed.

P6.02 | A Spatiotemporal Notch Interaction Map from Membrane to Nucleus

Marian Kalocsay[1] (MKalocsay@mdanderson.org), Alexandre Martin[2], Gary A. Bradshaw[2], Robyn J. Eisert[2], Stephen C. Blacklow[2]
[1]MD Anderson Cancer Center, Houston, TX, [2]Harvard Medical School, Boston, MA

Notch signaling relies on ligand-induced proteolysis to liberate a nuclear effector that drives cell fate decisions. The location and timing of individual steps required for proteolysis and movement of Notch from membrane to nucleus, however, remain unclear. Here, we use proximity labeling with quantitative multiplexed mass spectrometry to monitor the microenvironment of endogenous Notch2 after ligand stimulation in the presence of a gamma secretase inhibitor and then as a function of time after inhibitor removal. Our studies show that gamma secretase cleavage of Notch2 occurs in an intracellular compartment and that formation of nuclear complexes and recruitment of chromatin-modifying enzymes occurs within 45 minutes of inhibitor washout. This work provides a spatiotemporal map of unprecedented detail tracking the itinerary of Notch from membrane to nucleus after activation and identifies molecular events in signal transmission that are potential targets for modulating Notch signaling activity.

P6.03 | The role of Myristoylated, Alanine-rich C-kinase Substrate (MARCKS) in Macrophages upon stimulation of Toll-like receptor 4

Jiraphorn Issara Amphorn[1] (jiraphorn.issara-amphorn@nih.gov), Virginie Sjoelund[2], Aleksandra Nita-Lazar[2]
[1]National Institutes of Health, Bethesda, MD, [2]Functional Cellular Networks Section, Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD

MARCKS (Myristoylated Alanine-rich C-kinase Substrate) is a membrane protein expressed in many cell types including macrophages and functionally related to cell adhesion, phagocytosis and inflammatory responses. LPS (lipopolysaccharide), one of the strongest PAMPs (Pathogen-associated molecular patterns), triggers inflammation via TLR4 (Toll like receptor 4). During TLR4 stimulation, MARCKS is phosphorylated by PKC (Protein kinase C) resulting in its release to the cytosol followed by activation of inflammatory signal transduction pathways. The phosphorylation site of MARCKS (phospho-MARCKS) on serine (S163) may have a regulatory role, since we found it changes upon TLR stimulation. Serine phosphorylation serves as a key regulator of many physiological processes including innate and adaptive immune responses. Although MARCKS and the formation of phospho-MARCKS in macrophages have been described, the cellular role(s) of MARCKS and phospho-MARCKS in regulating macrophage

functions remain unclear. As a proof-of-concept study, we activated macrophages with LPS with or without addition of a PKC inhibitor. We found that PKC inhibition substantially decreased macrophage IL6 and TNF cytokine production. In addition, confocal microscopy showed that MARCKS and phospho-MARCKS increased localization to endosomes in response to LPS stimulation. Moreover, CRISPR-CAS9 mediated knockout of MARCKS in macrophages downregulated TNF and IL6 cytokine production following LPS stimulation, suggesting a role for MARCKS in inflammatory responses. Our comprehensive proteomics analysis comparing LPS-stimulation of WT and CRISPR-CAS9 MARCKS knock-out macrophages provides insight into the involvement of MARCKS in specific biological processes and KEGG pathway analysis including inflammatory responses, cytokine-cytokine receptor interaction and oxidative phosphorylation, uncovering the specific proteins involved in regulating MARCKS activity upon LPS stimulation. The discovery of the mechanism by which MARCKS contributes to the inflammatory response may provide new strategies to manipulate inflammation-related diseases.
This research was supported by the Intramural Research Program of NIAID, NIH.

P6.04 | Structural Neuromics: Proteomics and Crystallography for Substrate Discovery

Keith D. Philibert[1] (keith.philibert@rosalindfranklin.edu), Marc Glucksman[2], Xiaomeng Shao[3], Charlie Yang[3]
[1]Rosalind Franklun University (RFUMS), North Chicago, IL, [2]RFUMS/Chicago Medical School, North Chicago, IL, [3]Rosalind Franklin University (RFUMS), North Chicago, IL

Neuropeptides modulate intercellular communication for proper metabolic activity, cell differentiation and growth. Alterations in expression or processing of neuropeptides produces deleterious neuropathophysiologic consequences. Neuropeptide processing enzymes form a nexus in the synapse for regulating signaling in the extracellular milieu through its control of peptide neurotransmitters. Our Neuropeptide Discovery Pipeline (NDP) is utilized to examine novel substrates for two closely related neuropeptide processing enzymes: EC3.4.24.15 (EP24.15; thimet oligopeptidase) and EC3.4.24.16 (EP24.16; neurolysin). These enzymes were originally distinguished solely by an alteration in the bond hydrolyzed in neurotensin; hence, the designation Neurolysin. Remarkably, these enzymes are homologous to angiotensin converting enzyme 2 (ACE2), the SARS CoV2 cellular receptor, share angiotensin as a substrate and may be targets for COVID-19 intervention.
Through the NDP, new potential substrates were unveiled for processing enzymes as the repertoire of neuropeptide regulation has expanded beyond the initial half dozen substrates. Utilizing biochemical, structural, and systems-based proteomic approaches including *in silico* molecular modeling, mass spectrometry, enzyme kinetics, and X-Ray crystallography, we have discovered novel substrates for EP24.16 that differ substantially from the closely related EP 24.15 both in detailed cleavage sites and in the enzyme active site as a true substrate. Several examples include the neuroprotective, hybrid neuropeptide substrate Colivelin and the EP 24.15 cleaved reproductive and cardiac peptide Phoenixin. Additionally, the gonadotropin inhibitory neuropeptide RFRP-3 exhibits different cleavage sites and is a novel peptide. Interestingly, Xenin 8, a gut peptide, with a C-terminal sequence very similar to Neurotensin, is cleaved at the same site with both EP24.15 and EP24.16 enzymes.
With the discovery of novel neuropeptides and their functions coupled with advanced technological methodologies, one must be cautious with utilizing older, misleading enzyme designations and limited substrates for the processing of the expanding repertoire of neuropeptides.

P6.05 | Identification of RNA-binding proteome in mouse macrophages by total RNA-associated protein purification

Deepali Rathore[1] (deepali.rathore@nih.gov), Aleksandra Nita-Lazar[1]
[1]National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD

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), which results in a ribonucleoprotein complex (RNP). Any perturbations in RNA-protein interactions can lead 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 SILAC mass spectrometry-based approach to identify and quantify RNA associated proteome at a global level in mouse macrophages. This entails 254nm ultraviolet crosslinking of RBPs to RNA in live cells, followed by their capture on silica and selective elution of peptides. The RBPs are then enzymatically digested and analyzed with LC-MS/MS. The experimental design included comparisons between UV crosslinked and non-crosslinked mouse macrophages as controls as well as comparisons between lipopolysaccharide (an endotoxin that initiates inflammatory response through Toll-like receptor 4) activated vs non-activated samples, at three different time points: 0 ,30, 60 min. The GO term analysis showed significant enrichment for RNA binding proteins, with RNA transport, spliceosome and ribosome as the most enriched KEGG pathways, in the UV crosslinked samples. RBPs identified were also found to be associated with inflammatory responses, such as IL1α, Irg1, wdr26, Hdac1, Marcks and Eifs. Upon differential expression analysis, >90% reduction in the RNA binding of proteins was observed in the LPS treated samples, indicating their possible dissociation from the stress granules upon phosphorylation in the MAPK pathway. The analysis also identified some novel candidates not previously annotated as RBPs but could be possible regulators in TLR signaling pathways. Experiments are underway to identify the amino acids at the crosslinking site in the RNA oligonucleotide moiety as well as specific RNAs bound to the proteins of interest. This research was supported by the Intramural Research Program of NIAID, NIH.

P6.06 | Comprehensive Proteomic Analysis of Human 3D Cartilage Tissues with Induced Senescence Provides Insight into Drivers of Joint Disease

Charles A. Schurman[1] (cschurman@buckinstitute.org), Jacob P. Rose[1], Christina D. King[1], Sandip K. Patel[1], Joanna Bons[1], Samah Shah[1], Judith Campisi[1], Birgit Schilling[1]
[1]*Buck Institute for Research on Aging, Novato, CA*

Proteomic studies analyzing the skeleton are crucial for understanding molecular mechanisms behind age-related diseases like osteoarthritis (OA). Dysregulation of cell signaling and increases in senescence within cartilage is implicated in OA, thus understanding skeletal senescence and its associated secretory phenotype (SASP) will help identify interactions that drive disease. We generated 3D tissues from human cartilage tissue donors, investigated the intracellular proteome and the SASP produced from chondrocyte spheroids, and compared senescent chondrocytes (irradiation induced) to control, quiescent chondrocytes. Protein lysates were digested using S-Trap (Protifi) and desalted by HLB columns. Samples were analyzed on an Orbitrap Eclipse Tribrid. Digested sample aliquots were pooled and fractionated using high pH fractionation to generate a comprehensive sample-specific spectral library. All samples were analyzed using data-independent acquisition (DIA).
Protein extracts from chondrocyte 3D tissues showed enrichment for collagen types II, IX, and XI accounting for 45% of all collagen peptides despite representing only 5 out of 18 subtypes detected, confirming a primarily cartilaginous ECM. Important cartilaginous proteoglycans, such as Aggrecan and Chondroitin sulfate proteoglycan 4, were detected further confirming the cartilage tissue origin.
Preliminary results investigating the role of senescence in joint tissue degradation showed significant regulation of 847 proteins, 72% of which mapped to existing SASP proteins as identified by the SASP atlas. Within this proteome, 534 proteins were down regulated and involved in KEGG pathways involving ribosomal and amino acid synthesis and glycolytic function. These metabolic changes were accompanied by increases in pathways for galactose and pyruvate metabolism and an increased reliance on proteasome degradation. These changes imply cell-intrinsic metabolic shifts that may drive joint degeneration.
Monitoring proteins related to the SASP throughout OA progression provides an opportunity to discover novel biomarkers. Here, we demonstrate the susceptibility of human chondrocytes to senescence and changes in cellular metabolism during degenerative disease.

P6.07 | Proteomic profiling reveals dysregulation of Chlamydomonas CLiP strain under stress conditions

Amanda L. Smythers (asmyther@email.unc.edu), UNC Chapel Hill, Chapel Hill, NC

Chlamydomonas reinhardtii is a model unicellular green alga that has facilitated detailed biomolecular investigations into fundamental biochemical processes. *C. reinhardtii* grows rapidly and can grow photoautotrophically, mixotrophically, and heterotrophically, facilitating comprehensive studies of photosynthetic function and regulation. Furthermore, genetic advances have advanced the accessibility and precision of mutant generation, with ballistic plasmid insertion, CRISPR, and high throughput genetic screening of insertional mutants enabling low cost strains that target biochemical pathways of interest. Research interest in *C. reinhardtii* led to the creation of the Chlamydomonas Library project, through which more than 60,000 haploid colonies are maintained and currently available for purchase. However, these mutants are not comprehensively characterized, and it is unclear in what ways they differ from Chlamydomonas wild type strains. To better understand how the CLiP background strain (CMJ030) compares to CC-2895, a common cell-wall replete wild type strain, we leveraged quantitative global and redox proteomics with physiological and photosynthetic characterization to probe for phenotypic differences. Although limited during basal growth, the differences became apparent following the inhibition of target of rapamycin (TOR) kinase, with largescale translational and post-translational dysregulation observed that matched disruptions in physiological conditions. Together, these data suggest that the CLiP background strain has a vastly altered stress response network that lead to significant differences in signaling cascades, emphasizing the importance of proteomic phenotyping prior to experimentating with novel genetic variants.

P7: Chemical Proteomics and Drug Discovery

P7.01 | Determining the Effects of Sphingosine Kinase 1 Targeting PROTACs on the Human Pancreatic Cancer Proteome

Tyler M. Bateman[1] (tboneb497@gmail.com), Rebecca Burlett[1], Brayden Wehrli[1], John C. Price[1], David J. Michaelis[1], Mariur Rodriguez Moreno[1]
[1]Brigham Young University, Provo, UT

Pancreatic cancer claims around 50,000 lives every year. Our prior research identified the Sphingosine Kinase 1 (SphK1) protein as a key oncogenic driver that could be targeted to counteract the development of pancreatic cancer. We are testing the efficacy of a new class of therapeutics known as proteolysis-targeting chimeras (PROTACs) to target Sphk1 for degradation by the ubiquitin proteasome system. Preliminary tests on cell growth and proliferation were used to identify leading compounds in a variety of cell lines including human pancreatic cancer cells (PSN-1). We are investigating the effect of targeted degradation on the proteome for some of these lead compounds. By analyzing the change in concentration and turnover for a variety of proteins across the proteome we can understand the changes that targeted enzyme degradation causes in these cells, including off-target effects and other possible side effects of drug. This presentation focuses on the analysis of proteomic data that has been corroborated with western blotting and cell proliferation assays to determine the effects of novel SphK1 targeting PROTACs. This proteomic data was collected using deuterium labelling to determine protein turnover rates and label free protein concentrations. The combination of these measurements will help us better understand the true effects of the candidate drugs on human pancreatic cancer cells and SphK1.

P7.02 | Cellular Protein Profiling to Identify Herbicide Exposure Targets

Joseph Genereux[1] (josephg@ucr.edu), Guy Quanrud[1]
[1]University of California, Riverside, Riverside, CA

Reactive environmental toxicants, such as electrophiles and oxidants, can damage proteins. If this damage lowers a protein's stability, then the protein can misfold and present a proteotoxic threat to the cell. Despite this risk, the protein targets of abundant toxicants are not generally known. To identify proteins that are destabilized by cellular toxicant exposures, we have developed a mass spectrometry-based Hsp40 affinity profiling assay. In this assay, a mutant DNAJB8 serves as a recognition element for misfolded cellular proteins, allowing them to be isolated and quantified. We used this assay to identify destabilized proteins following exposure of HEK293T cells to three structurally similar chloroacetanilide herbicides: alachlor, propachlor, and acetochlor. While hundreds of proteins are targeted by these herbicides, and most targets harbor experimentally reactive cysteine residues, the profiles of these herbicides are distinct and include many proteins that lack general haloacetamide reactivity. These finding suggest that neither warhead reactivity nor protein reactivity on their own primarily drive targeting and destabilization. Two of the most selective propachlor targets are GAPDH and PARK7, both of which are strongly implicated in Parkinson's Disease. We find that cellular activities of both proteins are inhibited by propachlor, with GAPDH inhibition driven by direct propachlor conjugation to C152. Destabilization of these proteins by propachlor might contribute to the strong association between occupational herbicide exposure and Parkinson's Disease risk.

P7.03 | Piperlongumine (PL) conjugates induce targeted protein degradation

Dongwen Lyu (lyud@uthscsa.edu), UT Health San Antonio, SAN ANTONIO, TX

PROteolysis Targeting Chimeras (PROTACs) are bifunctional molecules that degrade target proteins through recruiting E3 ligases. However, their application is limited in part because few E3 ligases can be recruited by known E3 ligase ligands. In this study, we identified piperlongumine (PL), a natural product, as a covalent E3 ligase recruiter, which induces CDK9 degradation when it is conjugated with SNS-032, a CDK9 inhibitor. The lead conjugate **955** can potently degrade CDK9 in a ubiquitin-proteasome-dependent manner and is much more potent than SNS-032 against various tumor cells *in vitro*. Mechanistically, we identified KEAP1 as the E3 ligase recruited by **955** to degrade CDK9 through a TurboID-based proteomics study, which was further confirmed by KEAP1 knockout and the nanoBRET ternary complex formation assay. In addition, PL-Ceritinib conjugate can degrade EML4-ALK fusion oncoprotein, suggesting that PL may have a broader application as a covalent E3 ligase ligand in targeted protein degradation.

P7.04 | Implementation of a Robust In-Cell Covalent Fragment Screening Platform

Alex Shannon (alex.shannon@abbvie.com), AbbVie, Inc., North Chicago, IL

Expanding the fragment-based drug discovery paradigm to covalent compound classes has enabled the identification of ligandable sites on “undruggable” protein targets. Using chemical probes to competitively label reactive amino acids in complex environments allows for this ligandability to be assessed in the context of lysates and intact cells. Here, we integrate recent advancements in proteomics—including laboratory automation, SP3 protein cleanup, data-independent acquisition, and automated data analysis and annotation—into a robust and high-throughput covalent fragment screening platform. We subsequently validated a set of kinase-binding covalent fragments discovered by this workflow via an orthogonal activity assay. These platform improvements have enabled wider library building and rapid identification of hits alongside considerable time savings.

Disclosure statement: All authors are employees of AbbVie. The design, study conduct, and financial support for this research were provided by AbbVie. AbbVie participated in the interpretation of data, review, and approval of the publication.

P7.05 | Proximity-labeling Chemoproteomics Defines the Subcellular Cysteine Redoxome

Tianyang Yan[1] (tyyan14@outlook.com), Heta Desai[1], Keriann Backus[1], Lisa M. Boatner[1], Ashley Julio[1], Miranda Villaneuva[1], Anthony Jones[1], Andrea Bell[1], Ajit Divakaruni[1], Alex Turmon[1]
[1]*University of California Los Angeles, Los Angeles, CA*

Proteinaceous cysteines function as essential sensors of cellular redox state. Defining the cysteine redoxome is a key challenge for functional proteomic studies. Recent methods, including our SP3-Rox platform, have improved the throughput and coverage of redox sensitive cysteines, yielding inventories of cysteine redox state for thousands of cysteines proteome-wide. As such methods assess bulk proteomes, they fail to capture subcellular differences in cysteine oxidation state, which are expected given the established spectrum of organelle redox potentials. Here we report the development of proximity-labeling based chemoproteomic profiling of cysteine redoxome, “PL-CysOx”, which enables the quantification of compartment-specific changes in cysteine oxidation. Our method relies on two key innovative steps. First, we developed a novel two-step biotinylation procedure that enables capture of subcellular cysteines, “PL-Cys”. This dual enrichment method combines TurboID proximity labeling with “click chemistry” based bioorthogonal handles to enrich cysteines. Implementation of PL-Cys identified > 1,000 cysteines that are uniquely identified for each compartment targeted, spanning endoplasmic reticulum (ER), mitochondria (Mito), golgi and nucleus. Second, we achieved enhanced subcellular proximity labeling specificity by combining TurboID with translation arrest. We observe that cycloheximide treatment simultaneously affords a decrease in TurboID mislocalization and an increase in specificity of mitochondrial captured cysteines, from 23% to 45%. By combining these two innovations with SP3-Rox method, we measure cell state dependent changes in mitochondrial cysteine oxidation. Using a TurboID expressing macrophage cell line, we compare cysteine oxidation in the presence and absence of treatment with lipopolysaccharide (LPS), which causes increased expression of nitric oxide synthase and decreased mitochondrial respiration. The PL-CysOx method quantified the oxidation state of 2,182 cysteines and identified 35 mitochondrial cysteines showing statistically significant LPS-induced oxidation. Collectively, PL-CysOX achieved chemoproteomic profiling of cell dependent changes of cysteine redoxome and provided a platform or analysis of functional redox sensitive cysteines with high subcellular specificity.

P8: Computational Proteomics and Data Science

P8.01 | Applying Automated Machine Learning to Accelerate Large-Scale Proteomics Data Analysis

Amir Alavi[1] (aalavi@seer.bio), Harendra Guturu[1], Guhan R. Venkataraman[1], Ryan W. Benz[1], Asim Siddiqui[1], Khatereh Motamedchaboki[1], Serafim Batzoglou[1]
[1]*Seer, Inc., Redwood City, CA*

Advances in unbiased, high-throughput proteomics technologies, including the Proteograph™ product suite (Seer Inc.) coupled with liquid chromatography mass spectrometry (LCMS)-based proteomics, have enabled the profiling of thousands of proteins from a single LCMS run, and the creation of very high dimensional datasets for downstream analysis. Machine learning tasks such as classification and regression require a long process of exploratory data analysis, data preprocessing, normalization, imputation, feature selection, feature transformation, and model selection. This yields a large search space of modeling choices that is infeasible to search exhaustively. The unique properties of each dataset can lead to bespoke solutions that poorly generalize across datasets. Conducting this search for each new dataset is impractical, as it can require both machine learning expertise and domain knowledge.
To address these challenges, we developed a machine learning software package (seerml) which extends the auto-sklearn software package for Auto ML, leveraging recent advances in Bayesian optimization, and tailors its use for large scale LCMS proteomics data analysis. We do this by incorporating custom feature selection strategies, custom search space restrictions and parameters, as well as novel model inspection functions and figures for analysis of the resulting models.
We apply seerml to data collected with the Proteograph™ workflow for plasma samples of multiple disease cohorts for both classification and regression tasks. Seerml can both select algorithms for data preprocessing and the machine learning task, while also tuning the hyperparameters. We found that in most cases the models selected by seerml perform at least as well as manually-tuned baselines while requiring minimal human input and machine learning expertise. In some cases, seerml outperforms even highly human optimized models by over 2.5%. We additionally investigated the stability of the Auto ML models and the features being identified to better understand their viability to support biomarker discovery.

P8.02 | Utilizing Data-Independent Acquisition as a Substitute for Data-Dependent Acquisition for the Detection and Quantification of Low-Abundant Biomarkers

Noah Earls[1] (noaheearls@gmail.com), Noah Moran[1], Nate Zuniga[1], John C. Price[1]
[1]*Brigham Young University, Provo, UT*

Data-Independent Acquisition (DIA) is a relatively new method for Mass Spectrometry (MS) experiments. DIA differs from Data-Dependent Acquisition (DDA) through the introduction of pooled samples, each designated for a unique m/z window. The primary issue concerning DDA analysis is that MS1 scans select only the highest MS1 signals for MS2 fragmentation. This biases the results towards high abundant peptides (HAPs) and often inhibits the signal from other ions. Low abundant peptides (LAPs), especially those that form biomarkers, may be detected in MS1 but obtain lower signals and are not selected for MS2 fragmentation. In DIA, small regions of the MS1 scan are selected for MS2 fragmentation. In principle this allows more peptides to be fragmented, which leads to increased protein count but also leads to increased complexity in the MS2 analysis. To compare DIA to DDA methods, we conducted two analyses using brain tissue harvested from a transgenic mouse model. We performed standard DDA analysis of the top 20 ions in each MS1 scan. For DIA analysis, we divided the MS1 range into several 100 m/z windows to provide thousands of additional MS2 signals. The increased complexity of the DIA analysis was handled by the creation of a gas-phase fractionated chromatogram library through WALNUT (Searle et al. 2018). Our DIA results demonstrate higher reproducibility of detection and quantitation when compared to DDA experiments. The DIA data consistently quantifies hundreds of additional proteins previously undetected by our DDA methods. This comparison indicates that shotgun proteomics with DIA techniques can improve upon DDA experiments and offer increased breadth and reproducibility of peptide detection. Our initial analysis also suggests that DIA could serve as a pathway-centric shotgun MS experiment that provides focused interpretation of pathway disruptions through the measurement of additional protein concentration changes and allows for the identification of previously undiscovered ontology-level interactions.

P8.03 | A systematic characterization of protein complex stability

Samantha Fischer[1] (sfisch6@uic.edu), Anna Mallam[2], Kevin Drew[1]
[1]University of Illinois of Chicago, Chicago, IL, [2]University of Texas at Austin, Austin, TX

Large macromolecular protein assemblies are responsible for carrying most functions within cells, such as protein degradation, DNA replication, and translation. Examples include the proteasome, the MCM complex, and the ribosome. To carry out their roles within the cell, complexes need to be assembled into stable architectures. The structures of many complexes are unknown and their stabilities are not well characterized. Here, we explore the stability of protein complexes using a high throughput differential-fractionation mass spectrometry (DIF-FRAC) framework treated with urea, a known denaturant. The resulting analysis identifies over 7,000 unique proteins and when compared to the 7,000 protein complexes of the Hu.MAP 2.0 resource, proteins from 4,500 potential protein complexes (and sub-complexes) are identified. We identify destabilization occurring between co-eluting proteins within known complexes through subsequent experiments of biochemical size exclusion chromatographic separation under a urea titration treatment. This high throughput technique allows us to track changes in stability across urea concentrations for thousands of complexes at once. We observe destabilization behavior both globally and at the level of individual complexes. We developed a scoring function, COSMIC - Complex Stability Metric, to quantify individual complex stability in each denaturant concentration condition, which uses pairwise correlations of complex subunits and entropy assessments to evaluate complex stability. Further, we capture differential stability behavior between and within complexes, such as the proteasome core exhibiting higher COSMIC scores than the greater holoenzyme and the MCM complex. We computationally modeled changes in stability behavior across urea concentrations to determine the half-maximal urea inhibitory concentration (urea-IC50) of each complex and we can rank the stability of ~1000 complexes and sub-complexes based on their changes in COSMIC over increasing urea concentrations. Assessment of complex stability will be a valuable tool for informing therapeutic design and will contribute to greater understanding of human disease.

P8.04 | A unifying, spectrum-centric approach for the analysis of peptide tandem mass spectra

Daniel Zolg[1], Florian Seefried[1], Tobias Schmidt[1], Siegfried Gessulat[2], Michael Graber[1], Samia Ben Fredj[1], Patroklos Samaras[1], Bernhard Kuster[3], Mathias Wilhelm[3], **Martin Frejno[1] (martin.frejno@msaid.de)**
[1]MSAID, Garching b.München, Germany, [2]MSAID, Berlin, Germany, [3]Technical University of Munich, Freising, Germany

Introduction: Mass spectrometry-based proteomics data is acquired using either data dependent (DDA), data independent (DIA) or targeted acquisition (PRM). Typically, the former is analyzed using spectrum-centric algorithms assuming that it generates non-chimeric spectra, while the latter two are analyzed in a peptide-centric fashion. Here, we introduce a spectrum-centric approach that deconvolutes spectra irrespective of isolation window size and demonstrate that it generalizes to any chimeric MS2 spectrum, unifying the analysis of DDA, DIA and PRM data.
Methods: Our library-free and spectrum-centric algorithm compares predicted and experimental MS2 spectra using various intensity-based scores. All promising peptides in each MS2 isolation window are considered simultaneously and compete for experimental fragment ion intensity in one concerted step. Our algorithm aims at explaining as much experimental intensity as possible with as few candidate peptides as possible and distributes the intensity of shared fragment ions to peptide spectrum matches given their estimated proportional contribution to the experimental MS2 spectrum, resulting in the deconvolution of chimeric spectra. FDR-control is performed using Mokapot.
Results: First, we demonstrate the benefits of our algorithm on 1h DDA measurements of HeLa cell lysates with varying isolation width, resulting in an identification rate of up to 80% and over 120k PSMs. Next, we compare DDA and DIA measurements of the same samples and benchmark various DIA tools against each other. We observe a large overlap in peptide identifications and a high correlation of quantitative values. Finally, we demonstrate the accuracy of our automated quantification method on a PRM dataset, achieving similar performance as manual peak integration using Skyline.
Conclusions: Conceptually, a DDA MS2 spectrum is indistinguishable from a DIA or PRM spectrum acquired with the same isolation window. Here, we developed a spectrum-centric approach that unifies the analysis of DDA, DIA and PRM data without the need for separate search algorithms.

P8.05 | Systematic analysis of DIA LC-MS protein rollup strategies and their impact on phenotype association and proteogenomic applications

Harendra Guturu[1] (hguturu@seer.bio), Jian Wang[1], Yingxiang Huang[1], Guhan R. Venkataraman[1], Amir Alavi[1], Ryan W. Benz[1], Khatereh Motamedchaboki[1], Matthijs B. De Geus[2], Sudeshna Das[2], Pia Kivisäkk[2], Steven E. Arnold[2], Asim Siddiqui[1], Serafim Batzoglou[1]
[1]Seer, Inc., Redwood City, CA, [2]Massachusetts General Hospital, Cambridge, MA

Nanoparticle-based sample preparation coupled with liquid chromatography–mass spectrometry (LC–MS) enables untargeted measurement of the proteome from biofluids at unprecedented depth.1 The deep proteome enables previously inaccessible applications such as phenotype diagnosis, biomarker discovery and proteogenomics from easy to obtain plasma samples. But, to improve the value of the measured data for these applications, we consider whether the LC-MS data is appropriately analyzed and summarized from the acquired peptide identifications.
To address peptide to protein roll-up challenges using DIA methods, we have analyzed a case/control cohort of several hundred samples analyzed with data-independent acquisition (DIA) using a 30-minute gradient on Bruker timsTOF Pro 2. We searched the DIA data using DIA-NN, which offers two different protein inference modes and three different quantification modes, including MaxLFQ. Combining the options provided by DIA-NN with Seer internally developed libraries and additional peptide to protein roll up strategies, we evaluated each strategy for summarizing the proteome by three metrics: 1) accuracy of the phenotype predictor, 2) number of significant marker associations with the phenotypes, and 3) number of protein quantitative trait loci (pQTLs) using the paired genotype data.
We found that the proteome measurements yield generally robust results across the applications irrespective of roll up strategies. Some of the rollup strategies yielded small improvements with maximum difference ~10%. Interestingly, we found even unrolled up peptide level measurements gave robust results, sometimes outperforming the rolled-up values.
1Blume, John E., et al. "Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona." Nature communications 11.1 (2020): 1-14.

P8.06 | SomaScan Platform Confirmation and Performance Validation

Ted Johnson[1] **(tjohnson@somalogic.com)**, Darryl Perry[2]
[1]SomaLogic Operating Company, Inc., Boulder, CO, [2]SomaLogic Operating Company, Inc., Longmont, CO

Introduction: The SomaScan® Platform for proteomic profiling uses 7288 (7k) SOMAmer® reagents, single-stranded DNA aptamers, to 6596 unique Human Protein Targets. The modified aptamer binding reagents, SomaScan assay, its performance characteristic for 5k and 7k content sets, and specificity to human targets have been previously described. We combine profiles of validation and performance metrics with orthogonal confirmation of specificity from published literature to provide a comprehensive view of the specificity and utility of the SomaScan Platform.
Methods: Validation of SOMAmer reagents results in a set of metrics that profile performance of the reagents to the protein standards used for discovery. These include linear ranges and affinity measures: dose-response 50 (apparent Kd), and solution Kd. Validation of the SomaScan Platform includes replicates of individual and pooled samples over 15 assay runs in both plasma and serum. Population ranges and performance are generated from matched plasma and serum drawn from more than 1,000 U.S. normal volunteers. Reproducibility and signaling metrics are summarized and reported. Production use of the SomaScan Platform includes replicates to monitor the accuracy and precision of the assay over time. Results from more than 3,000 replicates are aggregated and reported. Secondary confirmation of specificity is explored using published outcomes from alternative proteomic or genomic profiling methods. Results are extracted from the literature and assembled by reagent identifier. Alternative experiments that confirm protein targets are described and reported separately.
Results: Specificity and performance summaries from reagent validation, platform validation, and platform production data are presented alongside confirmation of specificity from published literature. A complete list of the proteins is included with primary metrics for review.
Conclusions: The SomaScan® Platform for proteomic profiling relies on a deep validation workflow for reagents and for the platform. Transparency in the methods and results is critical to help users interpret platform results.

P8.07 | Single-cell mass spectrometry-based proteomics enables causal inference in observational studies

Devon Kohler[1] **(kohler.d@northeastern.edu)**, Karen Sachs[2], Jeremy Zucker[3], Charles Tapley Hoyt[4], Olga Vitek[5]
[1]Northeastern University, Boston, MA, [2]Next Generation Analytics, Palo Alto, CA, [3]Pacific Northwest National Laboratory, Richland, WA, [4]Harvard Medical School, Bonn, Rheinland-Pfalz, Germany, [5]Northeastern University, Boston, MA

Characterization of proteome response to perturbations has long been a goal of proteomics, and an important step towards understanding protein function. Recently, two types of technological advances allowed us to gain additional insights. On one hand, statistical and machine learning causal inference methods allow us to estimate the impact of interventions from observational data (i.e., in absence of experimental interventions). These methods take as input protein abundance and can estimate the impact of an intervention on downstream proteins even when some proteins on the causal path are not quantified or other threats to validity are present such as unobserved confounding . On the other hand, methodological advances in single-cell proteomics improved both the scale and quality of the acquired data. These advances increased the number of replicate cells and increased the cellular resolution, focusing the investigations to more homogeneous cell types, making this technology particularly suitable for causal effect estimation. We propose to leverage the advances in causal inference and single cell proteomics to estimate the impact of interventions on downstream proteins from purely observational data, i.e. without implementing the interventions, with the goal of identifying causal effects that may not be easily detectable with traditional methods, such as regression and correlation. We demonstrate the utility of

the proposed approach on a real-world single-cell proteomic experiment investigating Melanoma . We illustrate ways in which estimation of outcomes of interventions can be used as guidance in planning future experiments. To highlight the value of single-cell experiments, we contrast this estimation with the estimation on a bulk LC/MS-MS experiment investigating Melanoma . Finally, we highlight some of the current limitations in the current state of single cell proteomics and causal inference and discuss how future work could increase the accuracy and scope of the estimation.

P8.08 | Assessing the Proteome Impact of Alternative Splicing using Structural and Functional Prediction Algorithms

Robert W. Ludwig[1] (robert.ludwig@cuanschutz.edu), Edward Lau[2], Maggie P. Lam[2]
[1]University of Colorado School of Medicine, Denver, CO, [2]University of Colorado School of Medicine, Aurora, CO

Alternative splicing expands the coding capacity of the genome, but the extent to which splicing contributes to proteome diversity is debated. Experimental characterization of isoform proteins remains challenging, which hinders fuller understanding of biological function. Recent advances in protein structural and functional prediction present a powerful new avenue for screening predicted protein structures for biological interest. Here, we report a computational workflow to identify protein isoform sequences from RNA-seq data, followed by structural and functional prediction to estimate their biological significance.
We used the JCAST (Junction Centric Alternative Splicing Translator) software we previously developed to perform in-silico translation of likely translatable isoform transcripts in a human heart RNA-seq data set, prioritizing 3,932 high-abundance, in-frame isoforms. As proof-of-concept, we targeted 224 isoforms shorter than 1,400 residues and belonging to proteins localized to mitochondria, to predict how isoform sequences change predicted protein structure and function. We assembled a pipeline where folded structures of isoform sequences were predicted using local-colabfold, and canonical structures were drawn from public databases. Molecular functions of all structures were predicted using the DeepFRI server, while sequences were submitted to the flDPnn server for disorder and disorder-function prediction. Custom scripts were used to visualize and compare canonical versus isoform predictions to categorize the potential structural and functional impacts of alternative splicing.
Of the 224 analyzed isoforms, 74 failed to produce predictions; 77 showed no predicted difference between isoform and canonical sequences; 41 isoforms showed decreased prediction certainty, with 18 falling below 50%. Contrarily, 23 isoforms showed increased certainty in predicted functions, and 7 gained new predicted functions. Comparing multiple isoforms from the same protein further provided opportunities to contrast the effects of sequence alterations at different protein sites. This method shows promise in screening large sets of isoform sequences for targets for confirmatory studies and functional assays.

P8.09 | Outcome Variance Between Sample Normalization Methods

Dylan Maxwell (dmaxwell@somalogic.com), SomaLogic, Denver, CO

Standard deliverables for SomaScan® data include two ADATs, flat files containing protein measurements: Pre-normalized and Final (with sample normalization). The Pre-normalized ADAT is provided for clients to apply their desired sample normalization. Client feedback was received regarding variation in statistical analysis and model outcomes, attributed to the use of varying sample normalization methods. Here we investigate how Adaptive Normalization by Maximum Likelihood (ANML), Median normalization, and Quantile normalization (a method used by various clients) affect the SomaScan® RFU data in EDTA-Plasma compared to the pre-normalized dataset, along with how the observed changes impact statistical analysis and model outcomes. The standard sample normalization is ANML normalization using a healthy U.S. adult population reference. If the study population has large variations in total protein compared to the population reference, study-specific internal Median normalization is applied. Quantile normalization is a global adjustment method, used under the assumption that each sample has the same statistical distribution and "is achieved by forcing the observed distributions to be the same and the average distribution, obtained by taking the average of each quantile across samples, is used as the reference." ANML normalization and Median normalization recapitulated previous results identifying SOMAmers® with a strong association with sex and age. Under Quantile normalization, it was observed that the fold change between the mean signal of males and females was greatly reduced in analytes previously shown to be associated with females. Additionally, the age linear model, when applied to the Quantile normalized dataset, saw large drops in p-value rank and absolute Beta coefficient of SOMAmers® previously shown to be negatively associated with age. We have shown that Quantile normalization is an inappropriate method for sample normalization of SomaScan® data. Further investigation into additional sample normalization methods and models is warranted to fully address client feedback and identify other problematic normalization methods.

P8.10 | MatrisomeDB: The Extracellular Matrix Proteomics Database

Xinhao Shao[1], Clarissa Gomez[1], Nandini Kapoor[1], James Considine[1], Chris Grams[1], Yu Gao[1], **Alexandra Naba[1] (anaba@uic.edu)**
[1]*University of Illinois at Chicago, Chicago, IL*

The extracellular matrix (ECM) is a complex assembly of proteins that constitutes the scaffold organizing cells, tissues, and organs. ECM remodeling is integral to physiological events, including development and aging. An imbalance in ECM, caused by mutations in ECM genes or by alterations in ECM metabolism, is a hallmark of many diseases such as cancer and fibrosis. The ECM thus constitutes an immense reservoir of disease biomarkers and therapeutic targets that has remained, until recently, largely under-explored. Over the past decade, mass-spectrometry-based proteomics has become the method of choice to profile the composition of the ECM, or the "matrisome", of tissues. To assist non-specialists with the reuse of ECM proteomic datasets, we released MatrisomeDB (https://matrisomedb.org), compiling ECM proteomic data generated and reprocessed using unified search parameters. The 2022 release of MatrisomeDB includes 42 curated studies on the ECM of 39 organs across physiological systems, ages, and disease states, including datasets on 8 cancer types (breast, colon, lung, pancreas, ovary, stomach, insulinoma, melanoma).

Using *de-novo* sequence analysis, we predicted that there is 1027 genes in the human genome encoding ECM and ECM-associated proteins. The aggregation of the datasets from MatrisomeDB resulted in the detection of 97.8% of the predicted matrisome. The recent expansion of MatrisomeDB also resulted in achieving near-complete protein sequence coverage for all matrisome protein families, like collagens or proteoglycans (93.2% on average). The latest development in MatrisomeDB also include the mapping of peptides and post-translational-modifications onto 2D domain-based representations and AlphaFold-generated 3D structures of ECM proteins, and referencing external resources to facilitate the design of targeted mass spectrometry assays for ECM proteins. MatrisomeDB has the potential to uncover novel facets of ECM proteins, like the detection of isoforms or single amino-acid variants, and support novel hypotheses regarding the roles of these proteins in pathophysiological contexts.

P8.11 | SEPepQuant Enables Comprehensive Protein Isoform Characterization in Shotgun Proteomics

Lindsey Olsen[1] (Lindsey.Olsen@bcm.edu), Yongchao Dou[1], Bing Zhang[1]
[1]*Baylor College of Medicine, Houston, TX*

Alternative splicing occurs in over 90% of human genes and is important to progression of diseases such as cancer. Large-scale analysis of splicing isoforms has been limited to transcriptomics studies. Protein isoform analysis in shotgun proteomics is challenging because many peptides map to multiple protein isoforms. Currently, the predominate approach to handling degenerate peptides is to assign them to the isoform with the most mapped peptides for quantification. This parsimonious method fails to accurately represent the underlying data and greatly limits or even incorrectly interprets protein isoform quantification. Here, we present a graph theory approach that models all identified peptides in a study, all proteins the peptides map to, and the proteins' host genes, in a tripartite graph in which the three types of vertices are connected via their mapping relationships. Peptides connecting to the same set of proteins in the graph are identified as Structurally Equivalent PEPtides (SEPEP), which is used as the quantification unit. In SEPEP based quantification (SEPepQuant), all identified SEPEPs passing SEPEP level false discovery rate (FDR) cutoff of 1% are quantified by rolling up the abundance of their associated peptides and then classified into five classes based on their source proteins and genes for reporting. We applied SEPepQuant to compare chemotherapy resistant and sensitive samples in a triple negative breast cancer (TNBC) dataset and a high-grade serous ovarian cancer (HGSOC) dataset, respectively, and identified more than 100 genes with significant isoform-level regulation between resistant and sensitive samples. For example, one SEPEP associated with a 5' exon inclusion in HNRNPM was significantly up in chemotherapy resistant compared to sensitive TNBCs. Overall, using SEPeps instead of parsimonious protein groups as the reporting and quantification units, SEPepQuant maximizes protein isoform information that can be extracted from shotgun proteomics data to boost discoveries in biological and translational research.

P8.12 | Resolving Proteome Complexity with Isotope Profiles for Protein Turnover Studies Using Heavy Water Metabolic Labeling and LC-MS

Rovshan G. Sadygov[1] (rovshan.sadygov@utmb.edu), Henock M. Deberneh[1]
[1]*University of Texas Medical Branch, Galveston, TX*

Protein turnover is an essential biological process that maintains the dynamic equilibrium of the cellular proteome. Metabolic labeling with a stable-isotope enriched diet followed by liquid chromatography coupled to mass spectrometry (LC-MS) has been a powerful tool to study the turnover of individual proteins *in vivo*. As a labeling agent, heavy water is cost-effective, easy to use, and

safe (in low doses). It labels all non-essential amino acids, lipids, and nucleic acids. The labeling with deuterium in heavy water results in a composite profile of labeled and unlabeled forms of a peptide. Traditional methods for the data processing required complete isotope profiles to characterize the label incorporation. Due to the complexity of the mammalian proteome, target peptides often co-elute with contaminants, even after powerful chromatographic separations. The co-elution results in distorted isotope profiles and reduces the accuracy of the estimation of label incorporation, retention time alignment, and the number of accessible hydrogens. We apply equations of dynamics of the isotope profiles during the label enrichment for data processing. The applications to real data show that the use of partial isotope profiles is less error-prone to contaminations from co-eluting species, as accurate measurements of abundances of only two mass isotopomers are sufficient for data processing. The resolution of the proteome complexity helps in the retention time alignment, label enrichment, and the number of accessible deuterium sites. It increases the proteome coverage in the labeling time domain and the number of quantified peptides. The work presents large-scale analyses of the protein turnover from murine liver, kidney, muscle, and heart tissues.

P8.13 | Amino acid sequence assignment from single molecule peptide sequencing data using a two-stage classifier

Matthew B. Smith[1] (mbsmith93@utexas.edu), Edward Marcotte[1], Zack Simpson[2]
[1]The University of Texas at Austin, Austin, TX, [2]Erisyon, Inc., Austin, TX

Tools for protein identification and quantification lag DNA and RNA sequencing techniques in sensitivity and throughput. To address this, our group invented fluorosequencing, a single molecule protein sequencing technology. 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 a significant elevation in noise and errors that must be addressed in subsequent computational analysis. We found that Hidden Markov Models representing the state changes of a peptide undergoing sequencing can provide excellent measures of the probability of a fluorosequencing read given that peptide, which we can in turn use for Bayesian classification. Naïve models did not scale to larger peptides with more labels, so we developed a number of novel algorithmic adjustments to our Hidden Markov Model implementation catered to address fluorosequencing data. Additionally we combined our brute-force Bayesian classifier with a k-Nearest-Neighbors classifier that reduces the number of Hidden Markov Models needed to be built and run.

P8.14 | PaSER 2024: a future-proofed renovation for better robustness, ease of use, flexibility, and scalability

Sebastian Wehner[1], Lennard Pfenning[1], Marc-Antoine Beauvais[2], Paolo Ballesteros[2], José Aubin[1], Duy M. Nguyen[2], Dennis Trede[1], Christoper Adams[3], Jonathan R. Krieger[4], **Tharan Srikumar[5] (tharan.srikumar@bruker.com)**, Sven Brehmer[1]
[1]Bruker Daltonics GmbH & Co. KG, Bremen, Bremen, Germany, [2]Bruker Scientific LLC, Billerica, MA, [3]Bruker Scientific LLC, San Jose, CA, [4]Bruker Canada, Toronto, ON, [5]Bruker Canada Ltd., Oshawa, ON

Parallel search engine in real-time or PaSER was developed together with the Yates lab to take advantage of GPU-powered database search. Since its initial conception, PaSER has been transforming into a proteomics data analysis platform that can integrate 3rd party tools while utilizing the concept of stream and stream-processors to realize fully customizable real-time processing workflows including on-the fly decision making based on the data generated. In order to overcome limitations regarding scalability (vertical and horizontal), optimal resource usage (CPU and GPU) and robustness a different system design was required. PaSER has now been re-written to encompass 4 fundamental principles: robustness, ease of use, flexibility, and scalability. The new PaSER platform utilizes an open-source, container first, cloud native technology stack that is well supported by the community and technology giants alike, providing a stable base for future growth. Using this platform, we can illustrate the optimal usage of system resources that are automatically balanced depending on the workload. We have also been able to illustrate the robustness of the system by not only processing a large sample cohort (>500 samples) but by utilizing chaos engineering for targeted disruption (or elimination) of specific workflow components during processing to ensure the system recovers and completes the tasks at hand. To illustrate the ease of use of the system, the new User Interface (UI), was provided to 5 proteomics users who were asked to carry out a list of 10 common tasks without any further instructions. The majority of tasks were completed by these users and provide further insight into the User Experience (UX) improvements. PaSER 2024 provides users with a proteomics analysis platform that is robust, easy to use, flexible, and scalable, for users as well as developers, making it an extensible\interrogatable system.

P8.15 | Realistic in silico generation and augmentation of mass spectrometry based proteomics data using generative adversarial networks

Vartika Tewari[1] (tewari.v@northeastern.edu), Sangtae Kim[2], Olga Vitek[1], Sunghyun Huh[3]
[1]Northeastern University, Boston, MA, [2]Bertis Bioscience, San Diego, CA, [3]Bertis R&D Division, Bertis Inc., Seongnamsi, South Korea

A fundamental problem in quantitative mass spectrometry-based proteomics is the low number of observations available, mostly due to a lack of available bioreplicates, prohibitive costs, or ethical reasons. Augmenting few real observations with generated in silico samples could lead to more robust analysis results and a higher reproducibility rate. Generating synthetic data representative of a particular experiment with desired structure can also help in experiment planning and analysis.
Neural networks (NNs) are a computational learning system that use a network of functions to understand and translate a data input of one form into a desired output in another form. Generative Adversarial Networks (GANs) are a class of NNs that learns to transform a simple, low-dimensional distribution into a high-dimensional distribution that is virtually indistinguishable from the real training distribution. GANs have been used to successfully generate high-quality artificial genomes, MRI scans, and electronic health records while mitigating privacy concerns. While data augmentation has been successfully applied in various fields of computer science, the development and usage of GANs for the augmentation of mass spectrometry proteomics data has yet to be investigated.
In this work, we examine data-independent acquisition and selected reaction monitoring datasets of serum samples from ovarian cancer patients and healthy controls. We use them to build and evaluate our GAN model to generate realistic synthetic mass spectrometry-based proteomic datasets. We further examine its ability to produce a specific subset of proteins/peptides on-demand with desired structure (such as depth of proteome coverage) while capturing correlations and generating any number of biological replicates. We show how these approaches can successfully augment sparse data sampled from populations to improve the quality and robustness of downstream conclusions.

P8.16 | Topic Modeling the Human Plasma Proteome: An Unsupervised Learning Method for Proteomic Analysis

Blake Williams[1] (bwilliams@somalogic.com), Robert K. DeLisle[1]
[1]SomaLogic, Boulder, CO

Introduction: The human plasma proteome is a complex system with dynamic protein levels, protein-protein interactions, and converging biological pathways. With the SomaScan® Assay we measure more than 7,000 proteins simultaneously, enabling advanced analytic methods. Here we share a proteomic application of topic modeling, an unsupervised machine learning technique appropriate for interpreting the rich data produced by SomaScan®.
Methods: Using the topic modeling technique, fit with non-negative matrix factorization, we reduced the proteome to a space of interpretable dimensions called topics, pooling information across all measured proteins into coherent biological signals. Topics were associated with relevant clinical traits, annotated through enrichment analyses, and used to cluster samples.
Results: Using topic models, we were able to identify biologically coherent topics in SomaScan® data, corresponding to prominent dimensions of the proteome. Hierarchical clustering patients on membership to these topics identified possible subtypes of diseases. Pathway analysis of prominent disease-associated topics identified relevant pathways and both known and potential biomarkers of disease.
Conclusion: These results show that advanced machine learning techniques may be applied successfully to SomaScan® measurements to explain the relationship between diseases and the proteome. Topic models enabled meaningful clustering of samples to suggest proteomic subtypes of diseases and relevant biomarkers. Topics identified by these models not only explain the differences between samples but represent prominent and functionally enriched dimensions in the proteome. These results demonstrate the wealth of information that may be gleaned from SomaScan® measurements through advanced modeling techniques.

P8.17 | QuickProteomics pipeline for mass-spectrometry based quantitative proteomics

Liang Xue[1] (liang.xue@pfizer.com), Robert Stanton[1]
[1]Pfizer, Cambridge, MA

The high diversity of mass spectrometry-based proteomics workflows in the pharmaceutical industry makes it challenging to assure reproducibility. Here we present an on-going pipeline development effort in our organization which aims to identify best practices for analysis of multiple types of proteomics datasets.
To optimize the pipeline, mass spectrometry data from public domain repositories was used for pipeline development. Later the

established pipeline was tested against internal data to demonstrate its applicability for real applications. Data generated from different instruments (Thermo orbitrap series and Bruker TimsTOF Pro) and data processing workflows (Label-free and Isobaric labelling) were analyzed using several popular software suites (Proteome Discoverer, Maxquant, IP2 and MSfragger). A cloud based statistical analysis framework was built upon MSstats in Seven Bridges Genomics platform, and further developed to fit within an in-house proteomics ecosystem. Pipeline reproducibility, documentation of processing parameters, data management, as well as user experience were considered as part of an evaluation matrix.
By providing an end-to-end solution to practitioners, the workflow has proved popular across the Pfizer research organization. Incorporating the pipeline with careful registration of study design and experimental details, as well as end-point data management not only improves the data lineage, but also the efficiency of pipeline automation. Additionally, it simplifies consumption and interpretation of the data in downstream analysis.

P8.18 | MIND-S: A Novel Deep Learning Prediction Model For Elucidating Protein PTMs In Human Diseases

Yu Yan[1] (yuyan666@g.ucla.edu), Peipei Ping[1], Wei Wang[1], Mingzhou Fu[1], Alexander Pelletier[1], Dibakar Sigdel[2]
[1]UCLA, Los Angeles, CA, [2]Truveta, Seattle

We report a novel AI platform, MIND-S, for protein PTM predictions. MIND-S is supported by multi-head attention and GNN. It utilizes multilabel strategy and ensembles 15-fold models from bootstrapped samples to enable simultaneous prediction of multiple PTMs with high performance and computation efficiency. MIND-S also features an interpret module, which provides the saliency of each amino acid for making the predictions and are validated with known motifs. The interpret module also captured PTM patterns without any supervision. Furthermore, MIND-S enables examination of mutation effects on PTMs. We document a workflow, its applications to 26 types of PTMs of two datasets consisting of ~50,000 proteins, and an example of MIND-S identifying a PTM-interrupting SNP with validation of biological data. We also include use case analyses of targeted proteins. Taken together, we have demonstrated that MIND-S is accurate, interpretable, and efficient to elucidate PTM-relevant biological processes in health and diseases. Highlights for this work:

- We developed a deep learning-based tool, MIND-S, consisting of multi-head attention and graph neural network (GNN) for protein PTM site prediction, enabling information integration from both protein sequence (1-dimensional) and structure (3-dimensional).
- MIND-S applies multi-label and bootstrapping strategies to achieve predictions on multiple PTM types and sites simultaneously with high accuracy and efficiency.
- MIND-S offers an interpretation approach for identifying important amino acids for PTM predictions at proteome-scale.
- MIND-S enables the examination of mutation effects from a PTM perspective to better understand disease pathology. We demonstrated the utility of MIND-S through use cases.

P8.19 | Increasing the HLA Repertoire Across Fragmentation Modes with MSBooster's Deep Learning PSM Rescoring

Kevin L. Yang[1] (yangkl@umich.edu), Alexey I. Nesvizhskii[1]
[1]University of Michigan, Ann Arbor, MI

T cell recognition of major histocompatibility complex (MHC) bound peptides on antigen presenting cells plays a key role in the immune response to bacterial and viral infections, as well as cancer recognition. Mass spectrometry-based proteomics has uncovered many of these peptide ligands, at least those with fragmentation patterns consistent with the theoretical backbones considered by database searching algorithms. In order to increase the number of identified human leukocyte antigen (HLA) peptides, we present several updates to the MSBooster deep learning component of the FragPipe computational proteomics workflow. First, we provide support for predicted spectral libraries from PredFull, which we show can predict immonium, internal, and neutral loss fragment ions. PSM rescoring with DIA-NN or Prosit's predicted y/b intensities along with PredFull's non-y/b intensities provides a >10% boost on top of the original MSBooster algorithm, which already showed a 30% boost in peptide numbers. The gentler methods of ETD/EThcD fragmentation of HLA peptides are another way to increase the known HLA peptide repertoire, especially those with labile post-translational modifications (PTMs). To generalize MSBooster to these two fragmentation methods, we performed transfer learning on PredFull and alphapeptdeep, a new transformer based MS2 spectral prediction model that supports many PTMs. We show a consistent increase in reported peptides across multiple HLA class I/II datasets. These new prediction models will be integrated into FragPipe to enable improved HLA peptide discovery in HCD, ETD, and EThcD workflows.

P9: Data Analysis and Visualization

P9.01 | Data to Information to Knowledge Effortlessly: timsTOF data with PaSER information and Mass Dynamics knowledge accelerates proteomic discoveries

Christoper Adams[1] (christopher.adams@bruker.com), Giuseppe Infusini[2], Brad Green[2], Aaron Triantafyllidis[2], Paula Burton[2], Sebastian Wehner[3], Lennard Pfenning[3], Marc-Antoine Beauvais[4], Paolo Ballesteros[4], Dennis Trede[3], Jonathan R. Krieger[5], Tharan Srikumar[6], Sven Brehmer[3], Andrew Webb[2]
[1]Bruker Scientific LLC, San Jose, CA, [2]Mass Dynamics, Melbourne, Victoria, Australia, [3]Bruker Daltonics GmbH & Co. KG, Bremen, Bremen, Germany, [4]Bruker Scientific LLC, Billerica, MA, [5]Bruker Canada, Toronto, ON, [6]Bruker Canada Ltd., Oshawa, ON

Mass spectrometers are generating data at an unprecedented rate, so much so that the bioinformatics of 'data to information' and then 'information to knowledge' is now the primary bottleneck in large cohort proteomic experiments. Data transfer and analysis by the most rigorous pipelines followed by statistical analysis, normalization and visualization adds many days if not weeks to complex dataset evaluation. In 2020, PaSER, developed together with the Yates lab, was launched to take advantage of GPU-powered database searching. Since its introduction, PaSER has been transforming into a proteomics data analysis platform that integrates 3rd party tools, while utilizing the concept of stream and stream-processors to realize real-time processing workflows. PaSER bridges the 'data to information' gap by including on-the-fly decision-making based on the data generated. In this work, we now complete the 'information to knowledge' gap by leveraging Mass Dynamics - a seamless and modular web-based and software environment that simplifies and templates complex proteomics analysis. It automates data normalization, statistical analyses, and integration of external and existing knowledge bases (e.g. Reactome) in the context of the experiment. Results are presented in a manner that allows more effective collaboration between proteomics experts and biologists, facilitating rapid insight generation and decision-making from proteomics datasets. Taken together, streaming data from a timsTOF mass spectrometer to a PaSER box, where on-the-fly information generation is performed, followed by a cloud-based platform accepting pre-processed data, results in biological knowledge in a seamless and expedited fashion. This accelerates research teams' abilities to extract interpretable information and maximize what their mass spectrometry datasets can deliver. This will be demonstrated using dia-PASEF data on complex biological samples including multispecies proteomes.

P9.02 | Mass Dynamics 2.0: A modular web-based platform for accelerated insight generation and decision-making for proteomics

Mark R. Condina[1] (mark@massdynamics.com), Giuseppe Infusini[1], Andrew Webb[1], Paula Burton[1], Aaron Triantafyllidis[1], Brad Green[1], Anna Quaglieri[1], Joseph Bloom[1]
[1]Mass Dynamics, Melbourne, Victoria, Australia

As the field of Proteomics matures and becomes more accessible to a broader scientific community, it is essential to ensure the reproducibility, quality and integrity of data analysis and its interpretation. The complexity of the data and the capabilities of a research team to adopt complex analysis pipelines have proven to be an obstacle to effective collaboration and more efficient biological insight generation. Launched in December 2022, MD 2.0 is a cloud- and web-based platform for quantitative proteomics data, which implements a novel analysis workspace where raw data processing, statistical analyses, visualizations, and external knowledge generation are integrated into a modular fashion. Drag-and-drop modules allow a user to easily and quickly assess different aspects of an uploaded experiment, including quality control, differential expression and enrichment analysis. This modularity enables researchers the flexibility to test different hypotheses and to customize and template complex proteomics analysis. This ability, coupled with a human-centred interface design, reduces the barrier to proteomics data analysis without compromising quality and depth and expedites insight generation for complex datasets. The extensible MD 2.0 environment has been built with a scalable architecture to allow rapid development of future analysis modules as well as enhanced tools for remote collaboration.

Since its launch, MD 2.0 has been further developed to offer new upload options for a range of raw and pre-processed LC-MS data, for seamless integration into existing processing workflows. The modularity of MD 2.0 continues to lay the foundation to support broader community-based template generation and optimized collaboration between proteomics experts and biologists, thereby accelerating research teams' abilities to extract knowledge from complex proteomics datasets. Case studies on published proteomics cohort analyses will be presented to highlight MD 2.0's ability to facilitate insight generation, collaboration and sharing.

P9.03 | Protein Contaminants Matter: Building Universal Protein Contaminant Libraries for DDA and DIA Proteomics

Ashley Frankenfield[1], Jiawei Ni[1], Mustafa Ahmed[1], **Ling Hao[1] (linghao@gwu.edu)**
[1]The George Washington University, Washington, DC

Mass spectrometry-based proteomics is constantly challenged by the presence of contaminant background signals. In particular, protein contaminants from reagents and sample handling are almost impossible to avoid. For data-dependent acquisition (DDA) proteomics, an exclusion list can be used to reduce the influence of protein contaminants. However, protein contamination has not been evaluated and is rarely addressed in data-independent acquisition (DIA). How protein contaminants influence proteomic data is also unclear. In this study, we established new protein contaminant FASTA and spectral libraries that are applicable to all proteomic workflows and evaluated the impact of protein contaminants on both DDA and DIA proteomics. We demonstrated that including our contaminant libraries can reduce false discoveries and increase protein identifications, without influencing the quantification accuracy in various proteomic software platforms. With the pressing need to standardize proteomic workflow in the research community, we highly recommend including our contaminant FASTA and spectral libraries in all bottom-up proteomic data analysis. Our contaminant libraries and a step-by-step tutorial to incorporate these libraries in various DDA and DIA data analysis platforms can be valuable resources for proteomic researchers, freely accessible at https://github.com/HaoGroup-ProtContLib.

P9.04 | Veneer is a web-based bioinformatic tool to address the need for standardized processing and reporting of mammalian cell surface glycoproteome data

Roneldine Mesidor[1], Jack Littrell[1], Linda Berg Luecke[2], **Rebekah L. Gundry[1] (rebekah.gundry@unmc.edu)**
[1]University of Nebraska Medical Center, Omaha, NE, [2]Medical College of Wisconsin, Milwaukee, WI

Defining the cell surface proteome is a critical aspect of studies of mammalian development, physiology, disease, drug discovery, and regenerative medicine. IDefining the cell surface proteome is a critical aspect of studies of mammalian development, physiology, disease, drug discovery, and regenerative medicine. In the past twenty years, >440 published studies have applied a variety of MS-based proteomic approaches to profile mammalian cell surface proteomes. To date, there is no consensus regarding the criteria required to designate a protein in a proteomic dataset as a cell surface protein, and most studies do not provide experimental evidence of location but instead rely on gene ontology or computational predictions. Thus, approaches for standardized analysis and reporting of cell surface proteome datasets would benefit comparisons among studies and methods, provide added confidence to assertions of localization, and promote the use of proteomic data by other scientists. Towards that goal, we developed a new web-based bioinformatic tool, Veneer, for analyzing data generated by glycopeptide capture, currently the most popular cell surface proteomics method that generates experimental evidence of protein subcellular location. Veneer is agnostic of MS instrument or search algorithm and applies stringent criteria for classification of identified proteins as cell surface N-glycoproteins and expedites data analysis and annotation. Veneer analysis of >100 published glycopeptide capture datasets reveals: 1) sub-standard enrichment specificity is observed in >40% of cell surface proteomics studies; 2) Veneer facilitates the extraction of experimental data useful for updating public protein repositories for >300 cell surface proteins that are currently incorrectly or incompletely annotated; 3) applying stringent criteria for surface protein data analysis reveals up to 40% of proteins reported as cell surface proteins in previous proteomic studies are lacking experimental evidence to meet this criteria. These analyses demonstrate the need for more stringent criteria for analyzing and reporting cell surface proteomics

P9.05 | In Vivo Protein Turnover Rates across the Proteome for Native Mouse and Engineered Human Tissues

Brian Frey[1] (bfrey@chem.wisc.edu), Zach Rolfs[1], Xudong Shi[1], Yoshitaka Kawai[1], Michael R. Shortreed[1], Lloyd M. Smith[1], Nathan V. Welham[1]
[1]University of Wisconsin-Madison, Madison, WI

Protein turnover is critical to cellular physiology as well as to the growth and maintenance of tissues. The unique synthesis and degradation rates of each protein help to define tissue phenotype, and knowledge of tissue- and protein-specific half-lives is directly relevant to protein-related drug development as well as the administration of medical therapies. After generating a cohort of isotopically heavy mice using $^{13}C_6$,$^{15}N_2$-lysine (Lys8) food for three generations, we devised two large experiments to measure protein turnover. In one experiment, we switched back to a regular diet (Lys0) and then sacrificed mice (n=22) after 3, 7, 14, 30, and 60 days. Proteomic measurements of heavy and light peptide abundances allowed calculation of *in vivo* turnover rates for thousands of proteins—including those of the extracellular matrix—in a set of biologically important mouse tissues. The other experiment involved surgically implanting engineered human tissues into the heavy mice (n=25) for 12, 24, 36, 48, or 60 days before surgical removal. The engineered tissues initially contained Lys0 proteins that degraded over time, while the human cells in these grafts

gradually synthesized human proteins with Lys8 (nutrient source in the heavy mice). Proteomic analysis of the grafts enabled protein-specific turnover measurements, increasing understanding of the fate of engineered human tissues during implantation. We additionally developed a freely-available, open-source data visualization platform, named ApplE Turnover, that enables facile searching for any protein or peptide of interest and then displays its half-life, confidence interval, and supporting measurements. These extensive datasets and the corresponding visualization software provide guidance to myriad future studies of mammalian protein turnover.

P9.06 | Prognostic Proteomic Models for Low Event Rates: A Case Study with Myocardial Infarction

Yolanda Hagar[1] **(yhagar@somalogic.com)**, Leigh Alexander[1], Jessica Chadwick[1], Gargi Datta[1], Michael Hinterberg[1]
[1]*SomaLogic, Inc., Boulder, CO*

Introduction: Large-scale clinical proteomics provide increasing opportunities for patient risk stratification, especially with multi-marker models derived using machine learning techniques. Prognostic models can be developed as binary risk classifiers, or by using time-to-event data. Survival modeling is ubiquitous in statistical literature, but support for machine learning optimization is more limited in comparison to other regression techniques. We have developed and assessed a novel prognostic model development method combining two statistical techniques -- survival analysis and subsampling – using existing machine learning tools in R. These methods were applied to a clinical dataset to identify a highly predictive proteomic model for myocardial infarction (MI) despite a low observed event rate.
Methods: Cox elastic net with subsampling tools were developed in R. Simulations were used to demonstrate the utility and accuracy of subsampling in a survival data context, with comparisons made to logistic regression, Cox elastic net, and SVM models. Following the validation of the approach via simulations, models were developed and assessed on the HUNT3 data set (n = 756), which had 61 (8.1%) MI events within four years of blood draw. Proteomic measurements were performed using SomaScan v4.0 technology.
Results: The subsampled survival model performance was superior to competing logistic regression, Cox elastic net, and SVM models (AUC=0.82 vs. 0.75, 0.61, 0.75, respectively). Sensitivity and specificity metrics were more balanced on hold-out test sets. Additionally, simulations showed that the proteins that were most highly correlated with MI were selected for final models, indicating that this method is a promising tool for clinical discovery and prognostic/diagnostic development.
Conclusions: Survival analysis with subsampling can be applied to proteomic data to discover novel individual protein associations with outcome, as well as develop prognostic models with superior performance to traditional statistical models, even in the context of relatively low incidence rates.

P9.07 | Efficient Development of Certified Prognostic Laboratory Developed Tests Using Proteomic Data

Yolanda Hagar[1] **(yhagar@somalogic.com)**, Michael Hinterberg[1], Leigh Alexander[1], Caleb Scheidel[1], Amy Zhang[1], Joe Gogain[1], Clare Paterson[1], Rachel Ostroff[1]
[1]*SomaLogic, Inc., Boulder, CO*

Introduction: Proteomic technology is a powerful biological tool with established methods for identifying proteomic biomarkers, but the development of certified prognostic clinical tests based on proteomic biomarkers can be time-consuming, prone to overfitting issues, and difficult to navigate. We demonstrate the utility of combining pipeline tools, statistical learning techniques, and a knowledge base of in-silico proteomic datasets into a reproducible workflow that allows for efficient development of LDT-certifiable tests using SomaScan technology.
Methods: Data pipeline and analysis tools were developed using R, in conjunction with proteomic measurements obtained using the SomaLogic platform. The tools take the analyst from data processing and QC through identification of optimized models for prediction of clinical endpoints, and then through validation on a hold-out test set. The tool performs an assessment of model robustness against sample handling issues, the impacts of assay noise on model performance, and risk of failure during CLIA validation in the lab. Real-life examples of clinical applications demonstrate the effectiveness of the tool in reducing analysis time and increasing model accuracy.
Results: Analysis time for identifying the optimal clinical model to validation was reduced by at least 80%, with decreased prediction variability by up to 90%. In at least 75% of cases, application of in-silico data allows for tuning of predictive models to ensure robustness in a variety of everyday settings. This tool has led to 16 LDT certified SomaLogic tests in the last 3 years, ranging from anthropometric measurements to cardiovascular- and cancer-risk predictions.
Conclusions: Not only are powerful, proteomics-driven, prognostic tests realizable, but they can be LDT certified in an efficient manner and made to be robust to real-life variability. Efficient analysis tools allow us to leverage proteomic technology in new ways, leading to tests that can be used for precision medicine applications.

P9.08 | Increasing the depth of single shot proteomics with enhanced data acquisition and processing strategies

Daniel Hermanson (daniel.hermanson@thermofisher.com), Thermo Fisher Scientific, San Marcos, CA

Advances in the online separation of complex proteomics samples including ultra-high performance liquid chromatography (UHPLC), ultra-high resolution separation columns, and high-field asymmetric waveform ion mobility spectrometry (FAIMS) enable a deeper mining of the proteome with single-shot methods. In addition to the improved separation of complex proteomes with single-shot methods described above, the CHIMERYS™ intelligent search algorithm unlocks the ability to deconvolute the chimeric spectra that still arise from the co-isolation and fragmentation of multiple peptides in tandem mass spectra from both Orbitrap and ion trap mass analyzers. To optimize the acquisition strategy for single-shot performance, 1 µg HeLa digest and a 1-hour gradient were used with isolation widths between 0.4 and 3 Th. Using an isolation width of 1.5 Th resulted in an average of 7,814 proteins, 59,213 peptides, and 128,647 peptide spectrum matches (PSMs) per run, increases of 13%, 31%, and 29% compared to a 0.4 Th isolation window, respectively. The use of a wider isolation window was enabled by the deconvolution capabilities of the CHIMERYS workflow, which provided 1.02 PSMs per spectra on average with the 1.5 Th isolation window versus 0.79 PSMs per spectra for the 0.4 Th isolation window. In contrast, processing the 1.5 Th isolation window results using a Sequest HT workflow provided an average of only 6,511 proteins, 39,549 peptides, and 53,289 PSMs. Thus, CHIMERYS provided improvements of 20% for proteins, 50% for peptides, and 141% for PSMs. These results demonstrate that coupling advanced capabilities in online separation to enhanced data acquisition and intelligent data processing allows for substantial improvements in single-shot proteomics performance to yield a more thorough coverage of biological pathways with higher throughput and minimal offline sample preparation.

P9.09 | Bioinformatics Characterization of SARS-CoV-2 Interactome reveal multiple insights into Potential Drug Discovery

Shubhdeep Kaur[1] (shubhdeep1456@gmail.com), Gaganjot Singh[2]
[1]Delhi University, Delhi, Delhi, India

The SARS-CoV-2-caused COVID-19 pandemic posed a significant threat to humanity. During infection, virus replication necessitates the presence of host cellular factors. By identifying virus-host protein-protein interactions (PPIs), we can better understand the severity of COVID-19 infection. In the study, 2812 protein-protein interactions were found between SARS-CoV-2 and human proteins. Bioinformatics and structural analysis have been used to identify and characterize the interactome based on physiochemical parameters such as molecular weight, primary sequence length and globularity. The results show that SARS-CoV-2 proteins are acidic, low molecular weight (below 50kDa), highly globular in structure while SARS-CoV-2 associated proteins are acidic and low molecular weight (below 50kDa). In addition, gene ontology analysis revealed cellular components, molecular functions and biological processes. A pathway enrichment analysis of virus-interacting proteins (KEGG and Reactome) identifies 14 potentially impacted pathways and 90 potential virus-targeted host proteins. Using drug–target–pathway–disease networks, we investigated the molecular details of the disease as well as potential new therapeutic candidate drugs. A total of 25 druggable human proteins were found to be targeted by 370 approved drugs using DGIdb. The STITCH database validated our findings and identified SUNITINIB, SORAFENIB, IMATINIB, PAZOPANIB, DASATINIB, and LAPATINIB as potentially novel drugs. Further, a therapeutic treatment for COVID-19 could be developed through various tests and experiments.

P9.10 | Machine Learning to the Rescue: A Software Suite for Real-Time Mass Spectrometry Data Analysis

Yun-En Chung[1], Giorgio Freije[1], Alona Petrova[1], Chloé Lavoie[1], Matthew Willetts[2], Jens Decker[3], Nagarjuna Nagaraj[3], Jonathan R. Krieger[4], Tharan Srikumar[5], **Mathieu Lavallée-Adam[6] (mlavall3@uottawa.ca)**
[1]University of Ottawa, Ottawa, ON, [2]Bruker, Billerica, MA, [3]Bruker Daltonics, Bremen, Bremen, Germany, [4]Bruker Canada, Toronto, ON, [5]Bruker Canada Ltd., Oshawa, ON, [6]University of Ottawa, Ottawa, QC

The ability to identify peptides and proteins using mass spectrometry is linked to data acquisition methods and experimental parameter selections. Such parameter choices involve the use of precursor ion exclusion lists and construction method thereof, as well as collision energy level used for peptide ion fragmentation. In most mass spectrometry experiments, these parameters are set prior to the start of the experiment and are constantly and uniformly applied to all peptide ions, irrespective of their characteristics. Such an experimental setting is used despite the fact that numerous peptide ions would optimally fragment when using a different

collision energy level than the predetermined one. Furthermore, in typical experiments, the mass spectrometer redundantly collects mass spectra for proteins detected several times during the experiment or that post-data acquisition analysis software cannot identify.
Herein, we propose a suite of machine learning algorithms that analyze mass spectrometry data in real-time to maximize peptide and protein identification capabilities. Specifically, we present an artificial neural network using precursor ion properties such as m/z, charge state and ion mobility coefficient to determine the collision energy that will optimize the fragmentation of a given precursor ion. When applied on a Bruker timsTOF Pro mass spectrometer, this method increased mass spectra identification by 5%. We also introduce a supervised learning algorithm that predicts, solely using precursor ion properties, whether such an ion is likely to generate a confident peptide identification (AUC=0.7). Finally, we show that by identifying proteins in real-time using our software MealTime-MS, we can prevent the acquisition of redundant mass spectra from previously identified proteins in the experiment and favor data acquisition from peptides of unidentified proteins.
Together, this suite of machine learning-based software will optimize mass spectrometry data acquisition in real-time to improve proteome characterization and therefore provide a better understanding of underlying biological processes.

P9.11 | A Cloud-Scalable Software Suite for Large-Scale Proteogenomics Data Analysis and Visualization

Taylor Page[1] (tpage@seer.bio), Harsharn Auluck[1], Margaret Donovan[1], Aaron S. Gajadhar[1], Yuandan Lou[1], Theo Platt[1], Serafim Batzoglou[1]
[1]Seer, Inc., Redwood City, CA

Introduction: Integration of large-scale proteomics and genomics data requires complex workflows spanning multiple data types and requires expertise across scientific domains, which may act as a barrier for researchers to perform multi-omics data analyses. The Proteograph™ Analysis Suite (PAS), a cloud-based, proteomics and proteogenomics data analysis software, enables the identification and exploration of proteins and peptide variants arising from allelic variation or other user-defined protein sequence altering genetic variants not captured in standard canonical reference databases through the integration of Proteograph proteomics data with NGS variant information.
Methods: PAS features integration of stablished proteomics database search algorithms, experiment data management system, analysis protocols, result visualizations, and setup wizards for seamless generation of results. PAS can support both Data Independent Acquisition and Data Dependent Acquisition workflows and is compatible with variant call format files. Using the Proteogenomics module in PAS, we performed a customized, sample-specific database search to identify variant peptides, including results interpretation using the Variant Peptide Browser and Proteogenomics Data Explorer.
Results: We demonstrate the Proteogenomic features of PAS by analyzing plasma proteome data generated from Proteograph™ Product Suite and corresponding variant data from healthy controls and Alzheimer's Disease (AD) patients. We highlight an interactive Variant Peptide Browser tool for examining peptide variation in the samples. These results include filtering and grouping features to identify qualitative and quantitative trends in single samples or healthy/disease groups. Further, we demonstrate the Proteogenomics Data Explorer tool, which provides a view of the genomic variant coordinates in relationship to the peptide/protein data. Detected variants can be reviewed in the context of gene structure, protein structure, and functional domains so biological insight can easily be gained from the proteogenomic data.
Conclusions: PAS provides an easy-to-use and efficient suite of tools for seamless and fast proteomics and proteogenomics data analysis enabling novel

P9.12 | Spectronaut 17's Novel directDIA+ Provides More Protein Identifications Than a Deep Project-Specific Library in a Plasma Prostate Cancer Study.

Roland Bruderer[1], Raphael Heilig[2], Marco Tognetti[2], Yuehan Feng[2], **Monika J. Puchalska[3] (monika.puchalska@biognosys.com)**, Lukas Reiter[1]
[1]Biognosys AG, Schlieren, Shiselweni, Swaziland, [2]Biognosys, Schlieren, Zurich, Switzerland, [3]Biognosys, Schlieren, Switzerland

Introduction: Biomarker discovery from plasma samples is challenging due to the high dynamic range and inter-individual variability of the proteome. This challenge was addressed by performing deep DIA analysis with optimized project-specific libraries, requiring substantial offline sample fractionation, acquisition, and computing times. This study presents the results of a novel directDIA+ analysis (Spectronaut 17, Biognosys) and an optimized project-specific library-based Spectronaut 16's DIA analysis applied to the plasma samples from prostate cancer patients and healthy individuals. The depth of plasma proteome coverage obtained with a new directDIA+ eliminates the need for a dedicated, deep library in plasma proteomics.
Methods: Plasma samples were depleted using a MARS-14 system (Agilent) automated with a UHPLC. Subsequently, samples underwent tryptic digestion, followed by C18 material cleanup. A FAIMS Pro device was connected to a Thermo Scientific Orbitrap Exploris 480, coupled to EASY-nLC 1200. Optimized, 3.5h FAIMS-DIA methods were generated. A DDA-based project-specific library was generated using HPRP fractionation (12 fractions). The DDA library was generated with Spectronaut 16.
Results: The DIA data of 15 prostate cancer patients and 15 matched healthy individuals were analyzed using Spectronaut 16 with

a project-specific library (85,000 precursors, 4200 proteins) and with a directDIA workflow. The first approach resulted in the identification of over 3500 protein groups in total with an average of 36,000 precursors per sample, whereas directDIA allowed the identification of 2700 protein groups (32,000 precursors per sample).

In contrast, Spectronaut 17's directDIA+ algorithm led to the identification of over 35,000 precursors per sample and over 3600 protein groups in total, exceeding plasma proteome coverage depth obtained with the deep project-specific library on the protein group level.
Conclusion: The novel directDIA+ algorithm preserves biological information obtained from analyzed samples, enabling plasma proteome discovery while reducing project costs and turnover time connected to the optimized library.

P9.13 | Integrating DIA-NN software analysis of SWATH DIA data into a cloud processing pipeline

Melanie Juba[1] (melanie.juba@sciex.com), Nick Morrice[2], Christie Hunter[3]
[1]SCIEX, Framingham, MA, [2]SCIEX, Nether Alderley, Cheshire, United Kingdom, [3]SCIEX, Redwood City, CA

Combined omics disciplines have proven to be more powerful than individual disciplines. However, multi-omics data analysis workflows can be time-consuming, as consolidating and interpreting various results outputs – especially when using different scoring schemas and criteria – is often challenging. The OneOmics suite was previously demonstrated to provide researchers tools for multi-omics data management, compound identification and quantification, statistical analysis and pathway analysis to streamline biomarker discovery studies. To further facilitate data processing for high-throughput proteomics workflows, the OneOmics suite extends its SWATH DIA (data independent acquisition) data processing by supporting visualization and statistical interpretation of DIA-NN software results. DIA-NN software is a widely used proteomics processing platform that leverages neural networks and powerful quantification and inference algorithms to achieve confident protein and peptide identifications. Here, a processing workflow for proteomics analysis is presented using Zeno DDA (data-dependent acquisition) and SWATH DIA data. In this workflow, a spectral library was generated from Zeno DDA data using the Ion Library app in OneOmics suite. SWATH DIA data were subsequently processed using this ion library in DIA-NN software and evaluated in the DIA Results app. Results generated using this workflow can be rapidly and securely shared with collaborators.

P9.14 | PIONEER: Harnessing multi-omics data to enhance immunotherapeutic target discovery and development

Amber K. Weiner[1] (weinera@chop.edu), Hemma Murali[2], Rawan Shraim[3], Karina L. Conkrite[3], Alexander B. Radaoui[3], Dan Martinez[3], Brian Mooney[4], Sandra Spencer Miko[4], Gian Negri[4], Alberto Delaidelli[4], Caitlyn de Jong[4], Yuankun Zhu[3], Allison P. Heath[3], Jennifer Pogoriler[3], Yael P. Mosse[3], Deanne M. Taylor[3], Poul H. Sorensen[4], Gregg B. Morin[4], Benjamin A. Garcia[5], John M. Maris[3], Sharon J. Diskin[3]
[1]Children's Hospital of Philadelphia, Warminster, PA, [2]University of Pennsylvania, Philadelphia, PA, [3]Children's Hospital of Philadelphia, Philadelphia, PA, [4]University of British Columbia, Vancouver, BC, [5]Washington University School of Medicine in St. Louis, St. Louis, MO

Introduction. Immunotherapeutic strategies have produced remarkable results in some malignancies. However, optimal cell surface targets in many childhood cancers remain elusive and tools for novel target discovery are limited. We developed a proteogenomic approach to identify high confidence cell surface proteins for immunotherapy development. Through the Pediatric Immunotherapy Discovery and Development Network (PI-DDN), we are applying this approach to 14 high-risk childhood cancers. This effort includes MS-based surfaceome data generation for 175 patient-derived xenograft (PDX) models, 30 primary patient tumors, and 10 human derived cell lines.
Methods/Results. To facilitate dissemination of data to the broader research community, we are developing a web-based application called PIONEER (Pediatric Integrative Omics Network Enhancing Early Research). PIONEER is designed to provide the necessary analysis, query and visualization tools to support target discovery and share immunotherapy development progress. A pilot version of PIONEER, a R Shiny application, was developed based on our neuroblastoma efforts. The application consists of two categories: (1) target discovery and prioritization (2) target validation and preclinical development. Modules include proteomics (MS), transcriptomics (RNA-sequencing), epigenomics (ChIP-sequencing), multi-omics, validation and pre-clinical drug development. Extensive normal tissue expression data from GTEx and mass spectrometry data will be integrated into the interface. Surface proteins are prioritized through an integrative multi-omic analysis of tumor and normal tissue data. We are currently adding functionality to support automatic data analysis and integration (SPACE-Surface Protein Analysis for Collaborative Efforts). Through the validation and preclinical development modules, users can view reagents, immunofluorescence, immunohistochemistry, drugs in development, and in-vivo efficacy. Data for additional histotypes will be incorporated as generated.
Conclusion. PIONEER is a web-based application for data integration, visualization and sharing. The interface provides a

comprehensive characterization of the surfaceome of high-risk pediatric cancers and facilitates the discovery of optimal immunotherapeutic drug targets.

P9.15 | Higher and higher N: SimpliFi for the masses now for takes masses of samples

Jim Palmeri[1], Darryl J. Pappin[2], **John Wilson[3] (john@protifi.com)**
[1]ProtiFi LLC, Deer Park, NY, [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, [3]ProtiFi LLC, Fairport, NY

Introduction: Ever-increasing amounts of omics data limit our ability to understand and make actionable decisions. Adding to the challenge, studies continue to grow both in the number of analyzed samples and in the kinds of omics analyses. Previously we developed SimpliFi, a first cloud-based, browser-accessible data-to-meaning engine that accepts all omics types and allows results to easily be shared, explored or published simply by sending a URL. We have now extensively upgraded and optimized both SimpliFi's front- and backend to accept hundreds to thousands of samples.
Methods: SimpliFi models biology using exclusively nonparametric 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. Confidence intervals are given 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 even for new-to-omics users. SimpliFi's backend analyses has been speed up by orders of magnitude and its frontend has been redone to allow for summary displays at all views enabling human understanding of the information present in large (>100s) numbers of samples.
Preliminary Data: SimpliFi uses non-parametric statistics where biological replicates yield their own empirical distributions. Using statistics based on these approaches, non-parametric p values often differ by several orders of magnitude compared to T-tests. This is a function first of the non-Gaussian nature of biology and omics data and general, combined with either oversampling of biological variability (with the presence of an outlier), yielding false negatives, or undersampling of variability (with tight clustering), yielding false positives. Non-parametric SimpliFi analyses of either mono- or integrated multiomics data are presented in clean interactive displays of pathways, states of tissues, disease, cells and molecular-level classifications.

P10: Glycoproteomics and Glycomics

P10.01 | A multi nanoparticle-based plasma proteomics workflow enabling investigation of the global glycoproteome

Moaraj Hasan[1] (mhasan@seer.bio), Shadi Ferdosi[1], Iman Mohtashemi[1], Evan O'Brian[1], Hongwei Xia[1], Jian Wang[1], Khatereh Motamedchaboki[1], Shao-Yung Chen[1], Harendra Guturu[1], Daniel Hornburg[1]
[1]*Seer, Redwood City, CA*

To overcome limitations of deep plasma proteomics in large cohorts, we have developed a fast and scalable technology that employs intricate protein-nano interactions. Introducing engineered nanoparticles (NPs) into a biofluid such as blood plasma can form 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. Since NP-coronas differentially interrogate complex samples at the proteoform level and plasma proteins are often glycosylated, here we investigate whether an automated NP-corona-based proteomics workflow can capture the plasma glycoproteome without subsequent downstream enrichment of glycosylated peptides. Additionally, we investigate the improvements in sensitivity, glycan localization and structure inference implementing biochemical and computational workflows upstream and downstream of NP-corona formation. Protein glycosylation states can provide diagnostic evidence where total protein abundance is uninformative but are not accessible in standardized large-scale deep plasma proteome studies. Improved methods for profiling the plasma glycoproteome, can thus potentially have a major impact on biomarker discovery.

We have investigated the utility of NPs in enrichment of low-abundant glycoproteins. Samples were analyzed with Orbitrap Lumos and timsTOF Pro mass spectrometry and UltiMate3000 Dionex LC system using 60min DDA and 30 min DIA runs. We evaluated depth, dynamic range, coverage, and precision of quantification at a wide range of concentrations for each NP. We deployed orthogonal chemistries of enriching glycopeptides and deglycosylation treatments, applying the same analysis to determine the relative performance of the NPs to established glycoproteomics methods.

In summary, NPs enable the detection of low-abundant proteins and corresponding peptides even without subsequent enrichment of glycopeptides. These data provide evidence that NP-protein coronascapture a substantial part of the glycoproteome without subsequent, glycopeptide-specific enrichment. Furthermore, the different NPs offer complementary views of the plasma glycoproteome due to their specificities for different proteins, and likely, different glycosylated proteoforms.

P10.02 | Mapping the Fragile X Syndrome N-Glycoproteome by Mass Spectrometry

Ashley Phetsanthad[1] (Phetsanthad@wisc.edu), Elliot Patrenets[1], Caroline Roycroft[1], Min Ma[1], Pamela R. Westmark[1], Cara J. Westmark[1], Lingjun Li[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Fragile X syndrome is the leading inherited cause of autism and can induce developmental delays and intellectual disabilities. Presently there is no cure. Although symptoms can be managed by combinations of different treatments, molecular understanding of Fragile X syndrome remains elusive. N-Glycoproteins are crucial for neuronal signaling and dysregulation is observed in several neurodegenerative diseases; however, the Fragile X glycoproteome has yet to be characterized. We use the Fragile X *Fmr1*^{KO} mouse model to characterize the altered N-glycoprotein expression from whole brain tissue. For characterization and quantitation of glycoproteins, strong anion exchange electrostatic repulsion-hydrophilic interaction chromatography (SAX-HILIC) enrichment is used at the peptide level; the addition of glycopeptide enrichment has been shown to greatly increase glycoprotein detection in various biological samples. As glycoproteins only comprise a small portion of total peptides, quantitative sensitivity is also increased using isobaric tagging techniques. For a comprehensive profiling of the glycoproteome, offline high pH reversed-phase fractionation is used for orthogonal separation of the samples prior to online acidic pH reversed-phase separation using a homemade C18 column. Together, these combined approaches will generate, for the first time, a comprehensive characterization of Fragile X glycoproteome and enable the detection of dysregulated glycoproteins for an improved understanding of the mechanisms of Fragile X syndrome.

P10.03 | Extracellular matrix structural alterations in the brain diseases

Manveen K. Sethi[1] (manveen.sethi1986@gmail.com), Margaret Downs[1], Jonathan S. Vogelgsang[2], Torsten Klengel[3], Sabina Berretta[3], Joanna Phillips[4], Joseph K. Zaia[1]
[1]Boston University, Boston, MA, [2]McLean Hospital, Harvard, Boston, MA, [3]McLean Hospital, Harvard Medical School, Boston, MA, [4]University of California, San Francisco, San Francisco, CA

Advancements in high-throughput technologies such as genomics, epigenetics, transcriptomics, metabolomics, and proteomics have provided promising approaches to understanding the mechanisms of brain diseases. However, these approaches do not address the critical roles of glycosylation in the pathogenesis and progression of neurological diseases. Thus, the potential of glycans and glycoproteins as biomarkers and disease intervention targets remains under-explored. The brain extracellular matrix (ECM) has gained prominence in the neurological field and there has been keen interest in mapping the structural alterations to ECM components, including glycosaminoglycans (GAG), proteoglycans, glycoproteins, collagenous and non-collagenous proteins in the progression and pathogenesis of brain diseases, including Alzheimer's disease (AD), Parkinson's disease (PD), and Glioblastoma (GBM). We have utilized state-of-the-art methods, including on-slide tissue digestion and serial in-solution digestion, and extracted the ECM components (glycans, proteins, glycoproteins) from the human postmortem tissues across disease types, including control vs. disease, and subtype comparisons for AD, PD, and GBM, and subjecting them to liquid-chromatography-tandem mass spectrometry (LC-MS/MS) glycomics, proteomic and glycoproteomic analysis. We have identified specific structural alterations in ECM components and related post-translational modifications (hydroxyprolination, glycosylation), including, altered expression of proteoglycans; versican and biglycan, and chondroitin sulfate-6-O sulfation (CS *6-O*) in AD, changes in collagen hydroxyprolinated peptides in PD, differential expression of CS *6-O-,* heparan sulfate *6-O-* sulfation, glypican-1, biglycan, decorin, chondroitin sulfate proteoglycan 4, and O-glycopeptides for brevican and versican in GBM. Thus, it is evident that ECM is altered during the brain disease pathogenesis, gaining deeper structural knowledge of these changes will assist in understanding the underlying molecular mechanisms driving the development of therapeutic strategies

P10.04 | Multienzyme, On-Trap Digestion-Enabled O-Glycoproteomic Profiling of the Human Pancreas

Dylan Tabang[1] (tabang@wisc.edu), Feixuan Wu[1], Megan Ford[1], Ashley Phetsanthad[1], Danqing Wang[1], Daniel Tremmel[1], Sara D. Sackett[1], Jon S. Odorico[1], Lingjun Li[1]
[1]University of Wisconsin-Madison, Madison, WI

O-GalNAc O-glycosylation modifies Ser and Thr residues and represents a major type of protein modification in the extracellular matrix. Unlike N-glycosylation, O-glycosylation lacks both a single core glycan template and a peptide consensus motif. This makes the study of O-glycosylation more difficult, though the development of O-glycoproteases has improved sample preparation. Abnormalities in O-glycosylation have been implicated in various diseases, including pancreatic cancer. Studying O-glycosylation, especially in the pancreas, may provide valuable insight into protein abnormalities in disease. To analyze O-glycosylation, we first sought to develop a workflow to remove interference from N-glycosylation, to generate O-glycopeptides, and finally to shorten potentially long peptide sequences that are difficult to analyze using mass spectrometry (MS). These criteria were met using the enzymes PNGase F, the O-glycoprotease IMPa, and trypsin. S-trap digestion before enrichment were used to minimize contamination from peptides originating from the enzymes. SDS-PAGE was performed to optimize digestion conditions, which revealed that overnight digestions were not necessary with any of the enzymes. The optimized workflow, which generates peptides that start with glyco-Ser/Thr or end with Lys/Arg, was then used to analyze O-glycosylation of the standard protein bovine fetuin. The resulting peptides were enriched using boronic acid solid phase extraction and analyzed using reversed-phase nanoflow-liquid chromatography-MS.
Removal of N-glycans and usage of IMPa helped to improve the enrichment efficiency to 84%, with 27 glycoforms confidently identified representing 4 of 6 known O-glycosites of bovine fetuin. Confident localization of O-glycosylation, a challenge due to poor fragmentation when O-glycosylated residues are far from peptide termini, was achieved using only collisional dissociation due to the cleavage specificity of IMPa.
Overall, this work demonstrates a workflow for protein extraction, de-N-glycosylation, and digestion to target O-glycopeptides as applied to human pancreatic tissue. It further lays the groundwork for future quantitative analysis of O-glycosylation in pancreatic disease.

P10.05 | Effective Mass Spectrometry-Based Methods to Comprehensively Study Glycoproteins on the Cell Surface and Their Dynamics

Ronghu Wu (ronghu.wu@chemistry.gatech.edu), Georgia Institute of Technology, Atlanta, GA

Surface glycoproteins are essential for mammalian cells to interact with other cells and the extracellular matrix. Surface glycoproteins are involved in nearly every extracellular activity. Aberrant protein glycosylation on the cell surface is directly related to human diseases such as cancer and infectious diseases. Comprehensive analysis of glycoproteins on the cell surface and their dynamics will help identify surface glycoproteins as effective biomarkers and drug targets. In our lab, we have worked on the development of effective mass spectrometry (MS)-based methods to globally and site-specifically analyze surface glycoproteins. Integrating metabolic labeling and bioorthogonal chemistry, we specifically tagged surface glycoproteins and then selectively enriched them. In combination with MS-based proteomics, we performed global identification and quantification of surface glycoproteins site-specifically. The copper-free click reaction is ideal to tag surface glycoproteins because it is quick and specific, and does not need cytotoxic metal ions. After cell lysis and protein extraction, the tagged surface glycoproteins were selectively enriched, and then released for MS analysis. One main advantage of this method is that the conditions are mild, which allows us to study the dynamics of surface glycoproteins. Therefore, this method was applied to quantify the dynamics of surface glycoproteins in monocytes and macrophages in response to the infection. Besides well-known surface proteins changed during the infection, we identified novel surface glycoproteins in response to the infection. Systematic investigation of surface glycoproteins and their dynamics results in a better understanding of glycoprotein functions and cellular activities.

P11: Metabolomics and Imaging Mass Spectrometry

P11.01 | Unsupervised co-registration of H&E microscopic images and mass spectrometry images with feature filtering

Sai Srikanth Lakkimsetty[1] (lakkimsetty.s@northeastern.edu), Olga Vitek[1]
[1]Northeastern University, Boston, MA

Joint analysis of Mass Spectrometry Imaging (MSI) and H&E images has the capabilities to improve performance of clinical applications of using MSI alone. However, integrating the two modalities should be prefaced by co-registration to correct for histological deformities arising from proximal tissue sections, tissue preparation and tissue handling. While there are supervised frameworks such as VoxelMorph, Deep Learning Image Registration framework that capitalize on neural network frameworks, they need prior training and their performance on multi-modal registration is untested. Un-supervised registration techniques for MSI and H&E images do exist but they are not comprehensive enough to handle idiosyncratic attributes of the images pairs such as differing aspect ratios, positioning of the tissue in the image, and the presence of background; and appropriate feature filtering prior to dimensionality reduction. We propose an improved framework that handles multi-modal registration and address the idiosyncratic attributes of the data. We address the multi-variate nature of the MSI by initally employing statistical filtering of the features (which reduces the noise and improves the contrast) followed by summarizing the MSI using t-SNE. The framework allows for removal of background in images of both modalities. The framework automatically performs an intial translation and affine registration so the local deformable registration starts from an advantageous initial state. Finally, we demonstrate the framework's performance using various performance metrics (Dice coefficient, Jacobian matrix and image gradient distance) and visual inspection.

P11.02 | Towards Single-Cell MALDI HiPLEX-IHC: Multiplexed, Multiomic and Multimodal Imaging of Targeted Intact Proteins and Untargeted Metabolites

Mark Lim[1] (mlim@ambergen.com), Gargey Yagnik[1], Zhi Wan[1], Joshua Fischer[4], Ethan Yang[4], Kate Stumpo[4], Kenneth Rothschild[1]
[1]AmberGen, Inc., Billerica, MA, [4]Bruker Scientific, LLC, Billerica, MA

An important goal in the biological sciences is to characterize the metabolomic, transcriptomic and proteomic profiles of intact biological specimens at the single cell level. Towards this goal, the emerging field of spatial multiomics is transforming our ability to understand human disease by providing a comprehensive picture of the immense complexity, heterogeneity, and interactions within biological systems. While significant progress has been made, existing imaging and spatial profiling methods are still not capable of combined metabolomic and proteomics measurements on the same sample, on a single instrument and using an integrated workflow. For example, immunohistochemistry (IHC) provides an important and widely used tool to image multiple protein biomarkers in tissue specimens. However, standard fluorescence based IHC is generally limited to 3-5 different biomarkers (hyperspectral/multispectral methods <10) without resorting to slow and complex iterative cycling methods which can also have a limited field of view. Critically, IHC also cannot image label-free endogenous small molecules such as metabolites and drugs. Conversely, mass spectrometric imaging (MSI) can image label-free small molecules but cannot image large intact macromolecules such as proteins. To overcome this barrier, MALDI HiPLEX-IHC combines the relative strengths of MSI and IHC. MALDI HiPLEX-IHC uses probes such as antibodies and lectins conjugated to novel photocleavable mass-tags (PC-MTs) for MSI of targeted proteins and other macromolecules in cells and tissues. Furthermore, MALDI HiPLEX-IHC workflows enable small molecule MSI (metabolites and drugs), high-resolution optical fluorescence imaging and highly multiplex targeted protein expression imaging on the same sample. We report here examples of MALDI HiPLEX-IHC applied to a variety of tissues including brain, lung, liver tonsil and cancer tissues. High resolution MALDI HiPLEX-IHC data at 5 µm was obtained using Bruker's microGRID system which combines MALDI stage and smartbeam 3D laser beam positioning technology to facilitate high quality imaging at the single cell level.

P11.03 | Optimization and on-tissue imaging of steroid hormones in mouse prostate using MALDI-2

Hannah N. Miles[1] (hnmiles@wisc.edu), Devin M. Tomlin[1], Alexis E. Adrian[1], William A. Ricke[1], Lingjun Li[1]
[1]University of Wisconsin-Madison, Madison, WI

Benign prostatic hyperplasia (BPH) impacts the majority of the aging male population, with associated lower urinary tract symptoms (LUTS) costing the healthcare industry around $4 billion annually. Steroid hormone imbalances are implicated in the development and progression of BPH/LUTS, with many treatment strategies targeting enzymes that metabolize these hormones. Recently, a

novel post-ionization technique was developed in tandem with matrix-assisted laser desorption ionization MS (MALDI-MS) to allow for increased detection of compounds with poor ionization efficiency, such as lipids and steroid hormones. Here, we set out to examine the utility of this technique for spatiotemporal localization of steroid hormones in rodent prostate tissues. Prior to on-tissue analysis, standards of 17β-estradiol, androstenedione, 5a-androstan-3a,17b-diol and 5a-androstan-3a,17a-diol were mixed with 2,5-dihydroxybenzoic acid (DHB) matrix and spotted to test limit of detection. All tissues were flash frozen using liquid nitrogen and embedded in gelatin. Tissue sections of 10µm were cut using a cryostat and placed onto ITO-coated slides. Slide images for future reference of tissue position were obtained prior to DHB matrix application using a robotic TM sprayer. MALDI-2 spectra were acquired using a Bruker timsTOF FleX with MALDI-2 capabilities. Initial experiments using 1µg of steroid hormone standards (17b-estradiol, androstenedione, and 5a-androstan-3a,17b-diol) show drastic improvements in steroid detection, as the protonated peak for b-estradiol (273.18 *m/z*) was detected using MALDI-2 yet undetected with MALDI alone. Additional hormone adducts were detected for b-estradiol, indicating good overall coverage and detection of steroid hormones using post-ionization. Limits of detection are currently being evaluated for these steroids and on-tissue imaging has been performed and is currently in processing. We expect limits of detection to be in the nanomolar or picomolar range, as this has been noted previously. We also expect good signal on our rodent tissue sections of various steroid hormones including cholesterol, estradiol and testosterone.

P11.04 | Evaluation of Multiple Extraction Methods for the Analysis of Human Heart Metabolites

Melissa R. Pergande[1] (mpergande@wisc.edu), Benjamin Wancewicz[1], Yanlong Zhu[1], Kalina Rossler[1], Zhan Gao[1], Zhuoxin Shi[1], Kylie Plouff[1], Ying Ge[1]
[1]*University of Wisconsin-Madison, Madison, WI*

The heart contracts incessantly and requires a constant supply of energy, utilizing numerous molecular substrates such as fatty acids, carbohydrates, lipids and amino acids to supply these high energy demands. Changes to the heart metabolome have been associated with cardiac failure. Therefore, the efficient extraction and detection of metabolites are essential in comprehensive analysis of the heart metabolome to understand their role in cardiac disease. With this, we sought to establish a robust method for the extraction and detection of metabolites from cardiac tissue. Human left ventricle tissue was dissected, snap frozen in liquid nitrogen, and stored at -80 °C under metabolite extraction. Multiple solvent extraction methods (100% methanol, 80% methanol, 3:3:2 acetonitrile:isopropanol:water, Bligh & Dyer, Matyash, and three phase liquid extraction (3PLE)) were evaluated for the extraction of metabolites. Extracted metabolites were analyzed using mass spectrometry (MS) via Fourier-transform ion cyclotron resonance (FTICR) and liquid chromatography (LC)-quadrupole-time-of-flight (Q-TOF) MS. Raw MS data was analyzed using the MetaboScape software program, where chemical formula generation, accurate mass annotation, and tandem mass spectrometry (MS/MS) matching to publicly available spectral libraries was performed. Molecular classification of the extracted metabolites was determined via ClassyFire. Here, we detected 9523 metabolic features using FTICR MS, of which 7699 had a chemical formula and 1,756 an accurate mass annotation. Using LC-MS/MS analysis, we detected 11667 metabolic features, of which 1965 had a chemical formula and 617 an accurate mass annotation. These metabolites covered a wide range of hydrophobicity including polar metabolite classes (benzenoids, alkaloids and derivatives and nucleosides) as well as non-polar metabolites (lipids) (phosphatidylcholines, phosphoserines and fatty acids). In summary, we show the efficient and reproducible extraction of metabolites from human heart tissue. The results of this study can be used to better understand metabolome changes that occur in concert with cardiac disease.

P11.05 | Database development of microbiome-derived metabolites and exploring biochemical landscapes of conventional and germ-free mice using ion mobility-mass

Xiang Tian (xiang.tian@astrazeneca.com), AstraZeneca, Gaithersburg, MD

Microbiome-derived metabolites play a significant role in affecting health and disease. The mechanisms through which this occurs, such as through the “microbiota-gut-brain axis”, however, remain largely unclear. Mass spectrometry-based metabolomics is a powerful approach to explore small molecule changes in biological samples, particularly when it comes to the host-microbiome axis. Here, we compared the global metabolome profile of conventional and germ-free mice, as measured in serum and feces samples, to gain insights into the impact of the microbiome on the host metabolome. To provide broad molecular coverage, multiple LC methods (HILIC (polar) and RP (non-polar)) and ionization modes were used on an Agilent 1290 Infinity LC coupled with 6560 IMQTOF. In addition, we generated an in-house database from more than 1300 metabolite standards, including CCS values, to enhance the throughput and confidence of metabolite annotation. The data was analyzed using MS-DIAL to generate metabolome profiles and R software to perform multiple data analysis methods (PCA and volcano plots) between the different groups. Based on these preliminary data analyses, we have observed large scale changes across multiple classes of metabolites. The in-house database increases the microbiome-derived metabolites annotation confidence and the comparison of metabolite profiles between conventional and germ-free mice provides more insight into the role of the microbiome in the context of the global metabolome. Analysis is ongoing using additional statistical approaches, such as network analysis, and altered metabolites will be further evaluated based on their biological roles

P11.06 | Comprehensive targeted metabolomics workflow utilizing plasma or cell sample preparation, HILIC chromatography, and a LC/TQ

Karen Yannell[1] **(karen.yannell@agilent.com)**, Genevieve Van de Bittner[1], Daniel J. Cuthbertson[2], Cate Simmermaker[1]
[1]*Agilent, Santa Clara, CA,* [2]*Agilent, Seattle, WA*

The goal of life science research is to understand the biological processes of living organisms. The small molecules that facilitate life's biological processes can be detected and quantified using LC/MS metabolomics studies. The most effective LC/MS metabolomics workflows are simple to implement, reliable, reproducible, and provide deep biological insights with ease. Herein is a comprehensive workflow for either plasma or cellular metabolomics aimed at non-expert metabolomics users. Metabolites are extracted using a solid phase extraction (SPE) method that can be optionally automated. Metabolites impacted by the biological process under investigation are targeted using a HILIC-based dynamic-MRM database with retention times. Biological insights are gained using a defined analysis process that includes peak integration, optional quantitation, data normalization, and statistical analysis. The automated sample preparation workflow for polar metabolite extraction from matrix provides consistently high recoveries, with approximately 90% of measured compounds having a recovery between 90% and 110%. Reproducibility of the metabolite recoveries is additionally excellent with an average recovery %RSD of approximately 5%.

Methodology and a database of over 500 metabolites was created for transferability to other labs. Over 300 analytes in this database were detected in pooled matrix. The detected metabolites largely represented compounds from core metabolism signaling pathways like glycolysis, TCA, amino acid, purine, and pyrimidine pathways. The sample-derived dynamic-MRM method demonstrated high sensitivity and reproducibility. RSD of unnormalized polar metabolites was <24% (average) as measured by triplicate injections of technical pooled matrix samples. Using isotopically labeled analytes, femtomole levels of metabolites were detected in the matrix. A longevity study was performed and showed reproducible retention time and peak shapes over the experiment. Different sample groups were measured. Statistical analysis identified those metabolites altered between the groups and the results were highlighted in volcano plots and heat map representations.

P12: Multi-omics Methods and Applications

P12.01 | Multi-omics Sample Preparation Workflow for Simultaneous Cleanup of DNA, RNA, and Proteins using ProMTag

Stephanie Biedka[1] **(stephanie.biedka@impactproteomics.com)**, Stephanie D. Byrum[2], Jonathan S. Minden[1]
[1]*Impact Proteomics, Pittsburgh, PA,* [2]*University of Arkansas for Medical Sciences, Little Rock, AR*

Sample preparation is the critical first step in any -omics workflow. Removal of contaminants, such as salts, detergents, and other molecules that interfere with analysis, while maintaining high yields of the desired product(s) is key to obtaining reproducible, high-quality results. Combining multiple -omics analyses to gain deeper insights into biological processes is becoming increasingly common. However, multiomics sample preparation typically requires several different sample preparation kits or workflows. We have developed a novel multi-omics sample preparation workflow that starts with a single cell or tissue sample and yields clean preparations of DNA, RNA, and protein ready for downstream analysis. This workflow relies on Impact Proteomics' reversible click chemistry tag, ProMTag. ProMTag reversibly modifies primary amines on the surface of proteins, allowing tagging of all proteins in a sample and capture of those proteins on a complimentary ProMTag capture resin. We have modified our foundational Universal Protein Extraction and Cleanup Kit (UPECK) workflow to enable capture of DNA and RNA on this same ProMTag capture resin. The resin-bound protein and nucleic acids can then be thoroughly washed to remove contaminants. The nucleic acids are eluted first by a solubilizing buffer. Proteins are then released from the capture resin by reversal of the ProMTag and digested in a 1-hour digest with our modified MT-Trypsin. The entire workflow requires less than 3 hours to complete. Starting with one aliquot of a cell lysate containing 100 µg of protein, we were able to produce high-quality DNA (average yield 0.35 µg), RNA (average yield 2.4 µg), and peptides (average yield 10.4 µg), which we utilized for whole-genome sequencing, mRNA sequencing, and mass spectrometry, respectively. This novel, streamlined workflow establishes a new standard for high-yield and reproducible simultaneous preparation of DNA, RNA, and peptides, allowing for simplification of multi-omics analyses and more efficient multi-omics data integration.

P12.02 | Integrating permethylation and data-independent acquisition mass spectrometry (DIA-MS) method to characterize nucleic acid modifications

Ye Chen[1] **(c.ye1@wustl.edu)**, Yixuan (Axe) Xie[1], Francisca Vitorino[1], Xingyu Liu[1], Zongtao Lin[1], Benjamin A. Garcia[1]
[1]*Washington University School of Medicine in St. Louis, St. Louis, MO*

Nucleic acids (DNA and RNA) play essential roles in storing, transcribing, and translating genetic information. Nucleic acids commonly carry a variety of chemical modifications to gain high stability and aid in evolution with great flexibility. Researchers have shown that the diverse modifications involve critical significance in many cellular processes, such as transcription regulation and dysregulation can lead to human diseases such as cancer development. While over 140 modifications have been found in nucleic acids, it is still challenging to simultaneously characterize and quantify a large variety of modifications using traditional genomic methods. Mass spectrometry (MS) has emerged as a powerful tool to analyze these nucleic acid modifications. Herein, we developed a method that integrates the permethylation reaction and data-independent acquisition mass spectrometry (DIA-MS) method to characterize nucleic acid modifications. The permethylation reaction replaces the hydrogens on hydrophilic groups (such as hydroxyl groups, amine groups, and carboxyl groups) with d3-methyl groups and yields hydrophobic permethylated nucleosides that are ideal for reverse-phase retention and separation. By applying the nano-flow LC with DIA-MS, superior sensitivity in identification and quantification was achieved. In addition, we fractionated different cellular fractions to examine the different nucleic acid modification distributions in cells as proof of the concept. Using this approach would enable us to explore the functions of the modifications as well as their metabolic mechanisms in cells based on their modification level, which potentially aids in new therapeutic discovery.

P12.03 | Chemoproteogenomics for stratification of expression and ligandability of missense variants in cancer

Heta Desai[1] **(hets@g.ucla.edu)**, Samuel B. Ofori[1], Nicholas Ung[1], Keriann Backus[1]
[1]*UCLA, Los Angeles, CA*

Despite the relative rarity of cysteine (2.3% of all residues in a human reference proteome), cysteine is remarkably the most commonly acquired amino acid due to missense mutations in human cancers. In fact, cysteine is frequently mutated in oncogenes. Given the unique chemistry of the cysteine thiol, including its nucleophilicity and sensitivity to oxidative stress, cysteine residues contribute to numerous biological processes, including enzymatic catalysis and redox signaling. While established proteogenomic

platforms have improved our understanding of proteome variations, the widespread adoption of these methods in chemical-proteomics applications is lacking. Here, we present a chemical proteogenomic workflow to identify ligandable and putative ligandable variants. By implementing a customized FragPipe computational pipeline, we achieve accurate variant identification and MS1 quantitation for activity based protein profiling. Application of this workflow to a panel of 11 cancer cell lines, including mismatch repair deficient and non-mismatch repair deficient lines, results in over 400 variant-containing peptides labeled with cysteine-reactive probes, including many uncharacterized variants.

P12.04 | Multiomic, mass spectrometry-based analysis of dried blood: toward deep phenotyping of sepsis

Matt Foster[1] **(mwfoster@duke.edu)**, J. Will Thompson[2], Brett Phinney[3], Youwei Chen[1], Hongmei Zhu[1], Timothy McMahon[1]
[1]Duke University, Durham, NC, [2]908 Devices, Carborro, NC, [3]UC Davis, Davis, CA

Introduction: The quantification of proteins and metabolites in blood can enable phenotyping of disease. However, most analyses are confined to plasma, which can suffer from pre-analytical variability and may not reflect cellular changes relevant to pathobiology. We have previously utilized a 1h microflowLC-MS/MS method to quantify ~1000 proteins from dried blood collected by volumetric absorptive microsampling, and we found that COVID-19 ICU patients exhibited disease-related changes in protein abundance. We are extending this work to PTMs and metabolites for multiomic analysis of sepsis biorepositories.
Methods: Blood collected from four normal human donors was treated ± phosphatase inhibitors, and twenty microliters was collected on Neoteryx Mitra tips, which were subjected to deoxycholate-aided trypsin digestion for 2 h. After deoxycholate precipiatation, ~1 mg of digests were adjusted to binding conditions and enriched for N-glycopeptides and phosphopeptides using dual functional Ti(IV)-IMAC or TiO. LC-MS/MS analyses utilized 15 min nanoflow gradients and DIA or stepped-collision energy DDA on an Exploris 480 for phosphopeptides and glycopeptides, respectively. Metabolites were extracted directly from Mitra tips and analyzed by ZipChip capillary electrophoresis coupled to high-resolution MS/MS.
Preliminary Results: We developed facile methods for plate-based processing of Mitra tips for PTMs and metabolites. Using a short nanoLC gradient, over 800 N-glycopeptide precursors and 2,900 phosphopeptides were identified. Over 300 phosphopeptides were significantly increased with phosphatase inhibitor, establishing a dynamically-regulated subset of the blood phosphoproteome. Over 200 polar metabolites were quantified using ZipChip CE by alignment against an indexed migration time library. A prototype ZipChip CE with integrated solid phase extraction, and an Evosep/timsTOF platform, are further being evaluated for analysis of phosphopeptides and glycopeptides from blood.
Conclusions: Multiomic analysis of dried blood can be accomplished in a relatively short amount of time and covers a large biological space that should aid in the phenotyping of sepsis.
Funding: GM146142

P12.05 | Identifying Genetic and Environmental Factors on Proteins in Age and Cardiovascular Risk Using SomaScan™

Brendan Epstein[1] **(bepstein@somalogic.com)**, Tina Lai[2], Ted Johnson[2], Michael Hinterberg[3], David Astling[4]
[1]SomaLogic, Inc., Watkins Glen, NY, [2]SomaLogic Operating Company, Inc., Boulder, CO, [3]SomaLogic, Inc., Boulder, [4]SomaLogic Operating, Inc., Boulder, CO

Background: Protein quantitative trait locus (“pQTL”) studies are genome-wide association studies that identify genetic associations with protein levels. Results from the growing number of pQTL studies can be combined with genome-wide association studies to identify proteins that underlie the genetic risk of disease, thus revealing the mechanisms of disease and and potential drug targets.
Methods: We compiled a list of pQTLs from two recent studies based on the SomaScan™ v4 platform, which measures the level of > 5,000 proteins—Ferkingstad et al. (2021) identified pQTLs in 35,559 members of the deCODE Icelandic study; and Zhang et al. (2022) identified pQTLs in 7,213 European Americans and 1,871 African Americans from the ARIC study. We calculated the proportion of analytes with a significant pQTL association in at least one of the studies and compared this to the proportion of pQTL-associated analytes that varied significantly (FDR < 10%) in response to age and four-year cardiovascular disease risk (“CVD”) using data described previously (Williams et al., 2019).
Results: Of the 4,924 analytes included in at least one of the two pQTL studies, about half (48.7%) were cis-pQTL-associated. By contrast, 70% of analytes strongly associated with age or CVD were pQTL-associated (~ 1,100 of 1,600 analytes and ~850 out of 1,200 analytes, for age and CVD, respectively). About half of the CVD-associated proteins were also associated with age, suggesting additional distinct and measurable environmental factors in cardiovascular disease.
Conclusions: Proteins that vary in response to age or CVD risk are more likely than not to have a pQTL association, supporting use of proteomic data in Mendelian randomization and proteome-wide association studies. However, there are still many proteins that had a strong association with an endpoint, but no cis-pQTL association, suggesting that the SomaScan platform also captures the effect of environmental variation on the proteome.

P12.06 | Simultaneous analysis of the proteome, histone PTMs and RNA modifications from FFPE and frozen tissue sections by MS

Joanna Lempiainen[1] **(ljoanna@wustl.edu)**, Yixuan (Axe) Xie[2], Angela C. Hirbe[2], Benjamin A. Garcia[2]
[1]*Washington University in St. Louis, St. Louis, MO,* [2]*Washington University School of Medicine in St. Louis, St. Louis, MO,*

Acquiring reliable proteomics and epigenetics data from limited volumes of tissue is essential for understanding dysregulated signaling in diseases such as malignant peripheral nerve sheath tumor (MPNST). Here, we develop a workflow to simultaneously characterize the total proteome, histone PTMs, and RNA modifications from formalin-fixed, paraffin-embedded (FFPE) core punches and frozen sections of MPNST tissue. FFPE core punches of 5 mg were lysed by sonication with adaptive focused acoustic sonication (Covaris) and de-crosslinked by heating. For total proteome, proteins were digested using S-Trap (Protifi) and analyzed by MS. The remaining lysate was split in two for both histone protein and RNA extraction. For histone extraction, proteins were acetone precipitated and acid extracted, followed by our propionylation protocol. RNA was extracted using TRIzol and further purified using the RNA Clean & Concentrator kit (Zymo Research). RNA was further processed to single nucleosides and permethylated before MS analysis. Frozen tissues of 5-300 mg were prepared with Quick-RNA Tissue Microprep kit (Zymo Research), where total proteome and histones were acetone precipitated from the flow-through. FFPE and frozen sections yielded 600-700 µg and 400-2800 µg of protein, respectively. MS analysis identified on average 2200 proteins in each sample. Proteins were identified from different cellular compartments with about 40% cytoplasmic, 30% nuclear, 10% cell membrane 10% secreted and 10% mitochondrial proteins. RNA yields after cleanup were from FFPE samples 1000-3000 ng and from frozen tissues 300-700 ng. For RNA modifications, MS analysis successfully identified the most prevalent modifications (m6A, i6A and t6A) in frozen sections. We are still processing the RNA from FFPE samples, and histones from both FFPE and frozen tissues. Taken together, the results suggest that these sample workflows may be suitable for the simultaneous analysis of the total proteome, histone PTMs, and RNA modifications by MS from limited volumes of tissue.

P12.07 | Characterizing the Isomeric Forms of PFAS in Serum and Liver with UHPLC and Ion Mobility Mass Spectrometry

Carrie McDonough[1], Jennifer Marciano[2], David Dukes[2], **Sheher B. Mohsin**[3] **(sheher_mohsin@agilent.com)**, David A. Weil[4], Emily Parry[5]
[1]*Department of Chemistry, Carnegie Mellon University, Stony Brook, NY,* [2]*Stony Brook University, Stony Brook, NY,* [3]*Agilent, Barrington, IL,* [4]*Agilent, Santa Clara, CA,* [5]*Agilent, Little Falls, DE*

Per- and polyfluoroalkyl substances (PFAS) are persistent organic pollutants that tend to bioaccumulate due to their chemical inertness and thermal stability. In this work, we demonstrate the advantages of using UHPLC separation with multiplexed ion mobility mass spectrometry followed by high resolution demultiplexing to separate isomeric PFAS Species. High resolution demultiplexing increases the signal intensity and resolution allowing us to separate isomers separated by less than 1 ms in drift time. In the initial study, we characterized PFAS standards to look for isomeric forms which include branched chain as well as linear and other structural isomers. Emerging PFAS possess both linear and branched structures. PFAS standards were analyzed over a wide dilution range in order to investigate the impact that total concentration has on the limit of detection for the structural isomers and labeled internal standards. This study was followed by analyzing human serum samples and liver tissue from mice dosed with PFAS-containing aqueous film-forming foam (AFFF) to look for linear and branched chain PFAS isomers.

Briefly, 100 uL aliquots of serum were transferred to Agilent Enhanced Matrix Removal (EMR) Lipid cartridges with mass-labeled internal standards and crash solvent (acidified cold acetonitrile) then eluted into polypropylene autosamplers for analysis. Liver tissue was diluted 2X in LCMS-grade water, homogenized by bead beater, and a 10 uL aliquot was fortified with internal standards, diluted with crash solvent, and centrifuged. Supernatant was transferred to a polypropylene autosampler vial for analysis.
The isomers were identified with retention times, order of elution, accurate mass and CCS values. Machine learning was used to increase the confidence of assigning structures with predicted CCS values. We find that a combination of drift time resolution along with chromatographic resolution is needed to thoroughly characterize isomeric PFAS species.

P12.08 | Integrated Multi-Omics Analysis of Brain Aging in Female Nonhuman Primates Reveals Altered Signaling Pathways Relevant to Age-Related Disorders

Laura Cox[1], Sobha Puppala[1], Jeannie Chan[1], Kip Zimmerman[1], Zeeshan Hamid[1], Isaac Ampong[1], Hillary Huber[2], Ge Li[1], Avi Jadhav[1], Benlian Wang[1], Cun Li[3], Carol Shively[1], Geoffrey Clarke[4], Tom Register[1], Peter Nathanielsz[3], **Michael Olivier**[1] **(molivier@wakehealth.edu)**
[1]*Wake Forest University School of Medicine, Winston-Salem, NC,* [2]*Texas Biomedical Research Institute, San Antonio, TX,*

[3]*University of Wyoming, Laramie, WY,* [4]*UT Health Sciences Center, San Antonio, TX*

The prefrontal cortex (PFC) is central to working memory, temporal processing, decision making, flexibility, and goal-oriented behavior, and has been implicated as a key brain region responsible for age-related cognitive decline. Although studies in humans and multiple nonhuman primate (NHP) species have shown reduction of PFC activity associated with cognitive decline, little is known about aging-related molecular changes in PFC that may mediate these effects. To date no studies have used untargeted discovery methods and integrated analyses to determine PFC molecular changes across the adult age span in healthy primates. The goal of this study was to quantify PFC molecular changes associated with healthy aging in female baboons (Papio), a NHP model of aging, by integrating RNA-Sequencing data with proteomics and metabolomics data from the same brain samples collected across the adult age span.
We used unbiased weighted gene co-expression network analysis (WGCNA) to integrate the multi-omics data, d identified 2 modules containing 587 transcripts and 13 proteins negatively correlated with age. In addition, we identified an additional 57 proteins and 20 metabolites associated with age using regression analyses. Pathway enrichment analysis revealed 25 overlapping, coordinated pathways negatively correlated with age. We identified pathways previously associated with PFC aging such as dopamine-DARPP32 feedback in cAMP signaling, and additional pathways not previously associated with aging-related PFC changes such as nitric oxide signaling.
These highly coordinated pathway changes during aging may represent early steps for aging-related decline in PFC functions, such as learning and memory, and provide potential biomarkers to assess cognitive status in humans.

P12.09 | Optimization of common lipid extraction methods towards efficient recovery of polar lipids

Valentina Petukhova[1] (valya@uic.edu), Stephanie M. Cologna[1]
[1]*University of Illinois Chicago, Chicago, IL*

Lipids are key components of biological systems representing a complex family of biomolecules significantly varying by lipophilicity. Cellular lipid composition is highly dynamic and changes with physiological or pathological conditions, and successful extraction is a key for potential biomarker identification. Common extraction methods allow for isolation of a wide range of lipid classes, but polar lipids bearing acidic functionalities often suffer from poor recovery, a serious challenge in analysis of biologicals samples. To address this problem, we explored a variety of different extraction techniques including organic solvents, aqueous buffers, and optimized LC-MS conditions to improve recovery of polar lipids including phospholipids and lysophospholipids. Optimization studies were carried out using HeLa cells while our current efforts are focused on expanding these methods to tissues lysates, ultimately focused on understanding lipidomic changes in the fatal neurodegenerative disorder, Niemann-Pick Type C.

P12.10 | A simplified orthogonal electrospray source setup for robust nanoflow or microflow proteomics analysis

Patrick Pribil[1] (patrick.pribil@sciex.com), Jose Castro-Perez[2], Paul R. Baker[3], Ihor Batruch[1]
[1]*SCIEX, Concord, ON,* [2]*SCIEX, Framingham, MA,* [3]*SCIEX, Redwood Shores, CA*

Most bottom-up proteomics analysis workflows are based on nanoliter-per-minute flow rates (nanoflow) during the chromatographic separation of analyzed peptides. This flow regime typically increases the sensitivity of the assay compared to higher flow rate regimes. However, there are significant challenges with nanoflow that impact productivity, such as the fragility of nanoflow spray emitters, their susceptibility to clogging and leaks in the plumbing connections. The OptiFlow TurboV ion source on the SCIEX ZenoTOF 7600 system gives users the capability to do electrospray ionization through either nanoflow (? 1 µL/min) or microflow (? 1 µL/min, ? 50 µL/min) LC connections through 2 separate dedicated electrospray probes/electrodes designed specifically for each of these workflows. In this work, we describe the application of a single probe/electrode for either flow rate during proteomics analysis. Here, the dedicated microflow probe/electrode was tested at flow rates commonly used for nanoflow. This setup has the advantages of (1) spraying orthogonally to the source inlet, thereby minimizing eventual contamination of the system through extended use, (2) not requiring any additional hardware or interface changes to the MS system and (3) simultaneous use of the built-in Calibrant Delivery System (CDS) for fast, easy calibration of the MS system. The performance of the MS system using the microflow probe at nanoflow flow rates was determined using both simple peptide mixtures and complex cell lysate digests, with different LC column types and on-column peptide loadings. At typical nanoflow flow rates (300 nL/min), the microflow probe yielded equivalent performance to when using the dedicated nanoflow probe on the OptiFlow TurboV ion source. The use of the microflow probe/electrode therefore offers users the option for a simplified setup while maintaining the high sensitivity performance required for impactful nanoflow proteomics workflows.

P12.11 | Multi-omics analysis of hepatocellular carcinoma using MALDI Imaging Mass Spectrometry

Kate Stumpo[1] (katherine.stumpo@bruker.com), Peggi Angel[2], Richard Drake[2]
[1]Bruker Scientific, Billerica, MA, [2]Medical University of South Carolina, Charleston, SC

The triple helical region (THR) of fibrillar collagens is sensitive to subtle alterations, such as the post-translational (PTM) conversion of proline to hydroxyproline (HYP), impacting biological status. In nearly all cancer types, HYP PTMs are known to change cell processes of signaling, recruitment, proliferation and migration status. A full systems biology analysis of cancer requires multiple levels of interrogation, including glycans, targeted proteins, and the extracellular stroma. In this study, we use a multiomics approach to visualizing changes in tissue, with multiple classes of molecules investigated, with MALDI-2 enhancement being critical for the detection of nonpolar peptides from collagen THR. Previously published methods using collagenase type III were used to target THR from hepatocellular carcinoma FFPE sections. Tissues were analyzed with conventional MALDI approaches followed by MALDI-2 imaging. MALDI-2 methods typically boost the laser power of the primary MALDI laser which can obscure enhancement effects due to the plume interacting with the MALDI-2 laser. A significant peak intensity enhancement was observed for collagen peptides mapped across tissue. The majority of peptides reported a two-fold increase in signal intensity such as the doubly hydroxylated proline COL1A2 peptide GPPGESGREGAP (m/z 1142.507). However, certain peptides were detected above the noise threshold that had previously not been detected using conventional MALDI. This included COL1A1 peptide GGPGSRGFP (m/z 863. 401), with only 1 HYP, detected only in specific stroma regions and also in a thin 200-µm ring around the tumor. The unmodified version of this peptide (m/z 847.400) was not detected in the MALDI-1 analyzed tissue and showed low detection in the MALDI-2 analysis. Glycan analysis within stroma and the tumor microenvironment, as well as targeted protein imaging utilizing the MALDI HiPLEX-IHC workflow, will also be presented.

P12.12 | Double Your Fun: High-Throughput Multi-Omics using Microchip Capillary Electrophoresis Under Unified Analysis Conditions

J. Will Thompson[1] (wthompson@908devices.com), J. Scott Mellors[1]
[1]908 Devices Inc, Carrboro, NC

Introduction: Researchers place high value on comprehensively evaluating genomic, proteomic and metabolomic alterations which occur to a system under stimulus; this is described as “multi-omic” analysis. Mass spectrometry is the most common technique for both proteomic and metabolomic measurements, however sample preparation and analysis are highly divergent between the two ‘omics approaches, requiring unique expertise and often dedicated instrumentation for each. Here we describe a workflow using prototype ZipChipTM microchip capillary electrophoresis (mCE) which unify the analysis conditions so that both types of measurement can be done back-to-back on the same system without interruption.
Methods: Metabolite samples were extracted with methanol containing NSK-A (Cambridge Isotope Laboratories) and 0.1M Ammonium acetate. HeLa lysate digest and PRTC standards (Pierce) were used for optimizing proteomics conditions, and standard ammonium bicarbonate-based trypsin digestion procedures were used for plasma and tissue digests. Analysis was performed on a prototype SPE ZipChip mCE system coupled to an Exploris 240 Orbitrap MS using Peptides BGE (908 Devices) for both proteomics and metabolomics.
Preliminary Results: We have previously demonstrated the inclusion of an integrated SPE bed into mCE resulting in over 20-fold sensitivity improvement for proteomics compared to standard ZipChip. We have also demonstrated that using indexed migration time libraries with ZipChip mCE-HRMS, we can quantify over 200 high value polar metabolites in less than 6 minutes. In this work, we demonstrate that the two workflows can be performed back-to-back using the same analysis conditions, by incorporating SPE for proteomics analysis and simply bypassing the SPE bed with strong sample solvent for metabolites. Additionally, we have improved the SPE bed geometry and ZipChip hardware, allowing analysis of samples containing less than 1 nanogram of material.
Conclusion: Multiomic analysis using ZipChip mCE allows quantification of hundreds of metabolites and ~1000 proteins in 25 minutes per sample under unified conditions.

P13: Post-Translational Modifications

P13.01 | Optimizing KINATEST-ID For Better Substrate Prediction and Readout Compatibility

Milton Andrews[1] **(andre913@umn.edu)**, Laurie Parker[1]
[1]*University of Minnesota, Minneapolis, MN*

We have successfully been using an R program, KINATESTID 2.0, to identify peptides that may bind more strongly to various kinases. The algorithm for this program is based on statistical analysis using Fisher's exact test. Using this method we determine which set of amino acids statistically appear to be more prevalent up to positions +7 and -7 relative to the putative phosphorylating residue. While this algorithm has generally worked well, there is an implicit assumption that gathering together all the statistically significant residues in the various positions will produce the best result, without considering that there may be some adverse results when placing a specified amino acid residue in proximity to another amino acid residue. Further, we need the designed substrates to be compatible with read-out techniques, especially mass spectrometry.

For specificity, we intend to take the KINATESTID 2.0 program to a level beyond purely statistical reasoning; we intend to add deterministic considerations to allow for more informed decisions as to whether or not a set of amino acids will make for a strong phosphorylating site. For ionization, we are adding a mass spectrometry ionization filter to KINATEST-ID using an algorithm such as ConSeQuence and/or PeptideRank. Preliminary analyses with both of these algorithms showed that ionization propensity could be combined with substrate prediction to provide peptide designs with favorable biochemical and physio-chemical features. These modifications to KINATESTID 2.0 will allow for more suitable candidates to be used in subsequent in vitro analysis.

P13.02 | Analysis of post-translational modifications using fast electron activated dissociation

Randy Arnold[1] **(randy.arnold@sciex.com)**, Nick Morrice[2], Alexandra Antonoplis[3], Christie Hunter[3]
[1]*SCIEX, Bloomington, IN,* [2]*SCIEX, Nether Alderley, Cheshire, United Kingdom,* [3]*SCIEX, Redwood City, CA*

Electron activated dissociation (EAD) is a fragmentation mode that involves capture of electrons by molecular ions to form radicals that dissociate into fragment ions. EAD often preserves labile post-translational modifications (PTMs) and provides peptide sequence information that is complementary to collision induced dissociation (CID). This work focuses on localization of peptide phosphorylation and glycation sites using EAD in a bottom-up liquid chromatography-mass spectrometry workflow. A total of 100 µg of digested HeLa cell lysate was fractionated using reverse-phase HPLC. The resulting 44 fractions were separated using a 21-min microflow LC gradient with a 6 µL/min flow rate. For glycated peptide analysis, human serum albumin was used with similar separation conditions. Data dependent acquisition was performed using the ZenoTOF 7600 system with Zeno MS/MS and either CID or EAD fragmentation. Mascot was used for database searching and results were imported into Scaffold for analysis. From the CID analysis of HeLa cell lysate, 93,866 peptides were identified at 1% FDR using Mascot as the database search engine. With EAD fragmentation, 52,905 peptides were identified at 1% FDR. While CID identified more peptides due to higher frequency of sampling, EAD provided complementary sequence information. EAD also enabled identification of unique peptides not found using CID, increasing the total number of identified peptides by 11%. Next, data were mined for PTMs with a specific focus on identifications that are challenging to achieve with CID data. EAD spectra from phosphopeptides identified the sequence and site of phosphorylation using complete c' and z• (z+1) ion series. Additionally, analysis of glycated peptides using EAD enabled site-specific localization of hexose modifications. Zeno EAD DDA analysis using microflow chromatography enabled confident identification and automatic site assignment of peptide phosphorylation and glycation sites.

P13.03 | Phosphoproteomics Studies using dia-PASEF and short gradients

Diego M. Assis[1] **(diego.assis@bruker.com)**, Denys Oliinyk[2], Stephanie Kaspar-Schoenefeld[3], Gary Kruppa[1], Matthew Willetts[1], Florian Meyer[2]
[1]*Bruker, Billerica, MA,* [2]*Jena University Hospital, Jena, Germany,* [3]*Bruker Daltonik GmbH & Co. KG, Bremen, Bremen, Germany*

Protein phosphorylation shows a prominent role in cellular signaling and many diseases. Mass spectrometry-based phosphoproteomics studies have collectively identified well over 100,000 phosphorylation sites in the human proteome. However, improvements in throughput and sensitivity are highly desirable to disentangle the underlying phosphorylation networks and further extend applications to primary clinical samples. Here, we demonstrate dia-PASEF for high-throughput and high-sensitivity phosphoproteomics. Thus, HeLa full proteome digests and enriched phosphorylated peptides from 200 µg protein mass were

prepared according to the EasyPhos protocol. Peptides were separated within 7- and 21-minute gradients on a 8cm x 75µm column (PepSep) using a nanoElute LC connected a trapped ion mobility – quadrupole time-of-flight MS (timsTOF Pro, Bruker Daltonik). The MS was operated in dia-PASEF mode with an optimized acquisition scheme for phosphopeptides using the py_diAID software and raw files were processed in Spectronaut (Biognosys) using library-free mode. We show the application of dia-PASEF for short-gradient phosphoproteomics from low sample amounts without the need for a spectral library where 12,500 phosphopeptides including ~9,200 class I phosphosites in 21-min gradients were identified. Additionally, we evaluated PTM localization score thresholds for the analysis of phosphopeptide positional isomers and unique phosphosites using collapse plugin tool for Perseus. Decreasing the gradient time to 7-min, we still quantified about 80% of the class I sites, while maintaining a virtually complete data matrix and quantitative accuracy with a median CV <10% in four replicates. Our data shows that TIMS compensates for the higher peptide density per time unit in shorter gradients. We also found examples of positional isomers of nearby phosphosites that were separated by ion mobility but remained unresolved by fast chromatography.

P13.04 | Temporal Phospho-Secretory Signaling of Insulin from Beta Cells

Mohammad Ovais Azizzanjani[1] **(moaziz@stanford.edu)**, Rachel E. Turn[1], Peter K. Jackson[1], Janos Demeter[1]
[1]*Stanford University, Palo Alto, CA*

The primary cilium is a tiny, specialized organelle (~3µm long, ~250nm diameter) that serves as a pivotal signaling hub to drive essential life functions. This cilium is highly organized, and failure to coordinate the generation of a functional cilium leads to severe pathologies including obesity and diabetes. Signaling within the cilium is understudied in part because of its size and limited tools to study it. Recent studies have demonstrated a critical function of primary cilia in the pancreatic islet, including insulin secretion by beta cells. Activation of both ciliary (e.g FFAR4) and non-ciliary (e.g FFAR1) GPCRswith w-3 fatty acid ligands drives tightly regulated insulin secretion. Loss of insulin secretion in beta cells is a major driver of Type 2 Diabetes. Yet, the signaling mechanisms directly downstream of GPCR activation remain largely unexplored. To tackle this challenge, we have monitored the kinetics of phospho-secretory signaling of beta cells after the activation of FFAR4, FFAR1, and GLP-1 receptors. We also examined combinations of FFAR4 and GLP-1 receptor activation both in low and high glucose. We looked at the secretion/phosphorylation signaling of Beta cells based on Glucose-Stimulated Insulin Secretion (GSIS) analysis at different timepoints. We were able to identify phosphorylation changes in different modes of exocytosis of secretory granules. So far, we have captured critical, specific signaling events. For instance, microtubule and ciliary trafficking components including PCM1 and several centrosomal protein were strongly phosphorylated within the first 10 minutes of FFAR4 activation. Also, secretory protein phosphorylation events including synapsins and synaptotagmins suggest the role of SNARE complex-mediated exocytosis in beta cell insulin secretion function. Taken together, we have a powerful model system for probing the phosphoproteomic signatures of insulin secretion in response to different stimuli and can work toward deciphering the downstream pathways that are disrupted in the context of disease.

P13.05 | Proteome and acetylome analyses reveal sirtuin 2 regulates mitochondrial fission-fusion dynamics during viral infection

Cora N. Betsinger[1] **(cnbetsinger7@gmail.com)**, Julia E. Edgar[1], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ*

Sirtuins (SIRTs) are evolutionary-conserved enzymes that are ubiquitously expressed in mammalian cells and function as multifaceted regulators of cellular homeostasis. We previously discovered that SIRTs are broad-spectrum viral restriction factors, protecting host cells against DNA and RNA viruses, including the prevalent pathogen human cytomegalovirus (HCMV). SIRT2, the only human SIRT which predominantly localizes to the cytoplasm, regulates numerous cellular processes that are important during the HCMV replication cycle, including cytoskeletal organization, metabolism, mitochondrial dynamics, and cell cycle progression. Here, we aimed to mechanistically define the function of SIRT2 in host defense during HCMV infection. To accomplish this, we treated cells with the SIRT2-specific inhibitor AGK2 prior to HCMV infection. Surprisingly, AGK2 treatment inhibited early stages of the HCMV replication cycle, suggesting that SIRT2 has a pro-viral function early during HCMV infection. To uncover this function, we analyzed SIRT2 protein interactions using immunoaffinity purification-mass spectrometry, identifying and validating SIRT2 interactions with mitochondrial fission factor (MFF) and the viral protein pUL13. MFF initiates mitochondrial fission, and we recently demonstrated that the previously uncharacterized protein, pUL13, regulates mitochondrial morphology. We next quantified temporal cellular proteome changes during infection and demonstrated that AGK2 treatment results in decreased abundance of proteins involved in mitochondrial fission. Mitochondrial fission is a hallmark of HCMV infection, occurring early during infection to inhibit mitochondrial-mediated immune signaling. However, the host factors involved in HCMV-induced mitochondrial fragmentation remain unknown. Hence, we hypothesized and tested whether SIRT2, through its deacetylation activity, promotes mitochondrial fragmentation during infection. We characterized SIRT2-dependent changes in the temporal acetylome, which we integrated with proteome, interactome, live-cell confocal microscopy, and molecular virology assays. Our findings address the outstanding question of host factor involvement in HCMV-induced mitochondrial fragmentation and point toward a potential therapeutic target for inhibiting early stages of the HCMV replication cycle.

P13.06 | Kinase-specific peptide substrate development for antibody-free high-throughput time-resolved luminescence assays

Dylan Boselli[1] (bosel014@umn.edu), Jason L. Heier[1], Laurie Parker[1]
[1]University of Minnesota, Minneapolis, MN

Homogeneous time-resolved fluorescence (HTRF) assays have quickly become a popular commercially available approach to monitor kinase activity. HTRF assays function by taking advantage of time-resolved fluorescence resonance energy transfer (FRET) between a donor, such as a lanthanide chelating antiphospho-antibody, and an acceptor chromophore, typically covalently bound to a peptide substrate. Available HTRF assays are nonradioactive, time-resolved to reduce background, require little enzyme, and are easily scalable to allow for high-throughput testing. However, they require specific antiphospho-antibodies which drive up the cost and limit the pool of substrates that can be used to only those with antiphospho-antibodies already available.
The Parker Lab has previously developed assays for tyrosine kinases featuring substrates designed to be both kinase-specific and to sensitize terbium upon tyrosine phosphorylation. These substrates allow for intramolecular time-resolved luminescence resonance energy transfer (TR-LRET) between phosphotyrosine and chelated terbium, thus eliminating the need for costly antibodies and fluorophores. Recently, the Zondlo lab developed a minimal peptide motif that sensitizes terbium when serine/threonine residues are phosphorylated, opening the door for similar kinase-specific substrates to be developed for serine/threonine kinases. Using this inducible motif in combination with consensus substrate sequences for CDK5 and its close homolog CDK2, previously reported by the Cantley Lab, we have developed a terbium-chelating peptide substrate for CDK5, a serine/threonine kinase that when dysregulated has been shown to play a role in Alzheimer's Disease progression. This substrate is remarkably specific for CDK5 and not phosphorylated by CDK2/cyclin A, despite the similarity in those kinases' substrate recognition motifs. The discovery of a CDK5-specific LRET-compatible substrate sequence could clear a path for high-throughput monitoring of cellular CDK5 activity under a variety of conditions which would shed light on its pathological role and streamline the development of a CDK5-specific inhibitor.

P13.07 | Integrative Proteomics Identifies Virus Remodeling of Phosphorylation and Ubiquitination in Cellular DNA Damage Response Pathways

Joseph M. Dybas[1] (dybasj@chop.edu), Katarzyna J. Kulej[1], Eui Tae Kim[1], Jennifer Liddle[1], Edwin Halko[1], Namrata Kumar[1], Benjamin A. Garcia[2], Matthew D. Weitzman[1]
[1]Children's Hospital of Philadelphia, Philadelphia, PA, [2]Washington University School of Medicine in St. Louis, St. Louis, MO

Viruses modulate the cellular proteome to subvert host immune responses and establish an environment conducive to infection. Controlling post-translational modifications (PTM) of ubiquitination and phosphorylation is an elegant mechanism for modifying the proteome since these modifications are direct, reversable, and affect diverse functional outcomes. However, a major hurdle in studying viral-mediated PTMs is quantifying the wide array of cellular substrates and range of interconnected pathways that are impacted. Proteomics techniques generate global, quantitative, and systematic analyses, which makes these approaches optimal strategies to elucidate the substrates and pathways targeted by viral-mediated PTMs. Crosstalk between viral-mediated phosphorylation and ubiquitination is a potential strategy for viruses to regulate redirection of PTMs and to generate targeted proteome modulation. However, crosstalk between viral-mediated PTMs has not been studied. Herpes simplex virus (HSV) infection activates the ATM kinase, which induces phosphorylation signaling within the DNA damage response (DDR) pathways. Additionally, the HSV viral ubiquitin ligase ICP0 targets some DDR proteins for ubiquitination. However, the global landscape and functions of HSV-mediated phosphorylation and ubiquitination within the DDR are not defined. I integrated ICP0 di-glycine remnant profiling ubiquitinomics data with HSV phosphoproteomics data to implement a systematic, global quantification of ubiquitination and phosphorylation in the HSV-activated DDR. I identified an intersection of ATM-induced phosphorylation and ICP0-mediated ubiquitination of DDR proteins, revealing a potential crosstalk between these PTMs. I hypothesize that HSV-induced activation of ATM regulates ICP0 function. These proteomics approaches address the questions of 1) what is the global landscape of phosphorylation within DDR pathways, and 2) what is the relationship between ICP0-mediated ubiquitination and DDR activation. This work quantifies a global interface between HSV-mediated PTMs and the cellular DDR pathways, reveals potential novel ATM signaling downstream of HSV infection, and suggests that HSV-induced phosphorylation may act as a regulatory mechanism of ICP0 E3 ubiquitin ligase function.

P13.08 | Epac2 deletion remodels phosphoproteomics in mouse mossy fibers

Kamil Gebis (kamilgebis@northwestern.edu), Northwestern University, Chicago, IL

Epac2 (encoded by gene *Rapgef 4*) is an exchange protein activated by cAMP, which regulates activations of a panel of kinases such as Rap1 and protein kinase B. Epac2 is highly expressed in the hippocampus and regulates a number of synaptic functions including synaptic plasticity. Therefore, we hypothesized that deletion of *Rapgef 4* may alter phosphoproteomics and further affect synaptic functions in mouse mossy fibers. To optimize the enrichment of phosphorylated peptides, we compared four different methods. Tryptic peptides samples were split evenly and underwent TiO_2 or Fe-NTA column enrichments independently. Flow-through (FT) from its respective column was collected (e.g. FT from TiO_2 column) and subject to the other enrichment method (e.g. Fe-NTA column). Then, we used LC-MS/MS analysis to identify differences in phosphorylated peptide yields between those methods. We found that the usage of TiO_2 column enrichment in conjunction with the Fe-NTA column is redundant, as the standalone Fe-NTA column enrichment provided the most consistent and largest yields. Furthermore, LC-MS/MS data shows that only 36.9%, 46.9% and 33.3% of phosphorylated threonine, serine and tyrosine sites are identified in both wild-type and *Rapgef 4* knockout samples, which strongly suggests that deletion of *Rapgef 4* significantly altered phosphoproteomics in mouse mossy fibers.

P13.09 | HypDB: A Functionally Annotated Web-based Database of the Proline Hydroxylation Proteome

Yao Gong[1] (gong0062@umn.edu), Gaurav Behera[1], Luke Erber[1], Luo Ang[1], Yue Chen[1]
[1]*University of Minnesota, Minneapolis, MN*

Proline hydroxylation (Hyp) regulates protein structure, stability, and protein-protein interaction. It is widely involved in diverse metabolic and physiological pathways in cells and diseases. To reveal the functional features of the Hyp proteome, we integrated various data sources for deep proteome profiling of the Hyp proteome in humans and developed HypDB (https://www.HypDB.site), an annotated database and web server for the Hyp proteome. HypDB provides site-specific evidence of modification based on extensive LC-MS analysis and literature mining with 14,413 nonredundant Hyp sites on 5,165 human proteins including 3,383 Class I and 4,335 Class II sites. Annotation analysis revealed significant enrichment of Hyp on key functional domains and tissue-specific distribution of Hyp abundance across 26 types of human organs and fluids and 6 cell lines. The network connectivity analysis further revealed a critical role of Hyp in mediating protein-protein interactions. Moreover, the spectral library generated by HypDB enabled data-independent analysis (DIA) of clinical tissues and the identification of novel Hyp biomarkers in lung cancer and kidney cancer. Taken together, our integrated analysis of human proteome with publicly accessible HypDB revealed the functional diversity of Hyp substrates and provided a quantitative data source to characterize Hyp in pathways and diseases.

P13.10 | Determining Proteoform Landscapes of MEK1 Against Various Drug Treatments for Accurately Localizing Post-Translational Modifications

Raveena Gupta[1] (raveenagupta2026@u.northwestern.edu), Bryon Drown[1], Jared O. Kafader[1], Neil L. Kelleher[1]
[1]*Northwestern University, Evanston, IL*

Various cellular processes that result in oncogenesis, metastasis, and drug resistance, are caused mainly by the dysregulation of signaling pathways modulated by post-translational modifications (PTMs). In this scenario, MEK1 is a dual specificity kinase within the mitogen-activated protein kinase (MAPK) cascade generally involved in metastatic melanoma progression. The MEK1 model is driven by mutant-specific inhibitors resulting in phosphorylations at different sites. However, it is prone to spontaneous resistance to existing drug treatments resulting in the reactivation of the MAPK pathway. Here we combine, for the first time, individual ion mass spectrometry (I^2MS) with HCD activation (I^2MS^2) to explore the MEK1 intact proteoforms landscape and unequivocally characterize the combination of phosphorylation sites past drug treatment. MEK1 was immune enriched from the melanoma cell line A375 after treatment with drugs targeting: upstream of MEK1 pathway (Vemurafenib), MEK1 (Trametinib), or off-pathway (PAC-1 and Everafenib). The intact mass spectra of MEK1 ensemble measurements show the extent of phosphorylation varied with drug treatment. MEK1 overall phosphorylation decreased upon MAPK pathway inhibition with vemurafenib (upstream BRAF mutation inhibitor) and trametinib (MEK inhibitor) treatments. Non-treated MEK1 was phosphorylated up to three times, and we see a decrease in one phosphorylation and a significant reduction in the intensities of the phosphorylations after MAPK targeting drug treatment. Meanwhile, MEK1 remained phosphorylated and had slightly increased phosphorylation levels at the off-target pathway with the procaspase-activating drug in clinical trials, PAC-1. The I^2MS^2 acquisition strategy improved overall sequence coverage to attain an understanding of the localization of phosphorylation sites. We are working on interpreting the fragmentation data to characterize highly modified MEK1 oncoproteoforms and understand drug target specificity. Our new workflow can address the

urgent need to understand the regulations of subcellular localization and function of cancer proteoforms upon different drug treatments by directly accessing PTM stoichiometry and localization.

P13.11 | A Suspension Trapping-based Sample Preparation Workflow for Sensitive Plant Phosphoproteomics

Chuan-Chih Hsu (cchsu@gate.sinica.edu.tw), Academia Sinica, Taipei City, Taiwan

Protein phosphorylation is an important cellular event in regulating plant defense mechanisms in response to environmental stresses. While mass spectrometry-based phosphoproteomics has transformed the way to dissect signal transduction in human cells, study of plant phosphoproteome has remained technically daunting especially for mass-limited samples. The main challenge is the presence of photosynthetic pigments and secondary metabolites produced by plant cells, which increases the complexity of plant lysate. Protein precipitation is a common approach to eliminate those interferences prior to phosphopeptide enrichment. However, protein precipitation steps not only cause significant sample losses but also are time-intensive, which limits the sensitivity and throughput of plant phosphoproteomics. To address the challenges, we have developed a suspension trapping (S-Trap)-based phosphoproteomics workflow, which successfully integrates S-Trap spin column with Fe-IMAC tip, termed tandem S-Trap-IMAC, for direct phosphopeptide enrichment after in S-Trap protein digestion without the need of sample transfer. This novel approach significantly minimizes the sample losses in the step of removing the interferences in plant lysate as well as reduces sample processing time. This method enables identification of more than 600 and 6,500 unique phosphopeptides from 10 and 100 µg of Arabidopsis protein inputs, which shows enhanced sensitivity and coverage of the plant phosphoproteome than the conventional protein precipitation-based workflow. Our results demonstrated this streamlined workflow may have potential for profiling the phosphoproteome of low-input plant samples in a high throughput manner, which is still inaccessible by current plant phosphoproteomics platforms.

P13.12 | Middle-Down Mass Spectrometry Reveals Activity Modifying Phosphorylation Barcode in a Class C G Protein-Coupled Receptor

Ashley N. Ives[1] (ashleyives2022@u.northwestern.edu), Henry Dunn[2], Hamid Afsari Samareh[1], Henrique dos Santos Seckler[1], Arash Foroutan[1], Erica Chavez[1], Rafael D. Donadelli Melani, Ryan T. Fellers[1], Richard LeDuc[1], Paul Thomas[1], Kirill Martemyanov[3], Neil L. Kelleher[1], Reza Vafabakhsh[1]
[1]Northwestern University, Evanston, IL, [2]University of Manitoba, Winnipeg, [3]The Scripps Institute, Jupiter, FL

mGluR2 is a class C G-protein-coupled receptor (GPCR) that mediates synaptic transmission and plasticity in the central nervous system. It has 14 possible phosphorylation sites located in the protein C-terminal tail. In this study, we developed a novel middle-down proteomic strategy coupled with parallel reaction monitoring (PRM) using an Orbitrap instrument as a generalizable method to measure intact phosphorylation states of the C-terminal tail of GPCRs. The new method was used to characterize the phosphorylation state of human mGluR2 from a mammalian cell line in different conditions. mGluR2 is subject to multiple co-occurring phosphorylations and can bear 0-4 detectable phosphorylations without agonist stimulation. These phosphorylations occur as many isomeric states as determined using fragmentation data acquired from PRM. >50% of the phosphorylation content was localized into sites proximal to TM7 and at the very C-terminus of the intracellular tail. mGluR2 activation via glutamate increases intracellular phosphorylation abundance and shifts positional phosphorylation abundance over time. After adding glutamate, the abundance of phosphorylated tails at the MS1 level increased. Subsequent PRM revealed that this increase was concentrated in phosphorylation sites at the C-terminus, while phosphorylation sites proximal to TM7 dropped in phosphorylation content. Finally, intracellular phosphorylation states fine-tune mGluR2 signaling. Given the identification of novel and L-glutamic acid-sensitive phosphosites, we developed mutants where identified phosphorylation sites were mutated to glutamic acid or alanine and measured signaling compared to wild-type receptors. We found that phosphomimetic mutations within 10 amino acids of TM7 lowered receptor sensitivity to L-glutamic acid compared to other tested sites. Overall, we quantified a previously unknown level of heterogeneity in both the degree of phosphorylation stoichiometry and positional distribution in mGluR2. Importantly, we found that phosphorylation is distributed non-uniformly throughout the tail, suggesting a non-random process controlling phosphorylation, and that phosphorylations proximal to TM7 tune receptor sensitivity.

P13.13 | mRNA Display for the selection of specific peptide substrates for Protein Tyrosine Kinases

Akaash Kannan (kanna111@umn.edu), University of Minnesota, Minneapolis, MN

Post-translational modifications (PTMs) serve to functionally diversify the proteome, affecting a protein's structure, stability, and dynamic interactions with other proteins. Every protein that is translated in a cell is likely to be subjected to a post-translational modification throughout its lifespan, resulting in emerging modified protein populations. Countless cellular activities are controlled by PTMs, the dysregulation of PTMs as one can imagine is linked to human disease progression. Thus, it is imperative to be able to create tools and assays to accurately monitor and detect PTMs. Genetically encoded peptide libraries pose several advantages: (i) they're larger in library size, (ii) they allow for rapid screening without requiring synthesis, and (iii) the hits discovered in rounds of selection can be further tuned through mutagenesis or recombination to improve function. These unique features lead to the rapid identification of specific binders, substrates, and inhibitors, along with lowering entry costs for their development.

mRNA display is a cell-free, *in-vitro* selection technique that allows an user to identify polypeptide sequences with desired phenotypes from large natural cDNA or combinatorial synthetic DNA libraries with complexities ranging from $10^{12\text{-}15}$ different clones. mRNA display has been previously used to identify protein-protein interactions, screen kinase or protease substrates, and evolve binders for targets. In this work, we employ mRNA display to select optimal, specific peptide substrates for a set of protein tyrosine kinases. We are applying this to compare substrate profiles for Lyn A and B isoforms. Furthermore, with later rounds of selection, we envision counter-selecting evolved peptides with panels of kinases.

P13.14 | Unraveling the effect of pharmacological inhibition of HDACs on combinatorial histone PTM profiles using mass spectrometry

Rashmi Karki[1] (karkir@wustl.edu), Natarajan Bhanu[1], Richard M. Searfoss[1], Benjamin A. Garcia[1]
[1]Washington University School of Medicine in St. Louis, St. Louis, MO

Histone acetylation, a reversible post-translational modification (PTM), is catalyzed by HATs that transfer the acetyl moiety to lysine (K) residues and leads to transcriptional activation. The removal of the acetyl group is mediated by histone deacetylase (HDAC) enzymes. Histone deacetylase HDAC inhibitors (HDACi) are potent epigenetic modifying drugs that have proven therapeutic use in treating various types of cancers, neurodegenerative diseases, inflammatory disorders, and immunological and cardiovascular diseases. Apart from their effect on transcriptional activation, HDAC inhibitors can indirectly affect gene expression through crosstalk with DNA and histone methylation, manipulation of polycomb group proteins and proteins within the SWI/SNF complex, and regulation of miRNA expression. However, the cross-talk between histone acetylation and methylation remains unclear. In this study, we assessed the effects of thirty HDACis, including the three FDA-approved drugs (Vorinostat, Panobinostat, and Belinostat), on histone PTMs. Native histones extracted from HeLa cells were treated for 24 hours with the recommended IC_{50} doses of HDACis. By using nano-LC with data-independent acquisition mass spectrometry (DIA-MS), both identification and quantification of histone modifications were achieved. Raw files were searched using in-house software, EpiProfile. As expected, several HDACis significantly modulated H3 and H4 acetylation, such as H3K9me2K14ac, H3K9acK14ac, H3K18acK23ac, and H4K44ac, in addition to acetylation of H2A variants. However, methylation marks were also impacted in addition to acetylation. Histone methylation have context-dependent effects for activation and repression not only based on site and degree of methylation, but also the crosstalk between acetylation and methylation. Our study dissecting this relationship through HDAC inhibition will help us explore the cross-talk between these modifications, which potentially aids to deploy these drugs as a combination therapy.

P13.15 | Deep Protein Arginine Asymmetric Di-Methylation Profiling of Breast Cancer Patient-derived Xenografts

Min Ma[1] (mma58@wisc.edu), Fabao Liu[2], Hannah Miles[1], Lauren Fields[1], Lingjun Li[1], Wei Xu[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Advanced Medical Research Institute, Shandong University, Shandong, Shandong, China

Protein arginine methylation catalyzed by arginine methyltransferase (PRMTs), is a prevalent posttranslational modification, which is implicated in a myriad of biological processes including transcription, mRNA splicing, and signaling. Abnormal expression of PRMTs in breast cancer has been reported, specifically, in PRMT1, PRMT4, and PRMT6. These three enzymes are mainly responsible for catalyzing asymmetric di-methylation (ADMA) of arginine and significantly altering the cellular pathways in breast cancer development and progression. Growing studies have been devoted to identifying novel substrates regulated by these three enzymes

in breast cancer cells. However, cell lines as *in vitro* models, do not possess the heterogeneous characteristics of original tumors and therefore exhibit limited correlation to the patient. Patient-derived xenograft (PDX) of breast cancer as an *in vivo* animal model has emerged as a powerful tool for understanding tumor characteristics. Here, we performed global mapping of ADMA modification, using LC-MS/MS along with pan-specific ADMA antibodies in five breast cancer PDX samples. In total, 415 methylated peptides were identified, corresponding to 213 proteins. Subsequently, *in vitro* methylation array demonstrated that approximately 70% of the implicated peptides were methylated by PRMT1, PRMT4, and PRMT6. It is worth noting that 48% of substrates were distinguished in one estrogen receptor type. In addition, in contrast to our previous *in vitro* findings, only 77 ADMA sites were shared between *in vitro* and PDX results. Importantly, some biological pathways, such as mRNA transport and signaling by Rho GTPases, were only observed as related to substrates detected in human tumors. To our knowledge, no report has systematically investigated ADMA in human tumors. Therefore, results from this study will be helpful for the comprehensive investigation of the biochemical features and pathways, permitting a better characterization of molecular mechanisms associated with breast cancer.

P13.16 | Prediction and validation of substrate preferences for the TAM family of kinases

Laurie Parker (llparker@umn.edu), University of Minnesota, Minneapolis, MN

Protein phosphorylation is a crucial post-translational modification in all cells, carried out by kinase enzymes and reversed by phosphatase enzymes. It is regulated by a broad range of factors including protein-ligand and protein-protein interactions, scaffolding, and subcellular localization. Dysregulation of kinase activity leads to cellular abnormalities and disease, making kinases a key target for drug discovery. While a subset of kinases have been heavily studied and ~10% have FDA-approved inhibitors available, there are hundreds of kinases that remain understudied—often called the “dark kinome.” Developing tools to study them is majorly hindered by a lack of information about their mechanisms and substrates. We have adapted an in vitro rephosphorylation phosphoproteomics approach initially developed by others to identify substrate preferences for the understudied Tyro3, Axl and Mer (TAM) family of tyrosine kinases, using our KINATEST-ID bioinformatic process to design novel peptide substrates. These experiments revealed subtle differences in the substrate preferences for these highly homologous family members, enabling the design of potentially selective candidate peptide substrates for each. Biochemical testing of the candidates showed that some do exhibit selectivity within the TAM family. Preliminary analyses with AlphaFold2 suggests that kinase substrate sequence predictions can be rationalized by structural modeling, by predicting the likelihood of peptide binding in the kinase active site with a favorable conformation for catalysis. In addition to providing novel substrates for the TAM family kinases, this work has also led to specific substrates for Tyro3, an understudied kinase with very little previously known about its substrate profile. These studies also position us to further incorporate structural modeling into refining predictions and designing efficient, selective substrates for a very broad range of kinases in the dark kinome.

P13.17 | Improving Analysis of Post-translationally Modified Peptides with MSFragger-Labile Search

Daniel Polasky[1] (dpolasky@umich.edu), Daniel Geiszler[1], Fengchao Yu[1], Kai Li[1], Guo Ci Teo[1], Alexey I. Nesvizhskii[1]
[1]*University of Michigan, Ann Arbor, MI*

Identifying protein post-translational modifications (PTMs) by mass spectrometry (MS) is a key task complicated by fact that many such modifications are labile, dissociating from peptides during tandem MS. Database search algorithms often struggle to identify peptides following PTM loss as the observed fragment ions do not match the expected fragments with intact modification(s). We have recently developed a flexible framework for labile PTM searches in the MSFragger search engine, termed MSFragger Labile. In labile searches, PTM-specific diagnostic ions are used to filter which spectra are searched for modifications, and peptide- and fragment-remainder ions can be used to improve search scores for peptides bearing labile modifications. Fragment remainder ions can also be used to localize a modification, even after partial dissociation. Using these features, we are able to identify more modified peptides and modification sites in wide variety of PTM searches, including of phosphopeptides, ADP-Ribosylated peptides, RNA-crosslinked peptides, and more. The degree of improvement from labile searches depends on the how labile a modification is, ranging from more than doubling the number of highly labile RNA-crosslinked spectra identified to more modest gains of 10-30% in phosphopeptides. Labile search can also be combined with conventional search for partially labile modifications, like phosphorylation, providing a flexible framework to maximize search performance across modifications and fragmentation conditions. MSFragger Labile mode is fully supported in the FragPipe search environment, where it can be combined with our recently described diagnostic ion mining module in PTM-Shepherd to automatically identify abundant fragment ions to use in labile mode search. Labeled and label-free quantitation of peptides bearing labile modifications is incorporated into TMT-Integrator and IonQuant, respectively, and visualization of labile fragment ions is also supported in the accompanying FP-PDV results viewer.

P13.18 | Quantitative Middle-down Proteomics Reveals Histone H3 Acetylation Hierarchy After HDAC Inhibition

Karl F. Poncha[1] **(Karl.Poncha@bcm.edu)**, Nicolas L. Young[1]
[1]*Baylor College of Medicine, Houston, TX*

The physiological response to perturbations depends on signaling pathways that must discriminate between different signals at a granular level. This type of signaling terminates in regulation at the chromatin level that is a consequence of and results in the writing of unique combinations of post-translational modifications (PTMs) on target proteins (e.g.: histone proteoforms). These histone proteoforms function as molecular barcodes, i.e., they serve as sites for complex signal integration in terms of information encoding by “writer or eraser” complexes and decoding by “reader” complexes.
The antagonistic dynamics between Histone Acetyl Transferases (HATs) and Histone Deacetylases (HDACs) control the pool of acetylated histone proteoforms. Dysregulated histone acetylation especially in cancer-stem-like cells (CSCs) results in traits and phenotypes that drive tumor expansion, resistance, and metastasis. Consequently, pan HDAC inhibitors have found use as clinically relevant, potent agents to reverse CSC phenotypes and overcome resistance. However, despite their use in both single and combination therapies, little is known about the effect of HDAC inhibitors on histone proteoforms and PTM Crosstalk.
In this study, we define PTM Crosstalk as how one PTM either potentiates or precludes the presence of another PTM as positive and negative crosstalk respectively. This formalization is used to develop two novel statistical scores to quantify PTM Crosstalk. This framework is then used in conjunction with quantitative middle-down proteomics of Histone H3 variants to uncover a {K23ac} driven hierarchy of histone acetylation after HDAC inhibition: {K23ac} > {K14ac} > {K9ac} > {K18ac}.

P13.19 | Directed Evolution of Acetyllysine Superbinder with Ribosome Display

Kyle Schultz[1] **(schu4312@umn.edu)**, Akaash Kannan[1], Laurie Parker[1], Samson J. Joseph[1], Jason L. Heier[1]
[1]*University of Minnesota, Minneapolis, MN*

High-affinity binding proteins have diverse applications as tools in research and in therapeutics. These molecules can be developed with protein engineering techniques such as ribosome display—an in vitro directed evolution technique optimal for the engineering of small proteins. We have created a vector that has allowed us to successfully display bromodomains on ribosomes. Bromodomains are a diverse group of acetyllysine binding folds with a conserved four-helix bundle. Native bromodomains interact with specific sequences such as acetylated histone tails and typically bind with poor affinity. We will use golden gate shuffling to generate an initial library of chimeric bromodomains encompassing a diverse sample of the available fitness landscape. Additional affinity maturation will be achieved through error-prone PCR. In our selection process, we intend to enrich for variants with high affinity and low sequence specificity. A successfully engineered protein would be a high-affinity binder of acetyl-Lysine. Since bromodomains are small and stable, they are well-suited for a variety of environments. As they are readily expressed in *E. coli*, we envision their utility in applications such as detection agents, affinity proteomics, in-vitro and in-vivo assays.

P13.20 | An Optimized Method of Propionylation and DIA Analysis for the Identification of Histone Post-Translational Modifications

Richard M. Searfoss[1] **(r.m.searfoss@wustl.edu)**, Rashmi Karki[1], Faith M. Robison[1], Zongtao Lin[1], Benjamin A. Garcia[1]
[1]*Washington University in St. Louis, St. Louis, MO*

Histones are DNA-binding proteins that play crucial roles in DNA packaging and gene regulation in eukaryotic cells. These functions are controlled through a highly coordinated network of enzymes that alter post-translational modification (PTM) levels, creating what has been termed the “Histone Code.” To date, over 200 modification sites have been annotated in histones, and alterations in histone PTM patterns have been implicated in a number of diseases. The application of liquid chromatography and mass spectrometry (LC-MS) to studying histone PTM’s has led to novel discoveries in their regulatory roles. However, histones are highly abundant in lysine and arginine residues, making the application of trypsin as a protease challenging. Further, histones and histone peptides are very soluble, making standard LC separation approaches difficult. To combat these challenges, histones are commonly derivatized with propionic anhydride in order to block free lysines from being cleaved by trypsin, while also increasing their hydrophobicity to better separate peptides via LC. Given the complexity of histones and the variability of PTM levels, the optimization of propionylation and LC-MS methods is of utmost importance to maximize identification while minimizing variability. Herein, we evaluated the effect of several variables in the propionylation protocol on overall identification and variation of histone peptides and PTMs. Further, we compared a number of LC flow rates and MS data-independent acquisition (DIA) approaches to test reproducibility of data. The resulting optimized protocol allows for increased peptide identification and quantitation, while boosting overall sample throughput to aid in further discoveries regarding histone PTMs.

P13.21 | Identification of 113 new histone marks by CHiMA, a tailored database search strategy

Jinjun Gao[1], **Xinlei Sheng[1] (xsheng1990@gmail.com)**, Jianfeng Du[1], Di Zhang[2], Chang Han[1], Yue Chen[3], Chu Wang[2], Yingming Zhao[1]
[1]The University of Chicago, Chicago, IL, [2]Peking University, Beijing, China, [3]University of Minnesota, Minneapolis, MN

Liquid chromatography coupled with mass spectrometry (LC-MS/MS) in combination with protein sequence database searching has been the gold standard for identification and quantification of histone marks. Conventional methods for searching databases often employ a "target-decoy" strategy to calculate the false discovery rate (FDR) and distinguish true peptide-spectrum matches (PSMs) from false ones. While powerful for large MS/MS datasets, this strategy, when analyzing histone marks, has the caveat of inaccurate FDR due to the small size of the MS/MS datasets. Small MS/MS datasets are often used when the PTMs of interest are of low abundance in the analyzed samples, which is common for recently discovered histone marks, e.g., lactylation. To address this challenge, we developed a tailored database-search strategy, named "Comprehensive Histone Mark Analysis (CHiMA)", to identify histone PTM sites in small MS/MS datasets. Instead of implementing the target-decoy-based FDR, our method uses 50% matched fragment ions as the key criterion to identify high-confidence PSMs.

We manually analyzed the histone lysine lactylation (Kla) datasets that were previously published, and we compared the identified Kla sites by the conventional methods and CHiMA. Using CHiMA, we identified twice as many histone modification sites as the conventional method. The newly identified histone marks were manually verified by MS/MS analysis of the corresponding synthetic peptides. We also discovered that integrating three most common co-existing modifications (Kac, Kme1, and Rme1) into the database search could markedly improve Kla identification. To demonstrate the utility of CHiMA, we applied the method to the analysis of several in-house proteomic datasets generated to identify histone lysine acylation marks. This analysis led to the identification of 113 new histone marks for four types of lysine acylations, greatly expanding the repertoire of these histone acylation marks.

P13.22 | Characterization of three lactylation isomers, KL-la, KD-la and Kce

Di Zhang[1], Jinjun Gao[2], Zhijun Zhu[3], **Xinlei Sheng[2] (xsheng1990@gmail.com)**, Qianying Mao[1], Zhiqiang Xu[5], Pankaj Singh[6], Cornelius Rimayi[6], Carlos Moreno-Yruela[7], Shuling Xu[3], Gongyu Li[3], Yue Chen[4], Christian A. Olsen[7], Nathaniel W. Snyder[6], Lingjun Li[3], Lunzhi Dai[5], Yingming Zhao[2]
[1]Peking University, Beijing, China, [2]The University of Chicago, Chicago, IL, [3]University of Wisconsin-Madison, Madison, WI, [4]University of Minnesota, Minneapolis, MN, [5]Sichuan University, Chengdu, Sichuan, China, [6]Temple University, Philadelphia, PA, [7]University of Copenhagen, Copenhagen, Denmark

Histone post translational modifications (PTMs) and cell metabolism are intricately linked. We recently identified lysine lactylation which regulates gene expression in bacteria-induced macrophage polarization. This modification was proposed to be L-form ($K_{L\text{-}la}$) derived from hypoxia-enhanced glycolysis, but two isomeric modifications, lysine D-lactylation ($K_{D\text{-}la}$) and N-ε-(carboxyethyl) lysine (K_{ce}), can also derive from glycolysis and decorate histone proteins. Therefore, we aim to distinguish these modifications and understand their regulatory mechanisms. To characterize these PTMs, we developed pan anti-$K_{L\text{-}la}$, pan anti-$K_{D\text{-}la}$, and pan anti-K_{ce} antibodies. We tested these antibodies using PTM-specific antigens for $K_{L\text{-}la}$, $K_{D\text{-}la}$ and K_{ce}, including peptide libraries, modified BSA, and sequence-specific histone peptides. Our immunoblots showed that these pan antibodies were specific for their corresponding PTMs. In addition, we subjected synthetic histone peptides bearing $K_{L\text{-}la}$, $K_{D\text{-}la}$ and K_{ce} to a C18 HPLC column connected to an EASY-nLC 1000 system coupled to a Q-Eactive mass spectrometer. $K_{L\text{-}la}$ and $K_{D\text{-}la}$ display similar retention time, while CEL can be separated.

Using the pan antibodies, we discovered that $K_{L\text{-}la}$, $K_{D\text{-}la}$ and K_{ce} exhibit distinct dynamics in response to metabolic alterations. Histone $K_{L\text{-}la}$ was induced by glucose, but not by glyoxalases. In contrast, cellular $K_{D\text{-}la}$ levels were induced in the absence of GLO2, but insensitive to glucose concentration. Under the same conditions, no obvious change in K_{ce} levels was observed.
We previously proposed lactyl-CoA as a cofactor for cellular $K_{L\text{-}la}$. To testify this hypothesis, we measured the concentrations of lactyl-CoA, acetyl-CoA, and propionyl-CoA during $^{13}C_3$ sodium L-lactate or U-$^{13}C_6$ D-glucose treatment. By assessing isotope tracing, we found that lactyl-CoA was dynamically regulated by lactate and glucose metabolism. Further substantiating our findings, we observed that lactyl-coA levels were increased during hypoxia, in agreement with histone $K_{L\text{-}la}$. These results suggest that the histone lactylation observed in our previous study was de facto L-lactylation mediated by L-lactyl-CoA.

P13.23 | Temporal Phosphoproteomics of Ciliogenesis in Human Nasal Epithelia

Rachel E. Turn[1] **(rturn@stanford.edu)**, Mohammad Ovais Azizzanjani[1], Janos Demeter[1], Peter K. Jackson[1]
[1]*Stanford University, Palo Alto, CA*

The cilium, a small organelle protruding from the cell surface, is a critical regulator of signaling and cellular function. When disrupted, a number of specific pathologies can arise, including multi-system developmental disorders called ciliopathies, metabolic disorders including diabetes and obesity, and cancers. Yet, these cellular signaling compartments remain poorly understood, in particular because of their tiny size (~3µm long, ~250 nm, versus ~100 µm diameter of cell). There is an urgent need to develop novel approaches to tease apart the ciliary mechanisms that are disrupted in pathologies. In particular, one challenge is deciphering the early mechanisms that drive the assembly of new cilia, called ciliogenesis. Several key regulators have been identified to date. How these factors interact temporally and spatially and how specific post-translational modifications including phosphorylation drive these interactions remains unclear because of the lack of sensitivity of current approaches, especially in the context of proteomics. To address this challenge, our lab established a bottom-up proteomics protocol for the study of multiciliated human nasal epithelial (HNE) cell differentiation over 3 weeks. Simultaneously, we harvested HNE for immunofluorescence to track the stages of ciliogenesis. We examined the phosphoproteome at specific stages of airway development. Using only small (6.5mm) trans well inserts containing ~20,000 cells, we performed high sensitivity phosphoproteomics, achieving an unprecedented phosphoproteome coverage of 10,000 phosphopeptides. This high sensitivity permitted us to identify key phosphorylation events of known ciliogenesis regulators as well as generate strong models for previously unidentified mechanisms. Using Kinase-Substrate Enrichment analysis and bioinformatics, we can impute activation of controllers of the cell cycle, apical basal polarity, cell adhesion, and other pathways. Taken together, the identification of many new candidate ciliary proteins and their modifications in the HNE will allow us to expand our understanding of airway biology and how disruption of cilia drives disease.

P13.24 | Sirtuin 6 induces an antiviral DNA-damage response to restrict replication of DNA viruses

Matthew Tyl[1] **(mtyl@princeton.edu)**, James Kostas[1], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ*

Sirtuins are NAD^+-dependent enzymes that function as “sensors” of the cellular environment by transmitting information through the addition or removal of posttranslational modifications. Initially designated as Class III histone deacetylases, it is now clear that sirtuins have a range of enzymatic activities and substrates. An important example is sirtuin 6 (SIRT6). Originally characterized as a nuclear lysine deacetylase acting on histone 3, SIRT6 was found to also function as a long-chain deacylase and an ADP-ribosyltransferase. Through these enzymatic activities, SIRT6 performs a series of critical cellular functions, including the regulation of transcription and DNA repair pathways. Our lab previously demonstrated that SIRT6 restricts the replication of several viruses, including human cytomegalovirus (HCMV). However, the mechanism underlying its antiviral function remained unknown. Here, we bridge molecular virology with MS-based acetylome, interaction, and proteome studies to characterize the role of SIRT6 in viral infection. We demonstrate a requirement for SIRT6 deacetylase activity in restricting HCMV replication. We determine that SIRT6 antiviral function occurs during the intermediate stages of HCMV replication, inducing defects in genome replication, capsid nuclear egress, and infectivity of released virions. Complementing these molecular virology findings, we define the impact of SIRT6 on the cellular proteome by quantitative MS, revealing a reduction in genome replication proteins and cell cycle regulators. Using microscopy, we show this impact is through SIRT6 dysregulation of DNA repair pathways required for DNA virus genome replication. Lastly, we performed the first unbiased search for SIRT6 deacetylase substrates by profiling the cellular acetylome after CRISPR-mediated SIRT6 knockout and integrating with SIRT6 protein interactions we acquired through IP-MS. This integration pointed to likely SIRT6-mediated acetylation sites on SIRT6 interacting partners, including lysine residues on key DNA-damage repair proteins. Altogether, our results lead us to a model whereby SIRT6 deacetylates DNA repair proteins to inhibit proviral DNA repair pathways.

P13.25 | Ion-mobility-based Gas-Phase Fractionation dramatically improves Ubiquitylated and Acetylated peptide identifications and chimeric resolution

Sebastian Vaca[1] **(sebastian.vaca@bruker.com)**, Elizabeth Gordon[1], Diego M. Assis[2], Tharan Srikumar[3], Jonathan R. Krieger[4], Matthew Willetts[5], Christoper Adams[6], Barry Zee[7], Kathryn Abell[7]
[1]*Bruker, San Jose, CA,* [2]*Bruker Scientific LLC, Billerica, MA,* [3]*Bruker Canada Ltd., Oshawa, ON,* [4]*Bruker Canada, Toronto, ON,* [5]*Bruker, Billerica, MA,* [6]*Bruker Scientific LLC, San Jose, CA,* [7]*Cell Signal Technology, Boston, MA*

Ion mobility is a powerful extension to mass spectrometry that delivers information about the three-dimensional structure of a peptide. Trapped ion-mobility spectrometry (TIMS) increases selectivity and peak capacity, allowing for a more confident

identifications and generating a CCS value for each analyte measured. Here we used of ion-mobility-based Gas-phase Fractionation (GPF) combined with parallel accumulation-serial fragmentation (PASEF) on a Bruker timsTOF SCP to evaluate advantages in K-GG and K-Ac peptide characterization. GPF optimization was performed with K562 digest evaluating the use of a single narrow (0.7-1.3 V s/cm^2) and large ion-mobility ranges (0.6-1.6 V s/cm^2) as compared to ion-mobility fractions (>2 gas-phase fractions). The peptides were separated using a Bruker NanoElute system with an IonOpticks Aurora C18 column (25cmx75um, 1.7um) couple to a TimsTOF SCP system. The data was analyzed using PASER 2022c and MsFragger v18.
In an 85 min gradient, injecting 50ng on column, we observed GPF methods produced more than 8,000 proteins groups and more than 119,000 precursors. This represented more than 80% more identifications than an unfractionated run analyzed with the same LC-MS conditions, which identified 7,300 proteins and 64,000 precursors.
This approach was then applied to the analysis of enriched samples of ubiquitylated and acetylated peptides from three different tissues (brain, embryo and liver). The ion-mobility based GPF allowed identification of a combined >24,000 ubiquitylated peptides and >21,000 acetylated peptides. This represents an increase of ~20% and ~40% more peptide identifications, respectively, as compared to the unfractionated runs. Additionally, GPF using TIMS for K-GG and K-Ac peptides resolves thousands of co-eluting chimeric species by mobility offset mass alignment (MOMA) where we demonstrate how these CCS values can be used in quantification and to train deep learning algorithms.

P13.26 | High-throughput simultaneous quantification of glycopeptides and phosphopeptides by 12-plex DiLeu isobaric tags and dual-functional Ti(IV)-IMAC

Feixuan Wu[1] (feixuan.wu@wisc.edu), Danqing Wang[1], Dylan Tabang[1], Junfeng Huang[2], Lingjun Li[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Chinese Academy of Sciences Guangzhou Institutes of Biomedicine and Health: Guangzhou, Guangzhou, China

Pancreatic cancer is the seventh-leading cause of cancer-related deaths worldwide. The characterization of post-translational modifications (PTMs) of proteins in this disease is essential because those modifications can alter molecular structure and protein activity. Glycosylation and phosphorylation are the most common, and important, PTMs. Alteration of glycosylation correlates with the progression of cancer and many other disease states. For instance, aberrant mannose-6-phosphorylation (M6P) of lysosomal acid hydrolase causes lysosomal storage diseases. Phosphorylation regulates various cellular signaling pathways, especially in tumors. However, without proper fractionation, substantial phosphopeptides that coelute in the enrichment step can cause suppression of the glycopeptide signals in subsequent liquid chromatography tandem mass spectrometry (LC−MS/MS) analysis. Here, glycosylated and phosphorylated peptides were simultaneously enriched using a dual-functional Ti(IV)-immobilized metal affinity chromatography (IMAC) approach and analyzed with MS. Tissues were cryo-pulverized into a fine powder before protein extraction using RapiGest. Then proteins were subjected to enzymatic digestion with trypsin. The peptide samples were labeled with 12-plex DiLeu isobaric tags, followed by Ti(IV)-IMAC enrichment. The dual-mode fractionation was applied in the elution. The enriched glycopeptides and phosphopeptides were separated using nanoflow-reversed phase liquid chromatography-mass spectrometry (nRPLC-MS) using a Q-Exactive HF Quadrupole-Orbitrap mass spectrometer. Raw data files were analyzed using Byonic and MaxQuant.

Before applying on complex biological samples, dual-functional Ti(IV)-IMAC without labeling was used to simultaneously enrich phosphopeptides and M6P glycopeptides from human PANC-1 cell lysate, and loaded onto nanoLC coupled to an Orbitrap Fusion Lumos Tribrid mass spectrometer. In total, 5439 phosphopeptides were identified and 244 intact M6P glycopeptides with site-specific glycoforms were profiled, which corresponded to 1903 phosphorylated proteins and 83 M6P glycosylated proteins from PANC-1 cells. This result suggests the potential for high-throughput quantitative analysis. Our study provides a powerful tool that enables simultaneous separation and in-depth characterization of glycopeptides and phosphopeptides from human pancreatic cancer samples.

P14: Protein Complexes and Interactomics

P14.01 | Destabilized Domain Methodology to Study Time-Resolved Host-Pathogen Interactomics

Crissey Cameron[1] (crissey.d.cameron@vanderbilt.edu), R. Mason Clark[2], Adam Metts[1], Runze Jiang[1], LaToya Scaggs[1], Lars Plate[1]
[1]Vanderbilt University, Nashville, TN, [2]University of Illinois College of Medicine, Chicago, IL

Protein-protein interactions are integral to almost all cellular processes, from ligand-receptor binding for signal transduction to large complex formation for proteolysis. Most current methods to study protein interaction networks are limited to steady state, or all of the interactions during the entire protein lifespan, missing the dynamics and transient nature of many protein interactions. This work presents a methodology to selectively label a protein population and identify associated interactions with time resolution. We utilize destabilized dihydrofolate reductase (dDHFR) tagged to the N-terminus of a protein of interest (POI) causing the dDHFR-POI complex to be constitutively degraded by the proteosome. The dDHFR-POI complex can be rescued from degradation using a trimethoprim (TMP) moiety that stabilizes the dDHFR. We generate a multifunctional TMP probe compound containing a Cys reactive group for covalent dDHFR labeling and a terminal alkyne. This Click handle allows for further functionalization with an affinity purification group and fluorophore. Our methodology was first validated using yellow fluorescent protein (YFP) as a model, showing time-dependent accumulation of YFP and the capability of enriching the labeled population through affinity purification. Using this technology, a pulse-chase assay was used to determine the time-resolved interactome of the SARS-CoV-2 nonstructural protein 15, an endoribonuclease known to play a key role in host immune suppression.

P14.02 | Evaluation of the Structure of the TclIJN Complex to Understand the Role of Post Translational Modification in Novel Antibiotic Discovery and Creation

Clarissa D. Clark[1] (claydc17@gmail.com), Leena M. Patil[1], Diana Calvopina[1], Joel Griffitts[1], John C. Price[1]
[1]Brigham Young University, Provo, UT

An estimated 3 million infections and 35,900 deaths every year in the US are caused by antibiotic resistant bacteria. As the prevalence of antibiotic resistant bacteria grows, so does the need for novel antibiotics to combat them. One such potential antibiotic is TclE, an antibacterial peptide synthesized by the bacteria *Macrococcus caseolyticus*. The TclE peptide consists of a leader and a core sequence that undergo heavy post-translational modifications. TclE is modified by a protein complex, TclIJN, which attaches to TclE's leader sequence and alters a series of cysteines in the TclE core. Understanding the mechanisms of the post-translational modifications of TclE shows promise for aiding in the creation of novel antibiotic treatments. In order to better understand how the post-translational modifications of TclE impact its antibacterial properties, we are interested in the interactions between the proteins of the TclIJN complex (TclI, TclJ, and TclN). We are aware that the complex consists of a combination of the three proteins, but the exact stoichiometry of the complex has yet to be elucidated.

We initially ran the complex on an SDS page gel to determine the size of each protein. However, due to dissociation during the SDS page, we were unable to gain an accurate weight of the complex. As a result, we used High Performance Liquid Chromatography (HPLC) to perform Size Exclusion Chromatography (SEC) to determine the molecular weight of the TclIJN complex in a non-denatured state. After running purified samples on SEC, we used mass spectrometry to verify the identity of each eluted molecule. Understanding the nature of this complex will help us understand the creation of antibiotics like TclE and will aid in creating more effective compounds against bacteria. We will share our conclusions on stoichiometry and protein interactions.

P14.03 | Elucidating novel regulatory lncRNA-protein interactions in hematopoiesis by hybridization capture and mass spectrometry

Yuling Dai[1] (dai59@wisc.edu), Isabella T. Whitworth[1], Brian Frey[1], Mabel M. Jung[1], Emery H. Bresnick[1], Lloyd M. Smith[1]
[1]University of Wisconsin-Madison, Madison, WI

Long non-coding RNAs (lncRNAs) are widely expressed and have key roles in different mechanisms of gene and cell regulation. Over the past decade, research has shown that lncRNAs help coordinate gene expression to regulate cell fate in the hematopoietic system. Certain lncRNAs exert their regulatory functions through interacting with specific proteins, and thus characterizing these

interactions is essential in understanding the mechanisms of action. GATA2 has been shown to govern hematopoietic stem cell (HSC) genesis from hemogenic endothelium and the proliferation of hematopoietic progenitors, while reduced GATA2 expression creates a predisposition to develop myeloid leukemia. Although GATA2-regulated lncRNAs have been found to control human erythroid precursor cell function and erythrocyte development, many mechanistic questions remain unresolved. We aim to develop and integrate new tools to discover novel lncRNA-protein interactions, specifically focusing on discovery of GATA2-dependent regulatory lncRNA-protein interactions. First, we performed differential expression (DE) analysis by analyzing RNA-seq datasets, which were obtained from two cultured cell lines, wild type (WT) and *Gata2* upstream enhancer (-77)-mutant (KO) myeloerythroid progenitor cells. We discovered five GATA2-regulauted lncRNAs, and qPCR analysis validated the RNA-seq data. Using a prioritization strategy, we selected three lncRNAs that show sufficiently high transcript levels, as well as significant expression differences between WT and KO cells (i.e., they are GATA2 regulated). HyPR-MS is a hybridization capture strategy that is based upon the specific capture by nucleic acid hybridization of lncRNAs that have been subjected to i*n vivo* formaldehyde crosslinking followed by mass spectrometric identification of associated proteins. We applied HyPR-MS for purification of novel lncRNA?protein complexes and employed mass spectrometry to identify novel lncRNA-protein interactors that play an important role in hematopoiesis.

P14.04 | Multiplexed kinase interactome profiling quantifies cellular network activity and plasticity

Martin Golkowski[1] (martin.golkowski@utah.edu), Andrea Lius[2], Tanmay Sapre[2], Ho-Tak Lau[2], Taylor Moreno[2], Dustin Maly[3], Shao-En Ong[2]
[1]University of Utah, Salt Lake City, UT, [2]University of Washington, Department of Pharmacology, Seattle, WA, [3]University of Washington, Department of Chemistry, Seattle, WA

Dynamic perturbations of protein-protein interaction (PPI) networks underlie all physiological cellular functions and drive devastating human diseases. Profiling PPI network dynamics can, therefore, provide critical insight into disease mechanisms and identify new drug targets. Kinases are actionable, regulatory nodes in many PPI networks, yet facile methods to systematically study kinase interactome dynamics are lacking. We developed kinobead competition and correlation analysis (kiCCA), a quantitative mass spectrometry-based chemoproteomic method for rapid and highly multiplexed profiling kinase interactomes in native cell and tissue lysates. kiCCA utilizes competition-binding experiments with a panel of multi-targeted, soluble kinase probes and multiplexed inhibitor beads (kinobeads) to globally identify kinase PPIs based on the similarity of competition profiles of kinases and their interaction partners. We used kiCCA to globally map kinase interactomes in 18 diverse cancer cell lines, interrogating PPI rewiring in the context of cancer type, plasticity, and signaling states, and revealing that PPI network topologies are highly dynamic and context dependent. We identified and quantified 1,154 high-confidence PPIs between 238 kinases and 684 non-kinase proteins, which we compiled into an extensive and easily accessible kinase interactome knowledgebase (https://quantbiology.org/kiCCA). We found that kinase-mediated PPIs can be exploited to determine kinase functional states, and to study how kinases integrate with signaling pathways. We discovered that cancer cells that show epithelial-mesenchymal plasticity (EMP) and increased therapy resistance drastically rewired kinase PPI networks like endocytic and vesicle trafficking pathways controlled by an adapter-associated kinase 1 (AAK1) complex. We demonstrated that kiCCA can provide actionable leads for drug target discovery by knocking down AAK1 complex components, which greatly affected EMP marker expression and sensitized cells to targeted therapy. To conclude, our method and kinase interactome knowledgebase are invaluable tools for studying kinase-mediated PPI networks to better understand how signaling events influence disease states. Dynamic perturbations of protein-protein interaction (PPI) networks underlie

P14.05 | Identifying Ciliary Proteins in Mammalian Retinas using a Gentle Extraction Method

Adeline Fredrick[1] (afred3@uic.edu), Erin Claussen[1], Kevin Drew[1]
[1]University of Illinois Chicago, Chicago, IL

The retina, a complex structure in the internal layer of the eye, converts light signals to neural impulses to the brain. The structure consists of a junction between photoreceptors called the inner segment (IS)/outer segment (OS) line. This line has highly modified cilia that develop the primary cilia of the retina and connect the IS to the OS. The primary cilia structure is important because defects in ciliary proteins are associated with diseases such as retinitis pigmentosa (RP), a hereditary retinal degeneration. Currently, we lack an understanding of protein complexes involved in forming and maintaining primary cilia of the retina. We lack optimized protocols to obtain samples with primary cilia from retinal photoreceptors to use in high-throughput proteomic studies. Here, bovine eyes were used to conduct the research on ciliary proteins in the retina due to their availability and size which allows a high yield of proteins harvested per retina. The eyes were dissected in a darkroom to avoid light stimulation, a Mammalian ringer solution was used to gently release the retina from the back of the eye and optic nerve.1 A lysis buffer and glass dounce were used to break the retina cells. The lysed cells were centrifuged leaving a supernatant. Once the proteins were retrieved, detergent compatible protein assay was used followed by mass spectrometry to identify proteins present. Researching the retina allows an indepth look at the structure and function of a highly modified and specialized primary cilia. Future plans to extend this work include a co-fractionation

mass spectrometry pipeline to identify protein complexes in retinal photoreceptors; this will allow us to map the molecular network of disease relevant cells. 1 Panfoli,I etal.(2022).Maximizing the Rod Outer Segment Yield in Retinas Extracted from Cattle Eyes.Bio-protocol12(14):e4474.DOI:10.21769/BioProtoc.4474.

P14.06 | How to Fold Every Protein: Applying Proteomics to the Folding Problem

Philip To[1], Yingzi Xia[1], Divya Yadav[1], **Stephen D. Fried[1] (sdfried@jhu.edu)**
[1]Johns Hopkins University, Baltimore, MD

The striking accuracy of AI-based structure predictors has given new credence to Anfinsen's dogma that protein three-dimensional structures is encoded at the amino acid sequence level; nevertheless, the journey by which proteins navigate their energy landscapes to locate their native structures is complex and prone to failure. Protein folding biophysics seeks to understand how this process works, yet despite 60 years of research, our understanding of the topic is deep but narrow: We have detailed knowledge of folding pathways for a few small soluble reversibly-refolding proteins, but a general understanding of how complex proteins self-assemble and what cellular factors and processes are necessary for them to do so is lacking. To address this gap, we perform proteome-wide refolding experiments on denatured extracts and use limited proteolysis-mass spectrometry (LiP-MS) to interrogate the structures populated by proteins during (and after) refolding. We find that nonrefoldability is surprisingly common (more than one third of soluble E. coli proteins) and frequent for larger multi-domain proteins and particular fold topologies. We have also interrogated proteome-wide refolding assisted by the E. coli chaperones, GroEL (Hsp60) and DnaK (Hsp70), identifying which classes of proteins disproportionately lean on support from these molecular machines to fold to their native conformations. These experiments also highlight a cohort of proteins which do not refold even with chaperone assistance, which we hypothesize is because such proteins have a preference to fold cotranslationally on the ribosome. LiP-MS experiments conducted on reconstituted proteomes - synthesized by *in vitro* translation on total transcriptomes - provides support for this view. In summary, structural proteomics methods are powerful tools to explore protein folding globally, sensitively, and *in vivo,* and have helped move protein folding research from its traditional reliance on "model" proteins.

P14.07 | Reconstructing the vertebrate brain protein interactome using high-throughput proteomics

Vy Dang[1] (vyqtdang@utexas.edu), Claire McWhite[2]
[1]University of Texas at Austin, Austin, TX, [2]Princeton University, Princeton, NJ

Understanding the complexity of the vertebrate 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 vertebrates 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 vertebrate 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 vertebrates, namely pig, mouse, chicken and rabbit. 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 vertebrates in order to define a large-scale map of conserved protein complexes in the vertebrate brain. Additionally, we performed chemical crosslinking mass spectrometry (XL-MS) on synaptosome to identify protein-protein interactions that are unique to this highly complex and specialized organelle. Pairing mass spectrometry techniques with structure prediction AlphaFold-Multimer, we confirm known PPIs as well as identify novel PPIs that are highly conserved in the vertebrate brain.

P14.08 | Tissue Specific Interactome of the NPC1 Cholesterol Transporter

Roshan Javanshad[1] (roshanj@uic.edu), Thu T. Nguyen[1], Stephanie M. Cologna[1]
[1]University of Illinois Chicago, Chicago, IL

Niemann-Pick disease, type C (NPC) is a fatal, genetic, neurodegenerative disease. Mutations in the membrane-associated NPC cholesterol transporter 1 protein (NPC1) is responsible for 95% of diagnoses and results in cholesterol storage in the lysosome. While considerable effort has been focused on understanding the mechanisms of downstream events to the primary genetic defect in NPC, much less has been revealed about the NPC1 protein interactions beyond its role in cholesterol transport. In the present work, we aim to investigate protein-protein interactions of NPC1 to provide a deeper understanding of the possibility that this protein

has multiple functions, and to shed light on the lysosomal cholesterol efflux mechanism. To this end, we have determined the endogenous interactome of NPC1 in the cerebellum and liver, as these organs are the most affected by NPC disease. This study was done using the cerebellum and liver tissues of 9-week Balb/c *Npc1*^{nih} (*Npc1*-null) mouse model. Immunoprecipitation was performed using the Thermo Co-IP kit and the Abcam NPC1 primary antibody (ab134113). Eluents were digested using the S-Trap method and analyzed by nano-LC-MS/MS. Proteins were identified using Thermo Proteome Discoverer. Identifications were filtered to include proteins with ≥2 peptides and presence in all biological replicates of only wild type mice.

We observed more than 30 and 40 potential NPC1-interacting proteins in the cerebellum and liver, respectively, with one protein, GTP-binding nuclear protein Ran, found in both tissues. Several proteins were computationally predicted to interact with NPC1. Due to the high number of cytoskeletal and ribosomal proteins as possible interactors, we hypothesize NPC1 may interact with the lysosomal transport protein scaffold and affect ribosomal activity. Current efforts are focused on using immunostaining to examine the co-localization of NPC1 and potential interactors as well as to establish functional aspects of these new interactions.

P14.09 | Benchmarking Biotinylation Enrichment Methods for Chemical Biotinylation and Proximity Labeling Proteomics

Haorong Li[1] **(haorong@gwu.edu)**, Naomi Ni[1], Ling Hao[1]
[1]*George Washington University, Washington, DC*

Proximity labeling (PL) techniques characterize both stable and transient protein interactions and subcellular microenvironments *in vivo*. Proteins within a 10-20 nm radius of the bait protein can be biotinylated during the PL reaction. A crucial step in PL-proteomics is the enrichment of biotinylated proteins, commonly achieved by streptavidin-coated beads followed by on-beads tryptic digestion. Several alternative methods have been developed to enrich biotinylation. Here, we benchmark these enrichment methods and provide practical suggestions that can benefit both chemical and enzymatic PL-based biotinylation experiments.
Biotinylated proteins can be enriched at the protein level or peptide level. The three commonly used reagents to enrich biotinylation are streptavidin, neutravidin, and anti-biotin antibody. We systematically compared the performance of these three methods at both protein/peptide levels, regarding their enrichment efficiency, non-specific bindings, cost, and feasibility. To unambiguously evaluate enrichment efficiency and non-specific bindings, we incorporated a two-proteome model, where yeast-origin proteins/peptides were labeled with N-hydroxysuccinimidobiotin (NHS-biotin), then spiked in unlabeled proteins/peptides from human embryonic kidney (HEK) cells (Yeast: HEK = 1:50). This spike-in mimicked the low abundant biotinylated proteins and interference from unlabeled proteins in proximity labeling experiments. Moreover, we recently developed thiol-cleavable biotin as a novel probe for PL to reduce streptavidin contamination caused by on-bead digestion (Li *et al.*, 2021). The enrichment of thiol-cleavable biotinylated proteins/peptides was also benchmarked.
We also demonstrated that parallel multi-enzyme digestion dramatically improved biotinylated protein coverage and biotinylation site identification. Both biotin ligase-based PL and NHS-based labeling modify the lysine side chain, which impedes tryptic digestion efficiency. We digested biotinylated proteins with six enzymes in parallel, including Trypsin, Lys-C, Asp-N, Glu-C, Chymotrypsin, and Arg-C. Among 1070 commonly identified proteins, the median sequence coverage improved from 36% (trypsin only) to 65% (multi-enzyme). More importantly, we improved biotinylated protein identification three-fold from 428 (trypsin only) to 1263 (multi-enzyme).

P14.10 | Efficient isolation of RNA-protein complexes from formaldehyde-crosslinked cells for the characterization of the RNA-binding proteome

Samuel W. Markovich[1] **(smarkovich@wisc.edu)**, Rachel Knoener[2], Brian Frey[1], Lloyd M. Smith[1]
[1]*University of Wisconsin-Madison, Madison, WI,* [2]*Promega Corporation, Madison, WI*

Characterization of the RNA-binding proteome is important to understanding interactions between RNAs and proteins. A methodology capable of isolating RNA-protein complexes, Orthogonal Organic Phase Separation (OOPS), was developed for UV-crosslinked cells. However, UV-crosslinking is inefficient and does not reveal entire protein complexes, leading to increased sample requirements and missed protein identifications. We addressed these issues by reoptimizing OOPS parameters for formaldehyde-crosslinked cells. The resultant workflow, which we refer to as OOPS-FX, consistently recovers up to 90% of total RNA while retaining 30% of the protein. OOPS-FX contains several steps that differ from the original OOPS method to ensure RNA-protein complex purity. Preliminary results suggest that we will be able to estimate cellular crosslinking efficiencies, which are currently unquantifiable. Due to the increased efficiency of OOPS-FX, ten-fold fewer cells are needed to study the RNA-binding proteome compared to the original OOPS method, enabling the study of low quantity, precious samples. This method will be employed with our previously developed technique, Hybridization Purification of RNA-protein complexes followed by Mass Spectrometry (HyPR-MS). HyPR-MS is an RNA-centric proteomics approach, allowing the experimentalist to determine the protein interactors of specific RNAs present in cell lysate. The HyPR-MS protocol currently requires a large amount of starting material due to inconsistent and nonspecific captures of targeted RNAs, yielding low protein quantities for analysis. We hypothesize that the simpler sample OOPS-FX generates will lead to better capture efficiencies and specificities for targets of HyPR-MS and other similar approaches,

while lowering the cost of each experiment due to the reduced sample volume requirement. Ultimately, the adapted HyPR-MS protocol employing OOPS-FX will be applied to identify diagnostic markers in prostate cancer tumors.

P14.11 | Tapioca – A New Machine Learning Framework for Studying Global Protein-Protein Interactions Identifies NUCKS1 as a Broad-Spectrum Proviral Factor

Tavis J. Reed[1] (tjreed@princeton.edu), Matthew Tyl[1], Alicja Tadych[1], Olga G. Troyanskaya[1], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ*

The global dynamics of protein-protein interactions (PPIs) provide signatures of healthy and perturbed cellular states. For example, the progression of viral infections in host cells induces temporal alterations in host-host and virus-host PPIs on a systems level. Here, we describe *Tapioca*, a novel machine-learning framework for studying global PPI dynamics. *Tapioca* integrates dynamic thermal proximity coaggregation (TPCA) data with tissue-specific functional networks and protein properties, allowing it to outperform prior analytical methods. Additionally, we revisit the TPCA experimental workflow and perform a series of optimizations to define cell lysis and thermal denaturation conditions that improve PPI prediction and the ability to study transmembrane proteins. Next, we leverage our optimized experimental and *Tapioca* framework to characterize temporal virus-host and host-host PPIs during the reactivation from latency of the oncogenic gammaherpesvirus Kaposi's Sarcoma-associated herpesvirus (KSHV). *Tapioca*-driven functional analysis identified the dynamics of tens of thousands of novel interactions, providing a global picture of host biological processes modulated by KSHV. Further analysis led us to identify the nuclear ubiquitous casein and cyclin-dependent kinase substrate 1 (NUCKS1) as a hub protein that interacts with numerous host and KSHV proteins involved in innate immunity and viral genome replication. Follow-up IP-MS experiments validated NUCKS1 interactions with KSHV proteins RIR1/2, and knockdowns demonstrated its proviral function during KSHV replication. Next, we used *Tapioca* to integrate our TPCA datasets from KSHV with herpes simplex virus type I (HSV-1; an alphaherpesvirus) and human cytomegalovirus (HCMV; a beta-herpesvirus) infections. Using this integration, we predicted and then experimentally confirmed a broad-spectrum proviral role for NUCKS1. Additionally, we identified HSV-1, HCMV, and KSHV proteins with correlated temporally regulated interactomes, starting to assign function to uncharacterized viral proteins. Altogether, this work represents a significant advancement in our ability to study PPI dynamics at a global level and new insights into herpesvirus infections.

P14.12 | High-throughput and native detection of host-pathogen protein complexes by SEC-MS

Andrea Fosatti[1], Adrian Pelin[1], Nevan Krogan[1], **Danielle Swaney[1] (daniswan@gmail.com)**
[1]*UCSF, San Francisco, CA*

Host-pathogen interactions (HPI) play a major role in virulence, transmission, disease severity, and holds the potential to reveal novel therapeutic targets. Traditionally, HPIs have been studied by affinity-purification mass spectrometry (AP-MS), which involves the exogenous expression of individual pathogen proteins in host cells. AP-MS profiling of large pathogens, such as Poxviruses (~250 proteins) lacks scalability due to the necessity to generate hundreds of constructs and perform thousands of purifications. Furthermore, AP-MS limits our ability to measure HPIs under native expression levels, and how these HPIs are regulated in the context of the full pathogen protein repertoire during an infection time course. To overcome this, we have used size-exclusion chromatography mass spectrometry (SEC-MS) to probe the infection interactome in a systematic and unbiased fashion. As a proof-of-principle we have applied this approach to Vaccinia virus, due to its high sequence similarity (>85%) with an emerging global health threat, Monkeypox, and Vaccinia's large genome size which encodes several hundred proteins. In only 24 hrs of MS acquisition time, we were able to detect >200 Vaccinia proteins (~80% of the proteome) and identify > 50,000 HPIs from infected A549 cells using a DIA-PASEF approach. This comprehensive analysis both identified novel viral complexes as well as recapitulated known Vaccinia HPIs. Furthermore, this unbiased approach allows us to not only query HPIs, but also to quantify protein complex remodeling of the host proteome upon infection, which has revealed both novel rewiring and amplified existing knowledge of the role of Vaccina in interferon signaling and endosomal remodeling.

To summarize, we present a rapid and simple approach for the generation of deep host-pathogen interaction networks for large pathogens and showcase the deep insights that can be gained by this approach through its application to Vaccinia virus infection.

P14.13 | The BioGRID Database: A Curated Resource for Protein, Chemical and Genetic Interactions and CRISPR-based Phenotypic Screens

Jennifer M. Rust[1] **(jrust@princeton.edu)**, Rose Oughtred[2], Bobby-Joe Breitkreutz[3], Chris Stark[3], Lorrie Boucher[3], Christie Chang[2], Sonam Dolma[3], Genie Leung[3], Nadine Kolas[3], Frederick Zhang[3], Jasmin Coulombe-Huntington[4], Andrew Chatr-aryamontri[4], Kara Dolinski[2], Mike Tyers[5]
[1]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA, Colorado Springs, CO, [2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA, Princeton, NJ, [3]The Lunenfeld-Tanenbaum Research Institute, Toronto, Ontario, Canada, Toronto, ON, [4]Institute for Research in Immunology and Cancer, Université de Montréal, Quebec, Canada, Montreal, QC, [5]The Lunenfeld-Tanenbaum Research Institute, Toronto, Ontario, Canada; Institute for Research in Immunology and Cancer, Université de Montréal, Quebec, Canada; The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada, Toronto, ON

BioGRID (the Biological General Repository for Interaction Datasets, www.thebiogrid.org) is an open access database that curates genetic, protein, and chemical interaction data from the primary biomedical literature for humans and the major model organisms. As of December 2022, BioGRID contains over 2,579,500 interactions captured from more than 70,000 publications. In parallel, BioGRID has curated 1,128,339 post-translational protein modifications. BioGRID also curates an extensive collection of single mutant phenotype screens generated by genome-wide CRISPR screens. These high throughput CRISPR-derived datasets are available through the Open Repository for CRISPR Screens (ORCS) (orcs.thebiogrid.org), which currently houses 1,678 genome-wide mutant phenotype screens performed in human and model organism cell lines, as curated from almost 300 publications. BioGRID and ORCS data are disseminated through model organism database (MOD) collaborations including Saccharomyces Genome Database (SGD), WormBase, FlyBase, and PomBase, all of which are members of the Alliance of Genome Resources (www.alliancegenome.org). BioGRID data is also disseminated through key meta-database resources including UniProt, NCBI, and PubChem.To prioritize curation of the vast primary literature for human cells and to maximize the impact of this data, BioGRID undertakes comprehensive curation on central biological processes and diseases of interest. These data are provided through dedicated project pages for the Ubiquitin Proteasome System (UPS) and COVID-19, among others. Additionally, BioGRID curates drug and other bioactive compound interactions and currently houses 30,725 chemical interactions, in part imported through partner resources DrugBank and BindingDB. BioGRID provides unique visualization tools for the exploration of these interactions by users. BioGRID is updated on a monthly basis and all data in BioGRID and ORCS are freely available without restriction from the BioGRID search pages, download pages, and REST API. This project is supported by the National Institutes of Health Office of Research Infrastructure Programs [R01OD010929 to MT, KD].

P14.14 | Glycan-dependent Affinity Purification Mass Spectrometry (glyAP-MS) provides novel insights into glycoprotein interaction network

Yixuan (Axe) Xie[1] **(axexie@wustl.edu)**, Xingyu Liu[1], Siyu (Cathy) Chen[2], Carlito B. Lebrilla[2], Benjamin A. Garcia[1]
[1]Washington University School of Medicine in St. Louis, St. Louis, MO, [2]University of California, Davis, Davis, CA

The cell membrane is composed of a layer of the glycocalyx, which forms a highly interactive network. The glycoprotein network is investigated using several strategies, such as proximity labeling, cross-linking mass spectrometry (XL-MS), and affinity purification mass spectrometry (AP-MS). However, the role of glycan in mediating such interactions is barely considered due to the limitation of the methods. Herein, we report a novel method termed "glycan-dependent affinity purification mass spectrometry (glyAP-MS)" to study the glycoprotein interactome on the cell surface. This method offered valuable information on the overlooked role of glycans in glycoprotein interactions. We first used a toolbox of glycan mediators to treat colon cancer cells (HCT116) to generate different a variety of "glycan phenotypes." The resulting changes in glycoproteins were confirmed using glycomic and glycoproteomic analyses. Subsequently, we implemented this approach to investigate the glycan-dependent interactions of CD147 (Basigin) with an orthogonal HaloTag as a proof of concept. The interacting partners of CD147 were enriched by performing affinity purification followed by data-independent acquisition mass spectrometry (DIA-MS) to identify the interacting proteins. Overall, a comprehensive network of CD147 was revealed with the dimension interactions dependent on glycans. This is the first study that provides systematic insight into glycan-dependent cell surface glycoprotein interactomes. We expect this approach to elucidate a novel landscape of cell surface glycoprotein networks that can be applied to many other biological systems.

P15: Proteomics in Immunology and Infectious Diseases (COVID-19)

P15.01 | Conserved Coronavirus Nonstructural Protein Interactions with ER Proteostasis Factors Mediate Viral Replication

Jonathan Davies[1] (jonathan.p.davies@vanderbilt.edu), Lars Plate[1]
[1]Vanderbilt University, Nashville, TN

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 both shared and distinct infection mechanisms among CoVs is needed. Previously, we used quantitative comparative proteomics to profile the virus-host protein-protein interactions of nonstructural protein 2 and 4 (nsp2, nsp4) homologs from five different CoVs: SARS-CoV-2, SARS-CoV, MERS-CoV, hCoV-OC43, and MHV. Using a functional genetic screen, we have further characterized host interactors as pro- or anti-viral factors for CoV infection. We identified several conserved, pro-viral interactors associated with protein biogenesis, ER membrane remodeling, and ER-mitochondria contact sites. Further functional assays using infection-based proteomics and reporter viruses reveal these factors impact viral protein biogenesis and mediate early replication. Our results shed light on new roles for nsps in modifying host proteostasis processes during infection and identify new host factor-dependencies which may be shared across coronaviruses.

P15.02 | Ig-MS analysis of antibody repertoires produced in response to SARS-CoV-2 infection and/or vaccination in immunocompetent and immunosuppressed people

Eleonora Forte[1] (e-forte@northwestern.edu), Benjamin J. Des Soye[1], Rafael D. Donadelli Melani[1], Michael A. Hollas[2], Jiana A. Duan[1], Steven Patrie[1], Jared O. Kafader[1], Beverly E. Sha[3], Jeffrey R. Schneider[3], Daniela Ladner[1], Neil L. Kelleher[1]
[1]Northwestern University, Evanston, IL, [2]Proteomics Center of Excellence, Northwestern University, Evanston, IL, [3]Rush University, Chicago, IL

Ig-MS is a recently developed mass spectrometry-based serology platform that can characterize the total repertoire of antibodies against a specific antigen of interest with single-proteoform resolution. The platform workflow consists of an immunoprecipitation step to enrich antibodies that specifically interact with the antigen followed by individual ion mass spectrometry (I^2MS) analysis that can accurately analyze diluted and complex protein samples to generate a multiparametric readout with metrics that capture antibody repertoire complexity and relative abundances of individual antibody clones. In our first study, we applied Ig-MS to analyze the antibody response against the receptor-binding domain (RBD) of the viral Spike protein from SARS-CoV-2 in both COVID-19 convalescent patients and vaccinated individuals. This analysis showed that fully vaccinated individuals have an antibody response similar to hospitalized patients with severe COVID-19.
Here we used Ig-MS to determine how antibody repertoires evolve in response to the COVID-19 vaccine booster and natural viral infection with SARS-CoV-2 Omicron variant. We found that 1) a few antibodies generated against the Wuhan-RBD interact with the Omicron-RBD variant; 2) the booster increases antibody titers against both Wuhan- and Omicron-RBDs and elicits an Omicron-specific response; and 3) vaccination and infection act synergistically in generating anti-RBD antibody repertoires able to bind both Wuhan- and Omicron-RBDs with variant-specific antibodies.
Furthermore, we applied Ig-MS to COVID-19 transplant recipients to determine whether immunosuppression or the kind of organ received could affect the antibody response compared to the immunocompetent group. We found no statistically significant differences between the groups both in longitudinal and single-point analyses, despite differences in the individual antibody repertoire composition. However, differences occurred at the level of glycosylation.

P15.03 | Comparative Interactome Profiling of Nonstructural Protein 3 from Emerging SARS-CoV-2 Variants

Valeria Garcia Lopez[1] (valeria.a.garcia.lopez@vanderbilt.edu), Lars Plate[1]
[1]Vanderbilt University, Nashville, TN

SARS-CoV-2 continues to be a threat to global health due to ongoing mutations resulting in new strains. SARS-CoV-2 variants can display differences in virulence of COVID-19 and can result in reduced protection from vaccines. Significant progress has been made to identify how variants of SARS-CoV-2 structural proteins lead to molecular changes. However, investigating variants of

nonstructural proteins is important to understand the viral evolution and manipulation of protein-protein interactions within a host. Here, we focus on nonstructural protein 3 (nsp3), a viral protease that is implicated in host de-ubiquitylation and ISGylation. Affinity purification – mass spectrometry is used to identify the proteome of emerging variant mutations in two truncations of nsp3. Unique interactors are identified for each nsp3 fragment, with few overlapping interactors. The N-terminal fragment (nsp3.1) shows differing enrichment for RNA binding proteins between variants, including FXR1. The membrane and luminal fragment (nsp3.2) shows alternative variant enrichment for mitochondrial membrane components and cholesterol synthesis factors. Tracking the interaction changes resulting from nsp3 sequence variations is relevant for understanding the evolutionary arms race between host and viruses and for prioritization of cellular pathways as targets for antiviral therapeutics.

P15.04 | Utility of proteomic trajectories of cardiovascular risk and cardiorespiratory fitness to monitor adverse health states throughout post-COVID-19

Michael Hinterberg[1] (mhinterberg@somalogic.com), Clare Paterson[1], Jessica Kuzma[1], Andrew Morrow[2], Robert Sykes[2], Kenneth Mangion[2], Paul Welsh[2], Alex McConnachie[2], Naveed Sattar[2], Colin Berry[2], Stephen Williams[1]
[1]SomaLogic, Inc., Boulder, CO, [2]University of Glasgow, Glasgow, United Kingdom

Background: Poorer cardiovascular (CV) and cardiorespiratory (CR) health are risk factors for adverse outcomes from COVID-19, with a proportion of COVID-19 survivors exhibiting CV and CR dysfunction after initial recovery. However, established biomarker-driven predictors of CV and CR outcomes specific to COVID-19 are lacking. The current study determines if validated proteomic tests of CV risk (27 proteins) and CR fitness (52 proteins) associate with longer-term adverse health states (adjudicated myocarditis and rehospitalization) in COVID-19 survivors.
Methods: The SomaScan® platform was used for plasma proteomic phenotyping in CISCO-19 trial participants (n=154, mean age 55y, 43% Female), including protein model assessments of CV risk and CR fitness. Samples were taken at hospital discharge following COVID-19, and within 6-months post-discharge when comprehensive CV assessment including myocarditis adjudication was performed. N=20 (13%) were adjudicated as very likely to have myocarditis at follow-up. Complete clinical follow-up was achieved in all participants with n=23 patients being readmitted to hospital during longer-term follow-up post-COVID-19 (mean 419 days). We used linear mixed models to test proteomic CV and CR fitness scores at both visits against adjudicated myocarditis likelihood and rehospitalization across COVID-19 recovery.
Results: A 36% relative CV risk reduction and 5.6% CR fitness improvement was observed during short-term COVID-19 recovery. CV risk and CR fitness scores were significantly worse among those with very likely myocarditis at follow up ($p<0.05$ for effects). CV risk scores but not CR fitness scores were similarly significantly associated with longer-term rehospitalization, whereby COVID-19 survivors who were rehospitalized had higher CV risk scores at discharge and follow-up.
Conclusions: Multiplexed protein-based indicators of cardiovascular health have potential to support screening and monitoring of health states across the trajectory of COVID-19 recovery, demonstrating proteomic dynamics in the weeks and months following COVID-19, and with CV risk additionally capable of monitoring the likelihood of long-term

P15.05 | Deep, Unbiased and Quantitative mass spectrometry-based plasma proteome analyses of adaptive response to COVID-19 vaccine

Ting Huang[1] (thuang@seer.bio), Alexandre R. Campos[2], Ramón D. Peña[3], Svetlana Maurya[4], Khatereh Motamedchaboki[1], Laura R. Oliveira[5], Camila I. Alves[5], Rafael L. Rosa[5], Yohana P. Alves[5], Lucélia Santi[5], Walter O. Beys-da-Silva[5]
[1]Seer Bio, Redwood City, CA, [2]Plexium, San Diego, CA, [3]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, [4]Sanford Burnham Prebys, San Diego, CA, [5]Federal University of Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

Introduction: COVID-19 vaccines have been extensively used to immunize a large worldwide population. However, individuals respond to the vaccine differently, leading to distinct proteome changes and vaccination efficiency. At the same time, our understanding of the underlying molecular mechanisms behind the adaptive response to COVID-19 vaccine is still limited. Here, we used the ProteographTM workflow coupled with TMT labeling to quantify the proteome changes in 12 volunteers immunized with two doses of COVID-19 vaccine and gain insights about the personalized response.
Method: Plasma samples were collected before and after first and second vaccination from each volunteer. Samples were processed by Proteograph Product Suite (Seer Inc.) using five distinct nanoparticles. Tryptic peptides were then pooled into one single sample for TMT labeling. A total of 48 samples were allocated into three 18-plex TMT mixtures, followed by peptide fractionation by high pH RP and LC-MS/MS analysis, comprised of a Proxeon EASY nanoLC system coupled to an Orbitrap Fusion Lumos MS equipped with FAIMS Pro Interface (Thermo Fisher Scientific). Peptides were separated using an analytical C18 Aurora column (IonOpticks) into 24 fractions. A two-hour gradient with three FAIMS-Pro compensation voltages was used in this study. The raw spectra data were analyzed by SpectraMine (Biognosys) and MSstatsTMT to generate the normalized protein abundances.
Results: Hierarchical clustering of data showed that samples from the same volunteer were grouped together and demonstrated a high inter-individual variation. We found that 69 proteins were differentially expressed after second vaccination. Pathway enrichment analyses showed that 131 canonical pathways and 127 gene ontology biological processes pathways were significantly enriched.
Conclusion: Compared with the existing plasma studies about COVID-19 vaccine, the combination of Proteograph workflow with

TMT 18-Plex has provided a much deeper access to plasma proteome, i.e., 3,094 proteins, and revealed insights on inter-individual variation of response to COVID-19 vaccination.

P15.06 | Mechanistically defining the function of mitochondrial sirtuin 4 in metabolic remodeling and host defense during viral infection

Cora N. Betsinger[1], **James Kostas[1] (jkostas@princeton.edu)**, Elizabeth Rowland[1], Ileana Cristea[1]
[1]*Princeton University, Princeton, NJ*

Alterations of mitochondrial functions and cellular metabolism are hallmarks of all viral infections, as viruses rely on mitochondria to produce the metabolites and energy necessary for generating new virus particles. We previously discovered that the mitochondrial enzyme sirtuin 4 (SIRT4) is a broad-spectrum antiviral factor against DNA and RNA viruses. Among these, SIRT4 protects against human cytomegalovirus (HCMV), a beta-herpesvirus that globally alters cellular metabolism. Initially thought to be a deacetylase and ADP-ribosyltransferase, we discovered that SIRT4 is a lipoamidase. SIRT4 removes the rare but essential posttranslational modification lipoylation from the pyruvate dehydrogenase complex (PDH), inhibiting its activity and regulating carbon entry into the TCA cycle. Here, we integrated proteomics, metabolomics, and molecular virology to characterize SIRT4 antiviral function and regulation. We first monitored the SIRT4-mediated regulation of PDH lipoylation and activity throughout the HCMV replication cycle. By developing and applying a parallel reaction monitoring (PRM) assay targeting all lipoylated enzyme complexes, we found that SIRT4 lipoamidase activity is not exclusive to PDH but that it also targets additional lipoylated complexes. This expands our knowledge of enzyme complexes and metabolic processes regulated by SIRT4. Using molecular virology and genetic perturbations, we discovered that, while SIRT4 knockdown increases HCMV replication, SIRT4 overexpression has no impact on replication, suggesting that SIRT4 antiviral activity is inhibited during infection. Immunoaffinity purification mass spectrometry of temporal SIRT4 interactions led to identification of several candidate viral proteins as potential inhibitors of SIRT4. Follow-up functional analyses via generation of HCMV deletion strains combined with LC-MS-based metabolic flux and enzyme activity assays uncovered a mechanism acquired by HCMV to prevent SIRT4-mediated antiviral metabolic regulation. Altogether, these findings highlight the pivotal role of mitochondrial metabolic regulation in host defense during viral replication and uncover the interplay between SIRT4 and HCMV for control over this axis of cellular function.

P15.07 | IFI16 phase separation via multi-phosphorylation drives innate immune signaling

Dawei Liu (daweiliu@princeton.edu), Princeton University, Princeton, NJ

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) is a key sensor of foreign DNA in the nucleus. Its function is critical against nuclear-replicating viruses, such as herpes simplex virus 1 (HSV-1) and human cytomegalovirus, and even against retroviruses (HIV-1). During 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 suppressing viral gene expressions. However, little is known about mechanisms that initiate IFI16 antiviral functions or its regulation within the host DNA-filled nucleus. Here, we provide *in vitro, ex vivo,* and *in vivo* evidence to establish that IFI16 undergoes liquid-liquid phase separation (LLPS) nucleated by viral DNA. We show for the first time that IFI16 forms micrometer-sized liquid droplets that develop into filaments in cells and *in vitro*. We further demonstrate that LLPS governs IFI16 binding to viral DNA and cytokine induction during HSV-1 infection. Using *Ciona Intestinalis*, we provide *in vivo* evidence that IFI16 LLPS is regulated by an intrinsically disordered region (IDR). Using immunoaffinity purification, parallel reaction monitoring, and single-point mutants, we identify multiple IDR phosphorylation sites that act as a code to combinatorially activate IFI16 LLPS. Biochemical fractionation of the nuclear periphery, quantitative mass spectrometry, and motif analyses provided a temporal-spatial view of kinases early in infection. Validation, *in vitro* and functional assay led us to discover CDK2 and GSK3b as the major kinases that phosphorylate IFI16 IDR, providing a mechanism to toggle between active and inactive IFI16. These findings show how IFI16 switch-like phase transitions are achieved with temporal resolution for immune signaling and, more broadly, the multi-layered regulation of nuclear DNA sensors.

P15.08 | COVID-19: The turmoil inside the Brain

Sudipa Maity[1] (smaity@tulane.edu), Meredith Mayer[1], Qingbo Shu[1], Duran Bao, Robert Blair[2], Yanlin He[3], Christopher Lyon[1], Tony Hu[1], Tracy Fischer[2], Jia Fan
[1]Tulane University, New orleans, LA, [2]Tulane Primate Research Centre, Covington, LA, [3]Pennington Biomedical Research Centre, Baton Rouge, LA

Neurologic manifestations are among the most frequently reported complications of COVID-19. However, given the paucity of tissue samples and highly infectious nature of the etiologic agent of COVID-19, we have limited information to understand the neuropathogenesis of COVID-19. Therefore, to better understand the impact of COVID-19 in brain, we employed mass-spectrometry-based proteomics using data-independent acquisition mode (DIA) to investigate cerebrospinal fluid (CSF) proteins collected from two different non-human primates, Rhesus Macaque and African Green Monkeys for the neurologic effects of the infection. These monkeys exhibited minimal to mild pulmonary pathology but moderate to severe CNS pathology. Our results indicate that CSF proteome changes after infection resolution correspond with bronchial virus abundance during early infection and revealed substantial differences between the infected NHPs and their age-matched uninfected controls, suggesting these differences could reflect altered secretion of CNS factors in response to SARS-CoV-2-induced neuropathology. We also observed the infected animals to exhibit much scattered data distributions as compared to the tightly clustered corresponding controls which suggest the heterogeneity of the CSF proteome change and the host response to the viral infection. In addition, dysregulated CSF proteins were preferentially enriched in functional pathways associated with progressive neurodegenerative disorders, hemostasis and innate immune responses that could influence neuroinflammatory responses following COVID-19. Mapping dysregulated proteins to the Human Brain Protein Atlas found that these proteins tended to be enriched in brain regions that exhibit more frequent injury following COVID-19. It therefore appears reasonable to speculate that such CSF protein changes could serve as signatures for neurologic injury, identify important regulatory pathways in this process, and potentially reveal therapeutic targets to prevent or attenuate the development of neurologic injuries following COVID-19.

P15.09 | diaPASEF and Multi-Omic Analysis Reveal BRD4-driven Regulation of Alternative Splicing in Respiratory Syncytial Virus-Induced Innate Inflammation

Morgan W. Mann[1] (mwmann@wisc.edu), Yao Fu[1], David S. Roberts[1], Ying Ge[1], Allan R. Brasier[1]
[1]University of Wisconsin-Madison, Madison, WI

Respiratory Syncytial Virus (RSV), an airway virus that causes severe inflammation and airway pathology, represents the primary cause of pediatric hospitalizations, with cases surging in the wake of the COVID-19 pandemic. RSV infects the epithelial cells of the airway, resulting in toll-like receptor and NF-kB mediated innate inflammation. Importantly, NF-kB RelA/p65 translocates into the nucleus and interacts with BRD4 generating a protein complex triggers the coordinated expression of pro-inflammatory chemokines and interferons. However, the innate immune response is highly pleiotropic and stimulus specific, and the mechanisms by which the NF-kB:BRD4 complex control this pathway remain unclear. BRD4 inhibitors (BRD4i) have been shown to block virus-induced inflammation and fibrosis, implicating BRD4 as a key driver of innate activation. Moreover, recent works have demonstrated that BRD4 dynamically interacts with mRNA splicing factors in the context of RSV infection. Here we employ an integrated multi-omics approach to comprehensively characterize the role of BRD4-mediated alternative splicing in the regulation of innate inflammation. In this study, we utilize high-throughput RNA-sequencing to identify ~1500 RSV- and BRD4i-induced alternative splicing events in cultured airway epithelial cells. Data-Independent Analysis – Parallel Acquisition Serial Fragmentation (diaPASEF) was applied using a timsTOF Pro mass spectrometer (Bruker Daltonik) to quantify ~8300 proteins and identify splicing events with protein-level consequences. We highlight BRD4i-induced isoform switches to three genes with established roles in innate inflammation (ATP11A, IFRD1, XBP1), and explore their roles in cytokine activation using q-RT-PCR assays. We further leverage our combined proteomic/transcriptomic datasets to demonstrate that BRD4 controls the protein-level abundance of splicing factors, including the endoplasmic reticulum-bound endoribonuclease IRE1alpha. IRE1alpha uniquely splices full-length XBP1 transcripts into a short variant which codes for a potent transcriptional regulator of cytokine expression and the unfolded protein response. We conclude that BRD4 modulates innate activation in airway viral infection through regulation of alternative splicing.

P15.10 | Proteomic analysis of undepleted plasma from free-ranging cheetahs (Acinonyx jubatus)

Benjamin A. Neely[1] (benjamin.neely@nist.gov), Gábor Á. Czirják[2], Roland Bruderer[3], Claudia Moresi[4], Lukas Reiter[3], Bettina Wachter[5]
[1]NIST Charleston, Charleston, SC, [2]Department Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany, [3]Biognosys AG, Schlieren, Shiselweni, Swaziland, [4]Biognosys, Zurich, Switzerland, [5]Department Evolutionary Ecology, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

Plasma proteomics frequently relies on depletion techniques, but in some situations this is not possible (such as non-human systems). In these cases, state of the art data acquisition and analysis is required to overcome the challenge of the dynamic range of undepleted plasma. Previous studies of free-ranging cheetahs (Acinonyx jubatus), such as blood hormone levels, have proved inconclusive as to why fully grown territorial males have a significantly higher body mass index, based on a higher muscle mass, than non-territorial males. As an exploratory study, undepleted plasma from 45 free-ranging cheetahs from Namibia was analyzed using data independent acquisition (DIA) on a Thermo Orbitrap Exploris 480. The adult cheetahs were captured in box traps and blood samples taken from 10 females, 15 territorial and 15 non-territorial males ("floater"). A hybrid library approach using Spectronaut 16 was employed wherein eight fractions from a high pH separated plasma pool were analyzed using data dependent acquisition, and this data augmented directDIA analysis of the DIA data. We were able to identify 1367 protein groups across the experiment. Preliminary analysis identified 11 differentially abundant proteins between territorial and floater males, including elevated ORM1 in floater males, which is linked to improved muscle endurance, and elevated LECT2 in territorial males, which may help explain weight gain. Interestingly, the most pronounced change was a -5.7-fold decrease of IL1R2 in territorial males, possibly indicating improved immune function. This analysis is ongoing, but we expect continued insights into molecular changes driving the change in male phenotype, as well as improve our understanding of cheetah biology in general.

P15.11 | Phylogenetic Proteomics to Identify Species of Origin for Big Cat Fur Products

Brett Phinney[2] (bsphinney@ucdavis.edu), Nicole Nicole Slattengren[2], Ashley Spicer[1], Glendon Parker[3]
[1]California Department of Fish and Wildlife, Davis, CA, [2]University of California Davis, Davis, CA

Wildlife trafficking is a global issue with devastating ecological effects. The fur and pelt trade is a major component of wildlife trafficking, where many species are overharvested and threatened with extinction. The ability to prosecute poachers is dependent on the methods to include or exclude illegal versus legal species, yet fur DNA is highly during the chemically harsh production process. Protein however is considerably more stable and also contains phylogenetic information in the form of species-specific amino acid changes in protein sequences. These changes accumulate over evolutionary distance and result in family, genera, and species-specific sequences. Reference proteomes therefore become less efficient at assigning peptide spectra matches as evolutionary distance increases. In this study we demonstrate that this phenomenon can be used to resolve closely related species. Fur from four Felidae species (Tiger, Lion, Leopard and Cheetah) were processed, applied to an Exploris 480 and the resulting proteomes searched in PEAKs Software using Reference proteomes from Felidae and other mammalian species (Tiger, Lion, Leopard, Cheetah, Lynx, Domestic cat, Hyaena, Mouse and Platypus). In each case the reference proteome with the least evolutionary distance to the actual fur sample was assigned the most PSMs, including the closely related Lion and Leopard species (3.8 MYa). The process also identifies specifies-specific or -informative markers. To this date 48 candidates have been identified for development of resolution of Felidae species from processed fur alone.

P15.12 | Elucidation of the RNA-protein Interactomes of SARS CoV-2 Genomic and Subgenomic RNAs

Isabella T. Whitworth[1] (iwhitworth@wisc.edu), Rachel Knoener[2], Maritza Puray-Chavez[3], Peter J. Halfmann[1], M'bark Baddouh[1], Sofia Romero[1], Mark Scalf[1], Yoshihiro Kawaoka[1], Sebla B. Kutluay[3], Lloyd M. Smith[1], Nathan M. Sherer[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Promega Corporation, Madison, WI, [3]Washington University School of Medicine, St Louis, MO

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an RNA virus that has a 30-kb genome that produces 9 subgenomic RNAs and 28 viral proteins formed through discontinuous transcription of the genomic RNA (gRNA.) The viral life cycle is highly dependent on host cellular machineries that facilitate nearly all processes leading to virus particle production. Host proteins are also used by the cell to defend against virus replication. Hence, determining the SARS-CoV-2 RNA-protein interactome is integral to combatting the virus. 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. Here, we developed a novel strategy for elucidating the *in vivo* protein interactomes of SARS-CoV-2 gRNA and subgenomic RNAs that code for key structural proteins Spike (S) and Nucleocapsid (N), and the immunomodulatory accessory protein ORF8. Using HyPR-MS, we identified over 300 proteins from SARS-CoV-2 (Wuhan strain) infected Huh7 cells that associate with one or more of the SARS-CoV-2 gRNA, S, N, and ORF8 RNAs at 8 or 24 hours post infection (hpi). Among these were viral proteins N, S, and M interacting preferentially with the gRNA at 24 hpi and known viral replication antagonists APOBEC3F and ZC3HAV1 interacting with three of the four RNAs at both timepoints. Gene ontology enrichment analysis of the genomic, S, N, and ORF8 RNA protein interactomes showed enrichment of RNA binding proteins as well as proteins involved in splicing, RNA stability, gene silencing, and stress granule formation. Currently, siRNA knockdowns are underway on a subset of the identified proteins to evaluate knockdown effects on multiple aspects of the viral life cycle. Through this, we hope to identify key pathways used by the virus and establish the functional significance identified interactions.

P16: Proteomics of Diseases in Precision Medicine

P16.01 | Clinical applications of data independent acquisition mass spectrometry to support translational proteomics in drug development

Veronica G. Anania (ananiav@gene.com), Genentech, Inc., South San Francisco, CA

Biomarkers play an important role in the drug development process, including providing necessary insights into target engagement, dose selection, and mechanism of action for candidate therapeutics. Recent advances in data-independent mass spectrometry (DIA-MS) have made this technology more accessible and certain benefits of DIA-MS including reproducible label-free analysis of hundreds of samples, ability to capture low abundance ions over a high dynamic range, and deep proteome coverage makes this technology especially well suited to support translational proteomics. Here, results from two clinical biomarker studies will be presented. The first study investigated the fecal proteome of patients with Inflammatory Bowel Disease (IBD) to identify disease activity biomarkers that could be used to support clinical trial enrollment. This study identified >600 human proteins in IBD patient stool including some likely originating from infiltrating immune cells and other disease pathways. These proteins could serve as non-invasive biomarkers reflective of flare and may reduce the need for invasive procedures like colonoscopy. The second study explored the cerebrospinal fluid (CSF) proteome from patients with Multiple Sclerosis (MS) to understand residual disease progression in patients on high efficacy therapeutics. Here, >1500 proteins were measured in CSF from MS patients and comparisons with clinical disease activity metrics identified CSF proteins that may reflect oligodendrocyte or neuronal cell activation. These proteins may serve as valuable proof of activity biomarkers in studies using novel candidate therapeutics aimed at stimulating remyelination and repairing disease induced tissue damage. Altogether, these studies have demonstrated that DIA-MS methods generate high quality, robust results and utilizing DIA-MS has streamlined the process of going from biomarker discovery to hypothesis testing in large clinical cohorts. Furthermore, these studies have provided comprehensive proteomic information from valuable clinical trial samples that can be further interrogated against new biomarker hypotheses as knowledge in specific disease areas continues to evolve.

P16.02 | Spatial and Pathological Collagen Post-Translational Modifications in Hepatocellular Carcinoma Subtypes Defined By Clinical Outcomes

Jade Macdonald[1] (macdojad@musc.edu), Harrison B. Taylor[1], Andrew DelaCourt[1], Jake Griner[1], David Lewin[2], Christin Edge[1], Atsushi Ono[3], Takashi Nakahara[3], Hiroshi Hiroshi Aikata[3], Kazuaki Chayama[4], Noato Kubota[5], Noato Fujiwara[5], Fahmida Rasha[5], Cesia Marquez[5], Richard Drake[1], Yujin Hoshida[5], Anand S. Mehta[1], Peggi Angel[1]
[1]Medical University of South Carolina, Charleston, SC, [2]Medical University of South Carolina, Dept Pathology and Laboratory Medicine, Charleston, SC, [3]Department of Gastroenterology and Metabolism, Graduate School of Biomedical & Health Sciences, Hiroshima University, Hiroshima, Japan, Hiroshima, Japan, [4]Collaborative Research Laboratory of Medical Innovation, Research Center for Hepatology and Gastroenterology, Hiroshima University, Hiroshima, Japan RIKEN Center for Integrative Medical Sciences, Yokohama, Japan, Hiroshima, Japan, [5]Liver Tumor Translational Research Program, Simmons Comprehensive Cancer Center, Division of Digestive and Liver Diseases, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas, United States, Dallas, TX

Since 1980, hepatocellular carcinoma (HCC) incidence has tripled, while mortality has doubled in the US. Despite a surge of excellent research directed at diagnostic markers, there is no clear answer as to why, in a fraction of individuals, the disease results in an aggressive clinical phenotype with high mortality. Collagen deposition & composition play a role throughout the course of HCC. However, how collagen pathologically changes at the translational level and by post-translational modification (PTM) to advance disease remains undefined. Dr. Yujin Hoshida previously modeled HCC by transcriptomic patterns into three molecular subtypes (S) that differentiate clinically. Notably, S1 & S2 are characterized by poor outcomes, while S3 is characterized by favorable outcomes. To date, there are no investigations on the pathological collagen proteome contributing to subtype outcomes. Here, we hypothesize that collagen variation at the translational and post-translational level represents a novel, clinically significant contributor to HCC endpoints. HCC tissues subtyped as S1 or S3 (S1, n= 12; S3 n=24) were used to investigate collagen proteomic patterns relative to pathological annotation and outcome. Collagen-targeted mass spectrometry imaging (MSI) spatially mapped collagen peptide domains relative to pathology and by HCC subtypes. Proteomic sequencing by LC-MS/MS and IMS identified that most altered peptides were post-translationally modified and were from trypsin-resistant triple helical regions of fibrillar collagens. Collagen peptide PTM variations were uniquely linked to pathologist-annotated fibrosis, cirrhosis, and tumor regions. Multiple peptides could distinguish between subtype pathology by area under the receiver operating curve ≥ 0.7, p-value <0.01. S1 largely showed increases in tumor while S3 showed increases in fibrosis. Additional studies in cirrhosis and cirrhosis adjacent to tumor further illustrated potentially predictive differences by collagen PTMs. This study is the first to report potential pathological regulation of post-translational modifications from fibrillar collagen triple helical regions as linked to clinical outcomes.

P16.03 | Advanced glycation end products (AGE) PTM profiling on antigen processing machinery and MHC-II molecules in diabetes and T2DM syndrome

Cristina C. Clement (ccc4002@med.cornell.edu), Weill Cornell Medicine, New York, NY

Increased oxidized glucose, glyoxal, and methylglyoxal concentrations are observed in several metabolic diseases including diabetes and metabolic syndrome, mediating the non-enzymatic protein glycation followed by protein unfolding and aggregation. We investigated whether the protein modifications induced by glycation could affect the antigen processing and presentation by MHC I/II in PBMC from healthy and type-2 diabetics (T2D) patients and mice models for T2DM syndrome. As such, we purified PBMC from healthy and diabetic (T2D) patients, and CD11c+ dendritic cells from mice on high fat and high fructose diet (HFD), or from obese (Ob/Ob) mice. Profiling of AGE-PTM was performed by bottom-up label-free DDA/DIA proteomic analysis. Due to the selective AGE-targeted chemical modifications of Lys and Arg we employed multiple enzymatic digestions using a combination of LysC/trypsin/Glu-C/Asp-N that was previously shown to enable a higher sequence coverage, and better accuracy for PTM site identification in complex cellular proteomes. We identified and annotated AGE-PTM in each proteome using false discovery rate <2 % at the protein level, and <1% at the peptide level employing a combination of proteomic platforms. We observed a difference in the total number of unique AGE sites (485 versus 22) in the diabetic versus healthy controls from equal input of 10 ug tryptic peptides. Analysis of the type of AGE-PTM accumulated in the proteomes from T2D samples ranked the site-specific carboxymethylation of lysine (CML) as the most abundant AGE found in the glycated proteome, followed by formylation of Lys and carboxymethylation of Arg (CMA). Proteins involved in processing and MHC II peptide/loading were subject to AGEs modifications under conditions of elevated glucose, for humans in vivo, thus highlighting the potential role for accumulation of AGE-PTM in the in the pathophysiology of immune responses.

P16.04 | PEP-TORCH can illuminate mycobacterial identification

Jia Fan[1] (jfan5@tulane.edu), Sudipa Maity, Duran Bao[1], Qingbo Shu, Tony Hu, Adrian Zelazny[2]
[1]Tulane University, New Orleans, LA, [2]National Institutes of Health Clinical Center, Bethesda, MD

The incidence of human infections caused by mycobacteria species, a diverse family of environmental pathogens, has increased to become a major global public health concern. Rapid and accurate NTM identifications are crucial for effective care since treatment can be species-specific and standard methods can fail to discriminate closely related NTM taxa. We therefore developed and optimized an integrated sample processing method and mass spectrometry (MS)-based proteomics assay to provide comprehensive peptidomics information from filtered culture supernatants of clinical specimens to improve mycobacteria identification. This approach improved peptide identification rates in these samples by about 30% versus another highly sensitive method. We next developed an automated, algorithm-based (PEP)tide (T)axonomy/ (OR)ganism (CH)ecking (PEP-TORCH) pipeline PEP-TORCH that selects unique peptides or peptide combinations that permit species- and subspecies-specific mycobacteria identification. In a validation study performed with clinical samples, this method identified common clinical mycobacteria isolates (M. *tuberculsosis*, *M. abscessus* (*M.ab*), *M. kansasii*, *M. avium*, and *M. intracellulare*) with 100% accuracy. It also identified two *M.ab* subspecies (*M.ab. massiliense* and *M.ab. abscessus)* that are difficult to distinguish by standard methods but have different macrolide resistance phenotypes and require different treatments. To adapt this untargeted shotgun approach for clinical use, we next employed data from this study to select a set of species-specific peptide markers for each species that could be detected in targeted liquid chromatography−MS/MS assays more acceptable for use in clinical laboratories. Notably, our assay approach could detect species-specific peptides in a day 7 clinical culture which was not tagged positive until culture day 28 by a gold-standard method. Our preliminary data thus imply our method could substantially improve mycobacteria identification at the species and sub-species level to provide critical information required for appropriate treatment, and do so earlier than possible with current methods to reduce pathophysiology and improve outcomes.

P16.05 | Dysregulation of synaptic protein localization and protein phosphorylation in the human primary auditory cortex in schizophrenia

Vasin Dumrongprechachan[1] (vad54@pitt.edu), Jordan Gilardi[1], Ryan Salisbury[1], Cassandra Happe[1], Jason Newman[1], Lambertus Klei[1], Bernie Devlin[1], Ying Ding[1], David A. Lewis[1], Robert A. Sweet[1], Matthew L. MacDonald[1]
[1]University of Pittsburgh, Pittsburgh, PA

Schizophrenia (Sz) is a polygenic disorder that alters how individuals think and behave, affecting 1% of the population with few treatment options. For example, individuals with Sz have impairments in working memory and the processing of auditory sensory information. Alterations in neurochemistry, cytoarchitecture, and brain activity, have been consistently observed in Sz and are believed to underlie these cortical processing deficits. Consistent with genetic studies implicating synaptic proteins in Sz, our prior work using targeted proteomics found alternations in the synaptic protein network in Sz as well as reduced small spine density in the

primary auditory cortex. Here, we used TMT and offline fractionation to more thoroughly compare the bulk proteome, synaptic proteome, and phosphoproteome in the primary auditory cortex between Sz and neurotypical subjects matched for clinical covariates including age, sex, and postmortem intervals. Dendritic spine density and size were also assessed in this cohort. We found robust changes in synaptic proteome and phosphoproteome that cannot be accounted for by bulk protein levels. Co-expression network analysis primarily linked synaptic protein and phosphopeptide modules to the loss of small spines. In sum, the preliminary analysis of our multi-proteomic dataset emphasizes a dysregulation of synaptic protein localization and protein phosphorylation in the human primary auditory cortex in schizophrenia.

P16.07 | Plasma Proteome of Long-covid Patients Indicates Hypoxia-mediated Vasculo-proliferative Disease With Impact on Brain and Heart Function

Douglas Fraser[1] **(douglas.fraser@lhsc.on.ca)**, Cristiana Iosef[2]
[1]*Western University, London,* [2]*Western University, London, ON*

Aims: Long-COVID occurs after SARS-CoV-2 infection and results in diverse, prolonged symptoms. The present study aims to determine the underlying mechanisms, and to inform prognosis and treatment.
Methods: Plasma proteome from Long-COVID outpatients was analyzed in comparison to acutely ill COVID-19 (mild and severe) inpatients and healthy control subjects. The expression of approximately 3000 protein biomarkers was determined with proximity extension assays and then deconvoluted with multiple bioinformatics tools into both cell types and signaling mechanisms, as well as organ specificity.
Results: Compared to age- and sex-matched acutely ill COVID-19 inpatients and healthy control subjects, Long-COVID outpatients showed natural killer cells with a resting phenotype, as opposed to active, and neutrophils that formed extracellular traps. This resetting of cell phenotypes was reflected in vascular events mediated by both angiopoietin-1 (ANGPT1) and vascular-endothelial growth factor-A (VEGFA). Levels of ANGPT1 and VEGFA were validated by serological methods in different patient cohorts. Silent signaling of transforming growth factor-β1 with elevated EP300 favored not only vascular inflammation, but also tumor necrosis factor-α driven pathways. In addition, a vascular proliferative state associated with hypoxia inducible factor 1 pathway was predicted that progressed from COVID-19 to Long-COVID. The vasculo-proliferative process identified in Long-COVID was associated with significant changes in the organ-specific proteome reflective of neurological and cardiometabolic dysfunction.
Conclusions: Taken together, our study uncovered a vasculo-proliferative process in Long-COVID initiated by prior hypoxia, and identified potential organ-specific prognostic biomarkers and therapeutic targets.

P16.08 | PODIOMICS: Biomarker for Prediction of (non)Healing Chronic Diabetic Foot Ulcers

Keith D. Philibert[1], Xiaomeng Shao[1], Charlie Yang[1], Jacque Ortiz[1], Stephanie Wu[1], **Marc Glucksman**[1]
(marc.glucksman@rosalindfranklin.edu)
[1]*Rosalind Franklun University (RFUMS), North Chicago, IL*

Diabetic foot ulcers (DFUs) are a major and potentially life-threatening complication of diabetes mellitus. Ten million people in the United States (1/3 of diabetics) will have DFUs exhibiting significant impaired healing mechanisms due to cellular and biochemical processes involved in the injury response. The four phases of wound healing are inflammation, migration, proliferation, and remodeling. Acute wounds progress through these stages systematically but in chronic wounds, healing is “frozen” at the inflammation stage. In this study, we recruited 40 patients in three different categories: non-diabetic/non-healing ulcer, diabetic/healing ulcer, and diabetic/non-healing ulcer to examine the biologically relevant pathophysiology involved in wound healing. Blood samples were collected using either sharp debridement or hydro-scalpel and proteomic analyses involved elucidating altered levels of cytokines, and *de novo* protein discovery.
Cytokine analysis demonstrated altered levels in three cytokines: IL-1ß (a mediator of the inflammatory response by COX2 induction, proliferation, differentiation and apoptosis) and IL-6 (a pro-inflammatory, acute phase cytokine) both of which increased. RANTES (a CCL5 chemokine with T cell and eosinophil chemotaxis) decreased in diabetic patients with healing compared to non-healing patients.
Mass spectrometric data analyses revealed a decreased abundance in diabetic healers of MMP-9 (involved in proteolysis of extracellular matrix components), apolipoprotein B (involved in cholesterol metabolism) and SCARA3 (Scavenger Receptor Class A Member 3 that eliminates oxidative molecules). An increase in abundance in diabetic healers was exhibited by cytokeratin-e2 involved in keratinocyte activation and proliferation) and transgelin-2, a cadherin binding protein.
These potential biomarkers can be used to predict the healing capacity of DFUs and the need for varying treatment protocols. Identification of prominent biomarkers that are associated in the wound healing process of DFUs can provide a more accurate prediction of clinical diagnoses, prognoses, and creation of novel treatments in the care of these chronic wounds.

P16.09 | Bioinformatic Workflow for Metaproteomic Analysis of Host-Microbe Dynamics in Clinical Samples

Pratik D. Jagtap[1] (pjagtap@umn.edu), Monica Kruk[1], Surbhi Bihani[2], Subina P. Mehta[1], Aryan Gupta[2], Kevin D. Murray[1], Andrew D. Rajczewski[1], James Johnson[1], Reid Wagner[1], Kristin Boylan[3], Amy Skubitz[1], Theresa Laguna[4], Sanjeeva Srivastava[2], Timothy Griffin[1]
[1]University of Minnesota, Minneapolis, MN, [2]Indian Institute of Technology Bombay, Mumbai, India, Mumbai, Maharashtra, India, [3]University of Minnesota, Maple Grove, MN, [4]Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

Clinical metaproteomics has the potential to offer insights into host-microbiome interactions, although with challenges in the detection of low-abundance microbial proteins. As a solution, we have developed an integrated workflow coupling mass spectrometry-based analysis with customized bioinformatic processing of microbial proteins. We have utilized this workflow in ongoing projects to identify microbial peptide panels for: a) cystic fibrosis (CF) disease progression studies, b) co-infection status during COVID-19 pandemic waves, and c) ovarian cancer biomarker discovery. The bioinformatic workflow uses software tools within the Galaxy platform including: MetaNovo (to generate a reduced protein database), SearchGUI/PeptideShaker, MaxQuant and FragPipe (to generate peptide-spectral matches (PSMs) and quantitation), PepQuery (to verify the quality of PSMs), and Unipept (for taxonomy and functional annotation). Rigorous taxonomic, functional and quantitative analysis of verified microbial peptides generated a microbial peptide panel. In the CF study, broncho-alveolar samples from pediatric CF and disease control patients were characterized for their microbial diversity. Targeted analysis of microbial and host proteins from individual samples led to the detection of a peptide panel that includes known and novel peptide targets that can be used to track host-microbe protein dynamics during CF progression. In another study, nasopharyngeal swabs from SARS-CoV-2-infected patients from a hospital in India were analyzed to identify microbial peptides corresponding to potential co-infecting microorganisms during the pandemic waves in India (July 2020 and May 2021). The study detected several peptides belonging to opportunistic bacterial and fungal pathogens. Quantitative targeted analysis of the expression of these microbial proteins and their potential as co-infecting pathogens will be presented. Lastly, we analyzed TMT-labeled Papanicolaou test protein samples from benign, normal, and ovarian cancer patients for host and microbial proteins that are differentially expressed. We will discuss the results of the bioinformatics analysis; the potential use as biomarkers for cancer detection and the targeted investigation of host-microbe dynamics during this presentation.

P16.10 | Quantitative Multiplexed Comparison of Colon Cancer Cell Line Through Automated TMT Sample Preparation and Improved Tribrid Analysis

Woong Kim[1], Santosh Renuse[1], Steven Reeber[1], Kristan Bahten[1], David Horn[1], David Bergen[1], **Daniel Lopez-Ferrer[1] (daniel.lopezferrer@thermofisher.com)**
[1]Thermo Fisher Scientific, San Jose, CA

Introduction: Cancer cell lines have played a key role in understanding the biology underlying various cancers and as screening tools to identify drug targets. As technologies continue to improve to study cellular mechanisms from bulk cell analysis to single cell to single molecules, mass spectrometry based proteomics technologies have also continued to increase its sensitivity and speed. As the throughput for analysis is increasing, method for high quality samples preparation have also become essential. Here we described an end to end solution for quantitative MS analysis combining automated sample preparation with a new mass spectrometer.
Methods: Sixteen different colon cancer cell lines were grown in four different batches and harvested in four different dates as replicates. AccelerOme TMTpro™ kit was used together with the AccelerOme system. Each sample was lysed using 50µl of lysis buffer and transferred to a 96 well plate. The 96 well plate was placed in AccelerOme where the samples were reduced/alkylated, digested, TMT labeled, and cleaned up completely hands-free. Analysis was performed in Orbitrap Ascend Tribrid Mass Spectrometer coupled to Vanquish™ Neo UHPLC system. Acquired data was analyzed using Proteome Discoverer™ software.
Results: In quantitative analysis of large numbers of samples, quality of sample preparation, high confidence analysis with high throughput analysis are key aspects to address. AccelerOme system where the reaction chemistry and workflow have been optimized. Digestion efficiency >90% and TMT labeling efficiency >98% were achieved. The system reduced sample preparation time from 3 days of manual hands-on time to seven hour of hands-off time. The samples were analyzed in Orbitrap Ascend where architecture improvement has increased the ion management resulting in approximately 20% increase in total peptides quantified while reducing the analysis time 30%.

P16.11 | Immuno-proteomics: A Novel Workflow for Autoantigen Identification

Jonathan Minden[1] (minden@cmu.edu), Dana Ascherman[2]
[1]Impact Proteomics, Pittsburgh, PA, [2]University of Pittsburgh Medical Center, Pittsburgh, PA

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 >80 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 patient serums with scleroderma. This autoimmune biomarker discovery method can accelerate sample testing for known autoantigens and facilitate rapid discovery of novel autoantigens

P16.12 | Data-Independent Acquisitions Uncover Kidney Proteome Remodeling during Ischemic Reperfusion Kidney Injury using a Fast-scanning ZenoTOF 7600

Jordan Burton[1], Joanna Bons[1], Anne Silva-Barbosa[3], Jacob P. Rose[1], Katherine Pfister[3], Fabia Simona[4], Oliver Bernhardt[4], Eric Goetzman[3], Christie Hunter[2], Sunder Sims-Lucas[3], **Birgit Schilling[1] (bschilling@buckinstitute.org)**
[1]Buck Institute for Research on Aging, Novato, CA, [2]SCIEX, Redwood City, CA, [3]University of Pittsburgh, Pittsburg, PA, [4]Biognosys AG, Schlieren, Switzerland, [7]

Acute kidney injury (AKI) manifests a major health concern particularly for the elderly, and understanding related proteome changes is critical for prevention and development of novel therapeutics to recover kidney function, and to mitigate the susceptibility for recurrent AKI, or development of chronic kidney disease. In this study, mouse kidneys were subjected to ischemia-reperfusion injury while the contralateral kidneys remained uninjured to assess changes in the kidney proteome. A fast-scanning ZenoTOF 7600 mass spectrometer was used for comprehensive data-independent acquisition (DIA) for protein identification and quantification. Short microflow gradients, and the generation of a deep kidney-specific spectral library allowed for comprehensive and high-throughput assays. Upon AKI the kidney proteome was completely remodeled, and over half of the 3,945 quantified protein groups changed significantly. Downregulated proteins in the injured kidney were involved in energy production, including numerous peroxisomal matrix proteins, many related to lipid metabolism, such as ACOX1, CAT, EHHADH, ACOT4, ACOT8, and Scp2. Injured mice exhibited severely declined health. The comprehensive and sensitive kidney-specific DIA assays, highlighted in this work, feature high-throughput analytical capabilities to achieve deep coverage of the kidney proteome, and will serve as useful tool for the development of novel therapeutics aimed at remediating kidney function. Finally, we present proteome alterations comparing injured and control kidney to highlight affected biological pathways, senescence-associated secretory phenotype markers suggesting increased senescence burden upon AKI. In addition, we link the observed molecular changes with phenotypic mice behavior changes and physiological measurements affected after AKI. The developed high-throughput DIA-MS assays open many applications to monitor therapeutic interventions, potential drug dose responses, dynamic changes over time, timing of interventions after injury. This AKI mouse model together with a fast-scanning DIA-MS workflow will provide a relevant tool towards human translational work addressing human kidney injury in the clinic.

P16.13 | A new plasma proteomics approach to identify kidney-derived peptides and proteins

Sumaiya Nazli[1] (snazli@wakehealth.edu), Michael Olivier[1]
[1]Wake Forest University School of Medicine, Winston-Salem, NC

Proteins released from tissues affected by acute or chronic disease are often pursued as early disease biomarkers. However, identifying and quantifying disease tissue-derived proteins in plasma by mass spectrometry is often challenging. We have adapted an approach to use tandem mass tags (TMT) and tissue samples to identify and subsequently quantify tissue proteins in a plasma sample, even if present in low amounts. Proteins isolated from individual plasma samples are labeled with a different TMT tag and pooled with a reference tissue lysate. During MS analysis, the peptide identification is primarily driven by the more abundant proteins in the reference sample. Using this novel mass spectrometry-based approach, we demonstrate the ability to identify kidney-derived proteins in plasma samples. We used baboon kidney lysate as a tissue reference channel to examine baboon plasma sample in ratios 5:1, 5:2, and 10:1, and compared the identified proteins to the MS analysis of plasma samples directly. Analysis of unlabeled plasma identified 233±22 proteins (1527±266 peptides) and 1855 proteins (9402 peptides) in kidney tissue homogenate. Ratios of kidney tissue to plasma aided in identifying an average of 613±11 proteins, 2995±428 peptides (5:1 ratio, n=14), 539±14 proteins, 2338±232 peptides (10:1 ratio, n=14), 441±13 proteins, 2183±95 peptides (5:2 ratio, n=14). When the kidney lysate is labeled with the highest molecular weight TMT label (131), a ratio of 5:1 of kidney tissue to plasma identified an average of 707±44 proteins, 3408±309 peptides (n=14). Of the identified proteins in plasma samples, 470 proteins were also identified and quantified in the original LFQ analysis of kidney homogenate.

This represents a promising new approach that may lead to identification and quantification of potential disease biomarkers not just for kidney diseases but also other disease tissues that may not be easily accessible for direct analysis in patients.

P16.14 | Large-Scale IonStar Proteomics of Human Retinal Mitochondrial Fractions Reveals Key Regulators for Pathogenesis of Age-Related Macular Degeneration

Shichen Shen[1] (shichens@buffalo.edu), Ming Zhang[2], Deborah A. Ferrington[3], Jun Qu[2]
[1]SUNY at Buffalo, Buffalo, NY, [2]University at Buffalo, Buffalo, NY, [3]Doheny Eye Institute, Pasadena, CA

Age-related macular degeneration (AMD),the leading cause of elderly blindness, is a multifactorial disease with age, environment, and genetics as contributing factors. Multiple single nucleotide polymorphisms (SNP) in genes involved in the complement pathway have been associated with AMD, including the SNP in complement factor H (CFH) causing tyrosine-to-histidine substitution (Y402H). While AMD mechanisms have not been resolved, mitochondrial damage and dysfunction in the retinal pigment epithelium (RPE) has been frequently observed in AMD patients, especially those harboring the CFH high-risk allele. To address this knowledge gap, we examined the hypothesis that altered RPE mitochondrial proteome will provide insight into the loss of RPE mitochondrial function 1) during AMD progression from healthy aging; 2) with the AMD high-risk CFH genotype. In brief, RPE mitochondria-enriched organelle fractions were isolated from human patients with different CFH genotype and at different AMD stages (N = 273 in total), and were analyzed by UHR-IonStar featuring ultra-high-resolution MS1 acquisition and narrow-window MS1 feature extraction for quantification. In total, 6,996 unique proteins were quantified under a set of stringent criteria in all 273 samples. Statistical testing and functional analysis revealed prominent discrepancies of proteomic dysregulation in AMD patients w/ and w/o CFH polymorphism. Remarkably, more significant dysregulation of mitochondrial functions was observed in AMD patients w/ CFH polymorphism. Moreover, co-regulation of dysregulated proteins was identified by Fuzzy C-means (FCM) clustering. FCM clusters representing mitochondria matrix/membrane proteins and apoptosis-related proteins were identified with AMD progression-specific patterns. Interestingly, most mitochondria matrix/membrane proteins were down-regulated at later AMD stages. Besides, 14-3-3 protein family members, ERBB4, and RR2BC were up-regulated while AKT kinases were down-regulated at later AMD stages. Overall, these results provide novel insight into the molecular mechanisms underlying mitochondria dysfunction in AMD and how the CFH genotype influences these results, which are valuable in both treatment development.

P16.15 | High Throughput Plasma Proteomics: How to Deplete Robustly Thousands of Plasma Samples?

Arthur H. Viode[1], Benoit H. Fatou[1], Patrick van Zalm[2], Meenakshi H. Jha[2], David H. Stevenson[2], Saima H. Ahmed[1], Kinga H. Smolen[1], Judith H. Steen[1], **Hanno Steen[3] (hanno.steen@childrens.harvard.edu)**
[1]Boston Children's Hospital and Harvard Medical School, Boston, MA, [2]Boston Children's Hospital, Boston, MA, [3]Boston Children's Hospital and Harvard Medical School, Boston

Proteomics has the reputation of being a low throughput and expensive technology. When the world came to a halt at the beginning of our recent worldwide pandemic, the Steen Lab was tasked to increase its throughput 10-fold to support nationwide COVID-19 research efforts. Within 3 months, the Steen lab leveraged commercially available hardware and in-house developed methodology to establish a plasma proteomics pipeline to analyze thousands of plasma samples in a high-throughput fashion. Scaling-up the throughput of our pipeline 10-fold required optimization and streamlining of all aspects of the analysis pipeline with particular emphasis on cost efficiency, simplicity, and robustness.
Embracing the analytical challenges of plasma analysis, we developed a 2-pronged pipeline: one for neat plasma and one for depleted plasma. To this end, we leveraged selective protein precipitation with perchloric acid, a method developed in the 1940s, to cost-efficiently and robustly process thousands of samples. The cost-efficiency is underscored by the per sample costs of less than 10 cents/sample. The robustness was shown when processing >3,000 plasma samples in three batches over a 10-month period: a principal component analysis did not reveal any batch effects.
This biochemical plasma depletion method in combination with state-of-the-art LC/MS hardware (Evosep ONE + timsTOF Pro 2) and data-independent acquisition routines allows us to identify more than 1,500 plasma proteins even at a throughput of 60 samples per day. The improved detection limit is demonstrated by the ability to detect viral proteins when performing unbiased discovery proteomics on depleted plasma.
This platform covers the needs of current pilot studies with hundreds of samples, but also enables the analysis of statistically relevant sample numbers where 10,000 plasma samples per year are processed and subsequently analyzed on a single mass spectrometer.

P16.16 | Cell type specific proteome impairment in an ASD mouse model

Yi-Zhi Wang[1] (yi-zhi.wang@northwestern.edu), Jeff Savas[1]
[1]Northwestern University, Chicago, IL

Combinatorial expression of postsynaptic proteins underlies synapse diversity within and between neuron types. Thus, characterization of neuron-type-specific postsynaptic proteomes is key to obtaining a deeper understanding of discrete synaptic properties and how selective dysfunction manifests in synaptopathies. To overcome the limitations associated with bulk measures of synaptic protein abundance, we developed a biotin proximity protein tagging probe to characterize neuron-type-specific postsynaptic proteomes in vivo. We found Shank3 protein isoforms are differentially expressed by direct and indirect pathway spiny projection neurons (dSPNs and iSPNs). Studies in mice lacking Shank3 gene exons revealed a robust postsynaptic proteome alteration in iSPNs. We report unexpected cell-type specific synaptic protein isoform expression which could play a key causal role in specifying synapse diversity and selective synapse dysfunction in synaptopathies.

P16.17 | Quantitative Exosome Proteomics Approach for Precise Prediction of Tamoxifen and Palbociclib Response in Late-Stage HR+/HER2- Patients

Ziwei Zhang[1] (zzhan60@uic.edu), Xiuyuan Ma[2], Julia Ekiert[3], Kent Hoskins[3], Yu Gao[1]
[1]University of Illinois at Chicago, Chicago, IL, [2]Abbvie, Chicago, IL, [3]UIC, Chicago, IL

Hormone receptor-positive (HR+)/ HER2- breast cancer accounts for 73% of all breast cancer diagnoses. The combination of tamoxifen with cyclin-dependent kinase (CDK) inhibitors could work together to arrest breast cancer cells in G1 phase. This combination therapy has been proposed to treat late-stage HR+/HER2- patients. However, significant portion of the patients involved in the clinical trial does not respond well to CDK combination therapy and thus is at risk of missing precious treatment window. Exosomes are small extracellular vesicles secreted by cells and could provide valuable biomarkers for cancer drug response prediction. We have developed a sensitive and efficient exosome extraction method to obtain exosome protein from minimal patient plasma. Our optimized exosome isolation method could quantify >400 exosome derived proteins from 100 µl patient plasma sample.
Plasma samples from 39 patients (before and after each treatment cycle) were collected from the ongoing clinical trial (NCT02668666) by Pfizer and Big Ten Cancer Research Consortium. Different exosome protein isolation methods were evaluated and optimized. Exosome and plasma proteins were extracted, purified, digested and labeled with TMT 10plex.

A network model was developed to differentiate responder and non-responder to combination therapy from the exosome proteomics data even before the treatment. Our analysis could provide crucial information for predicting response to combination therapy with high precision (87.3%) and sensitivity (92.3%).
This finding could help classify and predict the prognosis of our patients involved in clinical trials and cancer treatments, providing an approach to avoid ineffective treatment. The underlying molecular mechanism was investigated with drug-resistant breast cancer cell lines. Our results have the potential to provide new insights into biomarkers for prognosis prediction and precision medicine.

P16.19 | Elucidating the mechanism of germline neurodevelopmental histone H3.3 mutations

Elizabeth G. Porter[1], Khadija D. Wilson[2], Joanna Lempiainen[1], Peter Klein[2], Elizabeth Bhoj[3], **Benjamin A. Garcia[1] (bagarcia@wustl.edu)**
[1]Washington University School of Medicine, St. Louis, MO, [2]University of Pennsylvania School of Medicine, Philadelphia, PA, [3]Children's Hospital of Philadelphia, Philadelphia, PA

Histone H3.3, the replication independent variant of histone H3, is monogenetically mutated in over 46 individuals with severe neurodevelopmental disorder and intellectual disabilities. Additional comorbidities include seizures, hypotonia, oculomotor abnormalities, cardiac defects, and thin corpus callosum. At present, 37 distinct mutations have been identified along both H3.3 genes, *H3F3A* and *H3F3B.* To understand how mutant H3.3 causes this novel neurological disorder, we began by screening H3.3 mutations in HEK293T cells. Histone PTM analysis by mass spectrometry demonstrated a global increase acetylation on H3, H2AV/Z, and H4 for cells expressing the core of H3.3 mutants (G90R and P121R), but little global changes for tail mutants (R8S and G13R). However, using a targeted method for mutant specific peptides, we observed increases in H3.3K4/14ac on R8S and H3.3K9me2 for G13R. Proteomic analysis showed significant changes for G90R and P121R, with strong overlap in dysregulated proteins levels, including SHMT1, LLGL1, ALKBH5 and DRHS7B. R8S and G13R had minor alterations. From this data we proceeded to examine how G90R and P121R mutations altered the development of *xenopus laevis.* We introduced H3.3WT, G90R, or P121R by microinjection of mRNA at the one cell stage, and collected embryos at stage 17, the neural crest stage. The neural crest is a population of cells that immigrate together and eventually form neurons, cardiac tissue, bones, and facial cartilage. These are all tissues effected in the H3.3 syndrome. Using this system we evaluated gene expression by RNA-seq and found H3.3G90R dysregulated genes involved in Wnt, Notch, and BMP signaling pathways, critical in neural crest migration. Additionally, we observed changes in facial cartilage and motility of stage 47 tadpoles. Here, we will focus on the initial molecular mechanisms of mutations at H3.3G90R and H3.3P121R to understand how these epigenetic mutations affect neurological development.

P17: Quantitative Proteomics and Systems Biology

P17.01 | Utilizing Proteomic-Based Strategies to Investigate Molecular Mechanisms of Fragile X Syndrome

Diana Abbasi[1] (Diana_Abbasi@rush.edu), Elizabeth Berry-Kravis[1], Stephanie M. Cologna[2]
[1]Rush University Medical Center, Chicago, IL, [2]University of Illinois Chicago, Chicago, IL

Fragile X syndrome (FXS) arises from a CGG trinucleotide expansion of fragile X messenger ribonucleoprotein 1 (FMR1) leading to transcriptional silencing of the gene and lack of production of its protein product, fragile X messenger ribonucleoprotein (FMRP). Since FMR1 was cloned in 1991, there has been great progress in understanding the genetics of FXS, FMRP function, and how its loss leads to many aberrant cellular, molecular, and behavioral observations in FXS. We understand that there are several genetic factors with respect to the mutation that influence how an individual with FXS is affected neurocognitively. However, in males with the same apparent mutation in FMR1 (e.g., male patients with a full mutation in FMR1 and no identifiable mosaicism), there is still high variability in intellectual ability and functional skills. This phenotypic variability has not been well defined and suggests that there are molecular factors other than the loss of FMRP influencing the fragile X phenotype. To investigate this, we quantitatively compared the proteome of patient fibroblast against age-matched controls using mass spectrometry. Linear regression analysis of the differential proteins revealed significant positive and inverse correlations between protein abundance and patient's Z-dev IQ measure. Additionally, we detected changes in ribosomal and ribosomal-associated proteins as well as mitochondrial proteins involved in oxidative phosphorylation, the tricarboxylic acid cycle, and sirtuin signing. Furthermore, alterations in proteins such as N-nicotinamide methyltransferase (NNMT), various isoforms of NADH dehydrogenases (NDUFV2, S5, A9, Y1), and solute carrier family 25 member 1 (SLC25A1) may suggest dysregulation in mitochondrial epigenetic crosstalk. Lastly, we carried out analysis of histone post-translational modifications by performing histone enrichment and mass spectrometry, which revealed five altered histone signatures: H3K4me1, H3K23ac, H3K27me1, H3.3K36me1, H3.3K27me1. Overall, this data presents potential biomarkers that can be used as prognostic measures and pathways that aid in our understanding of FXS pathophysiology.

P17.02 | Revealing Systematic Proteomic Remodeling through Library-Free DIA-MS Analysis of a Novel Progressive Prostate Cancer Cell Model

Daniel G. Delafield[1] (delafield@wisc.edu), Hannah N. Miles[1], Teresa G. Liu[1], William A. Ricke[1], Lingjun Li[1]
[1]University of Wisconsin-Madison, Madison, WI

Representing the second-most common form of cancer in men with the highest age-adjusted rate of incidence, prostate cancer (PCa) continues to occupy an area of extreme need in human health-related research. PCa is a progressive disease with each stage presenting distinct physiological characteristics, further hindering rapid advances in therapeutic strategies. Until recently, no suitable animal or cell model mirrored the progressive nature of PCa, forcing incongruous study of unique PCa stages and limiting logical proteomic comparisons. Through the development of the Benign prostate hyperplasia to Prostate Cancer (BCaP) cell line model – work provided by Ricke and colleagues – we are now able to directly investigate proteomic remodeling across six unique PCa pathogenesis and development conditions that span the four clinical stages. With this unique biological model in hand, we sought to characterize distinct physiological aberrations presented as PCa progresses from benign tumorigenic (control) to tumorigenic to metastatic tumorigenic states. Utilizing library-free data-independent acquisition mass spectrometry (DIA-MS) we identified 86,640 unique peptide sequences mapping to 9,877 protein precursors with excellent intra-sample reproducibility ($R^2>0.98$). Adjusting for a 1% protein-level FDR cutoff, we quantified 6,614 proteins with 1,245 shown to be dysregulated against control. Clustering analysis of dysregulated proteins reveal six distinct protein aberration profiles in accordance with cancer progression. Among others, these clusters suggest progressive cellular senescence, disruption of cytoskeletal organization, and reduced antigen presentation in advanced cancer states. Further incorporating next-generation sequencing into our analyses, we identified 485 genes significantly dysregulated in transcription and translation, presenting a pool of known and previously uncorrelated biomarker candidates for PCa detection and stratification. Further revealing 969 dysregulated proteins with normal transcription, these analyses reinforce the need for in-depth proteomic profiling in human-health research. In all, this report provides early reports of protein markers potentially useful for prognostic evaluation and novel therapy in progressive PCa.

P17.03 | ExclusionMS: Making a Smart Mass Spectrometer though dynamic exclusion lists

Patrick Garrett (pgarrett@scripps.edu), Scripps Research, San Diego, CA

With enough processing power, it is feasible to search mass spectra in real-time, i.e., at the same speed that the instrument can acquire them (real-time search (RTS)). RTS significantly increases the speed of data analysis by enabling a search to happen simultaneously with data acquisition. More importantly, it allows smart acquisition methods (SAM) to modify the mass spectrometer's behavior in real-time based on the real-time search results. To support the creation of novel SAMs, we have created ExclusionMS, a platform to enable dynamic exclusion lists (DEL) to exclude previously identified ions and schedule new unidentified ions.
ExclusionMS is three separate applications: 1) A docker container that runs a FastAPI server and handles all communication, 2) A Streamlit application that serves as the GUI interface, and 3) a python package enables easy integration and future development. These features make ExclusionMS easy to set up on any PaSER-enabled system and extend to other suitable RTS platforms.
We have extensively tested ExclusionMS on both a simulated and real mass spectrometer. We have observed that it can easily keep up with the speed of acquisition on the Timstof series instruments, handling ~500 additions/queries per second. The next experiment we will perform is to run a series of replicates with the dynamic exclusion list enabled. From this, we hope to see that subsequent experiments identify unique peptides/proteins, ultimately increasing the depth of analysis. Additionally, we would like to use ExclusionMS to selectively target proteins within a digest by excluding peptides from all other proteins.
SAMs have the potential to revolutionize mass spectrometry by making mass spectrometers intelligent. So it is crucial to understand and minimize the obstacles to their development. ExclusionMS offers a simple yet powerful solution to enable exclusion-based SAMs for any RTS platform.

P17.04 | Robust, high throughput and deep plasma proteomics with a multi nanoparticle-based workflow coupled with Orbitrap Exploris 480 mass spectrometry

Amirmansoor Hakimi[1] (amirmansoor.hakimi@thermofisher.com), Lucy Williamson[2], Michelle Dubuke[3], Satendra Prasad[4], Ryan Hill[4], Khatereh Motamedchaboki[5], Tabiwang Arrey[6], Daniel Lopez-Ferrer[1]
[1]Thermo Fisher Scientific, San Jose, CA, [2]Seer Inc., Redwood City, CA, [3]Seer, San Jose, CA, [4]Seer Inc., San Jose, CA, [5]Seer Bio, Redwood City, CA, [6]Thermo Fisher Scientific, Bremen, Germany

Introduction: Data Independent Acquisition (DIA) is a powerful method for cataloging thousands of proteins in complex biological samples like human plasma. However, for any large-scale proteomics studies, both LC and MS systems need to be robust, without compromising on peptide and protein coverage and analysis precision. Here we present a high-throughput single injection, label-free plasma proteomics workflow with Orbitrap Exploris 480 mass spectrometer coupled to a High-Field Asymmetric Waveform with Ion Mobility Spectrometry (FAIMS) as a robust analytical setup for in-depth analysis of plasma samples processed with Seer's Proteograph™ Product Suite utilizing a multi-nanoparticle-based approach for an unbiased and deep plasma proteomics analysis at scale.
Method: Neat plasma and Seer's enriched plasma proteins digested on Proteograph Product Suite, were analyzed using a 75cm, EasySpray PepMap-Neo column on Vanquish Neo UHPLC at 250nl/min, coupled to an Orbitrap Exploris 480 mass spectrometer with and without FAIMS Pro interface. Performance of the whole workflow was evaluated with a single injection using a 90 min effective gradient that allowed for 160 min total analysis time sample to sample. The data were analyzed with DIA-NN in Seer's Proteograph™ Analysis Suite and Spectronaut-16 (Biognosys), using a spectral library-based approach, providing enhanced peptide and protein coverage with a 1% FDR rate.
Results: Proteograph with Label-free DIA method resulted in 3423 protein groups. The addition of FAIMS improved the identifications by 10% and utilizing Spectronaut improved the identification further to 4236 protein groups. The single injection of 2ug pooled 5NPs plasma digest was used for improvement in throughput without compromise on protein and peptide coverage with nanoflow sensitivity. This optimized workflow provides an easy-to-use, robust workflow for high throughput, deep plasma proteomics analysis at scale. In addition, we foresee opportunities to further increase the throughput by further optimizing loading, equilibration and washing steps by up to 30% improvement.

P17.05 | Deletion of Translational Slippery Sites in CFTR Remodels The Proteostasis Network

Madeline F. Herwig[1] (mfherwig@gmail.com), Minsoo Kim[2], Eli F. McDonald[2], Lars Plate[2], Jonathan Schlebach[3]
[1]Augustana College, Vanderbilt University, Geneva, IL, [2]Vanderbilt University, Nashville, TN, [3]Indiana University, Bloomington, IN

Slippery sites are specific nucleotide sequences in mRNA that affect the rate of ribosomal translation. These sites cause the ribosome to "slip", resulting in ribosomal stalling and a frameshift in the ribosomal reading frame. These sites have recently been

discovered in the transcript of Cystic Fibrosis Transmembrane Conductance Regulator (CFTR). Mutations within this protein result in Cystic Fibrosis (CF), a lethal, monogenetic disease. CFTR is a membrane anion channel protein expressed in epithelial cells and is responsible for the transport of chloride ions across the plasma membrane. The most common mutation of CFTR is the deletion of phenylalanine 508 (F508del) and 90% of CF patients have at least one allele containing this mutation. Notably, slowing translation speed increases F508del expression. However, the relevance of these identified slippery sites in CFTR and their effect on its translation is largely unknown. Within this study, affinity purification and Liquid Chromatography Mass Spectrometry/Tandem Mass Spectrometry (LC-MS/MS) were used to quantify differences in protein interactors between Wild Type CFTR (WT), Wild Type CFTR with its slippery sites deleted (WT ΔSS), CFTR F508del (ΔF508), and CFTR F508del with its slippery sites deleted (ΔF508 ΔSS). The focus of this study was to determine whether the deletion of the CFTR slippery sites affects CFTR expression and if it changes interactors within the CFTR Proteostasis Network. We found that the deletion of slippery sites in the CFTR mRNA transcript results in distinct CFTR Proteostasis remodeling, most notably within the different intersections of ribosomal proteins.

P17.06 | Micro-scaffold Assisted Spatial Proteomics Version II (MASP-II): Substantially Improved Spatial Resolution and Throughput for Whole Tissue Mapping

Shihan Huo[1] (shihanhu@buffalo.edu), Min Ma[1], Ming Zhang[1], Shuo Qian[1], Jie Pu[2], Xiaoyu Zhu[1], Sailee Rasam[1], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY, [2]SUNY at Buffalo, Buffalo, NY

Reliable spatial mapping and characterization of region-to-region variability of proteins (*e.g.,* therapeutic drugs/targets and biomarkers) across the whole tissue provides critical information to inform engineering and therapeutic efforts. Previously, we have developed a *Micro-scaffold Assisted Spatial Proteomics(MASP)* pipeline, based on spatially-resolved tissue compartmentalization using a 3D-printed micro-scaffold(400-μm side), coupled to reproducible sample preparation(μ-SEPOD) and sensitive/robust LC-MS analysis(IonStar), capable of mapping >5,000 proteins in a mouse brain, with well-validated mapping quality. Here we developed a new generation of MASP pipeline(MASP-II) with substantially-improved spatial resolution(4-folds) and analytical throughput(4-16 folds), and demonstrated its application in mapping of the whole-brain distribution of mAbs and cerebral proteins post a direct CNS administration.
A novel micro-scaffold with hexagon-shaped cells, which is highly stable and carries excellent 3D-printing resolutions, was devised and fabricated for robust micro-compartmentalization of various tissues(*e.g.,* brain, liver, spleen, heart, etc.) at 100-μm-resolution. To substantially improve analytical throughput, which represents the major bottleneck for all spatial proteomics techniques, we adopted i)automatic sample preparation, by including automatic SP3 in the μSEPOD procedure and ii)multiplex TMTpro protein quantification.
As for the application, two non-targeted mAbs(IgG1&2) were dosed through intracerebroventricular(ICV)-dosing in mice. In total, we quantified >6,000 proteins, with ~3,000 proteins mapped in >80% of spatial locations. The mapping accuracy was confirmed by the high correlations between MASP-measured distributions and literature-reported distributions of certain markers. This novel study revealed high penetration of ICV-dosed mAbs in the lower cortex region and the two mAbs showed highly correlated distributions(r=0.93). Moreover, the distributions of many brain proteins were positively- or negatively- correlated with the two mAbs. These discoveries provide novel insights into mAb delivery into the brain.
To summarize, the MASP-II with higher resolution and throughput, facilitates a finer understanding of the spatially-organized biological/pharmaceutical processes in tissues. The method can be readily conjugated with various sample preparation and analytical platforms.

P17.07 | Peptide quantification of in vitro digested food protein in the presence of physiological surfactants

Katarzyna Klosowska[1] (katarzyna.klosowska@pg.edu.pl), Julia Bräcker[2], Jens Brockmeyer[2], Adam Macierzanka[1]
[1]Gdansk University of Technology, Gdansk, Pomorskie, Poland, [2]University of Stuttgart, Institute of Biochemistry and Technical Biochemistry, Stuttgart, Germany

Physiological surfactants may influence the rate of human gastrointestinal hydrolytic reactions. In the human intestine, the primary source of surfactants is bile. Its major surface active components are phospholipids (PL) and bile salts (BS). In proteolysis, BS can increase the rate of hydrolysis by pancreatic proteases. However, the function of PL is not well understood. This research aims to investigate the effect of the BS/PL ratio on the proteolysis of model food-grade protein material, Whey Protein Isolate (WPI). The in vitro proteolysis of WPI was performed regarding the conditions established by the INFOGEST protocol. The progress of WPI digestion was monitored by SDS-PAGE analysis. The time-point samples collected during the experiments were examined with HPLC−MS/MS using a bottom-up proteomic approach and label-free quantification.
The intestinal digestion performed in the presence of bile salts was rapid. Under the physiological concentration of biosurfactants (BS/PL), the protein is fully digested after 60 min. However, higher amounts of phospholipids in the digestion mixture led to delayed hydrolysis of the protein. As a result, the protein containing only PL remained undigested over 120 min of the experiment. The results provide insights into peptide profiles during digestion in various ratios of biliary surfactants. Depending on conditions, the number of peptides and their length varies.

The knowledge of food digestion peptidome provides essential information regarding food allergies. The BS/PL ratio can vary in the human small intestine for various physiological and pathological reasons. Therefore, PL may play a vital role in intestinal digestion.

P17.08 | Moving Proteomics towards the clinic: Realtime dda-PASEF and dia-PASEF Plasma Proteomics analysis with the timsTOF platform and PaSER

Jonathan R. Krieger[1] (Jonathan.krieger@bruker.com), Sander Willems[2], Tomasz Siwarga[3], Ignacio Jauregui[3], Edoardo Coronado[3], Sven Brehmer[3], Christoper Adams[6], Roman Fischer[7], Eduardo Chicano[5], Antonio Rivero-Juarez[5], Mario Frias Casas[5], Dennis Trede[3], Sebastian Wehner[3], Robin Park[4]
[1]Bruker Canada, Toronto, ON, [2]Bruker, Bremen, Germany, [3]Bruker Daltonics GmbH & Co. KG, Bremen, Germany, [4]Bruker, Billerica, MA, [5]IMIBIC Mass Spectrometry and Molecular Imaging Unit (IMSMI). Maimonides Biomedical Research Institute of Cordoba (IMIBIC), Reina Sofia University Hospital, University of Cordoba, Cordoba, Spain, [6]Bruker Scientific LLC, San Jose, CA, [7]Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

Purpose: Plasma proteomics is one of the main gateway's of entry for proteomics into the clinic, however plasma proteomics presents unique challenges. The high dynamic range of tens of proteins in plasma creates challenges in both identification and the quantification of the lower abundant but more clinically relevant proteins. The timsTOF platform of instruments is well situated to handle such challenges due to the added ion mobility dimension which increases identification while reducing chimeric spectra in both dda-PASEF and dia-PASEF modes. Subsequent to analysis, searching the data generated can be a slow and arduous process that can bottleneck results and be a barrier of entry to the clinic. Here we show that the PaSER platform can, in real time, consistently, confidently, and accurately identify and quantify plasma samples acquired in either dda-PASEF or dia-PASEF mode to a great level of depth.
Methods & Results: To demonstrate the plasma proteomic workflow, we analyzed 212 DDA plasma samples with label free quantification enabled. We identified greater than 1200 protein groups and observed a run-to-run correlation average of greater than 0.97 across all quality control runs. In this dataset we were able to quantify a dynamic range of protein intensity spanning 5 orders of magnitude. We also analyzed 400 different clinical plasma samples by dia-PASEF to investigate if candidate biomarkers could be identified between different clinical conditions. While several hundred proteins were identified in these short runs across all the samples, we clearly identified two proteins that show significant difference between disease and normal state. These two candidates will then be validated as potential biomarkers.
Conclusion: Taken together, DDA-PASEF and dia-PASEF coupled to the PaSER platform can provide deep, accurate plasma results in near real time.

P17.09 | Proteomics analysis combined with pulsed-metabolic labeling reveals new targets and mechanisms of host protein degradation mediated by Herpes simplex virus type 1

Katarzyna J. Kulej[1] (kulejk@chop.edu), Eui Tae Kim[1], Namrata Kumar[1], Edwin Halko[1], Simone Sidoli[3], Benjamin A. Garcia[2], Matthew D. Weitzman[1]
[1]Children's Hospital of Philadelphia, Philadelphia, PA, [2]Washington University School of Medicine in St. Louis, St. Louis, MO, [3]Albert Einstein College of Medicine, Bronx, NY

Herpes simplex virus-1 (HSV-1) causes contagious and persistent infections that affect about 60% of the human population. Our previous work demonstrated that HSV-1 infection leads to changes in histone marks of host chromatin and alters the host cell proteome (MCP 2017). We wished to define the viral activities that mediate selective host proteins and histone modifications alterations. Here we have developed a pulsed-metabolic labeling approach based on mass spectrometry and proteomics to determine how HSV-1 lytic infection induces degradation of host targets across the total cellular proteome. Analysis of histone post-translational modifications in infected cells revealed that host chromatin is enriched in histone marks of decondensed heterochromatin after infection i.e., loci decorated by silencing histone marks co-occur with markers of open chromatin (e.g., H3K9me3K14ac). We observed that some cellular proteins associated with heterochromatin are reduced in abundance during infection. To discriminate between protein degradation versus arrested gene expression, we compared protein abundance with labeling rate. Proteins reduced in abundance, but which maintained incorporation of heavy amino acids during infection, were classified as degraded rather than no longer translated. Importantly, this protein group included all known degraded targets of the viral ubiquitin ligase ICP0. Our method also allowed us to discriminate between proteasome- and lysosome-mediated degradation. Interestingly, we did not observe reduced abundance of histone deacetylases (HDACs) despite a significant increase in histone acetylation. Therefore, we performed phosphoproteomics analysis to identify regulation of chromatin via phosphorylation and inactivation of chromatin modifying enzymes. Our analysis detected a significant enrichment of phosphorylation on HDAC1 and HDAC2, which are known to be associated with their inhibition or subcellular localization. Our data demonstrate that combining pulsed-metabolic labeling of cellular proteome together with phospho-proteomics and analysis of histone marks, reveals multiple levels of host protein regulation during HSV-1 infection.

P17.10 | Cell size shapes the eukaryotic proteome

Michael C. Lanz[1] (mikelanz@stanford.edu), Josh Elias[2]
[1]Stanford University, San Mateo, CA, [2]CZ Biohub, San Francisco, CA

Cell size is tightly controlled in healthy tissues, but it is unclear how deviations in cell size affect cell physiology. To address this, we employed a high-accuracy mass spectrometry-based approach to measure how the proteome changes with cell size. Size-dependent protein concentration changes are widespread and predicted by subcellular localization, changes in mRNA concentrations, and protein turnover. As proliferating cells grow larger, concentration changes typically associated with cellular senescence are increasingly pronounced, suggesting that large size may be a cause rather than just a consequence of cell senescence. Size-dependent changes to the proteome, including those associated with senescence, are not observed when an increase in cell size is accompanied by an increase in ploidy. Our findings show how cell size could impact many aspects of cell physiology and provide a rationale for cell size control and polyploidization. Importantly, our re-analysis of published single-cell proteomic datasets suggests that cell size contributes to cell-to-cell proteome heterogeneity.

P17.11 | The S-Trap for Clean and Robust Automated Sample Preparation in Bottom-Up Proteomics

Stefan Loroch[1] (stefan@protifi.com), Eleftherios Panagiotidis[2], John Wilson[3]
[1]ProtiFi LLC, Dortmund, Nordrhein-Westfalen, Germany, [2]Leibniz-Institute für Analytische Wissenschaften - ISAS - e.V., Dortmund, Nordrhein-Westfalen, Germany, [3]ProtiFi LLC, Fairport, NY

Recent technological advancements have led LC-MS-based proteomics towards a new era of middle or even high throughput. This renders sample preparation in terms of manual pipetting of individual samples as uneconomic (extremely labor intense) and prone to human mistakes. Our integrated S-Trap-based sample preparation pipeline featuring a positive pressure device (Resolvex A200, Tecan) and a liquid handler (Microlab STAR, Hamilton) allows to keep up with the throughputs of modern LC-MS instrumentation enabling processing of > 450 samples (96-well format) or even > 1,900 samples (384-well format) per lab worker per week. Comparison of our S-Trap-based pipeline to a FASP-based approach revealed that the S-Trap fully removes all frequently used detergents such Triton-X 100 and Tween-20 as well as PEG contaminations. In contrast, detergents and PEG were well-retained upon FASP and co-eluted with the resulting peptide fraction rendering those samples not suitable for further LC-MS analysis (as detergents drastically reduce column lifetimes down to a few injections). Further, addition of detergents led to a significant reduction of reproducibility upon FASP but did not affect the S-Trap's reproducibility. Notably, the tests also confirmed that the S-Trap grants access to a substantially larger fraction of proteins, predominantly attributed to the ability of capturing smaller proteins (down to 5 kDa) and the compatibility with 10 % SDS which enhanced the solubility of proteins but led to disintegration of FASP-membranes (we determined polyethersulfone to tolerate up to 0.33 % SDS).

In summary, our automated S-Trap pipeline displayed excellent capturing and protein purification capabilities rendering it a perfect choice for robust automated sample preparation. Using the S-Trap omits the requirement for elaborating tailored purification methods for each individual sample or sample set and allows for routine sample preparation on larger scales without the danger for errors originating from contaminated samples.

P17.12 | A Well-controlled Strategy for High-quality Native Interstitial Fluid Procurement from Tissues with Minimal Intracellular Contamination

Min Ma[1] (minma@buffalo.edu), Shihan Huo[1], Shichen Shen[2], Ming Zhang[1], Yang Liu[2], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY, [2]SUNY at Buffalo, Buffalo, NY

Native interstitial fluid(IF) hosts a wealth of potential biomarkers that reflect tissue-specific biological or pathological activities. Quantitative characterization of the IF proteome is highly desirable for many biomedical and pharmaceutical studies but has long been impeded by the intractable technical challenges in the procurement of high-quality IF from tissues. Tissue centrifugation is the most prevalently used method for native IF collection. Nonetheless, poor reproducibility and contamination of intracellular proteins due to cell rupture, often severely compromise the quality of IF. The development of an optimal procedure guided by a reliable quality control(QC) approach is urgently needed. Here, we developed a novel strategy characterizing the dynamic proteomic changes of cell-rupture markers in IF as the function of centrifugation speeds. The IF samples were collected from freshly sampled mouse tissues under 9 levels of gradually-increased g-forces(66-2644g, n=3/level), and 3164 proteins(FDR<1%, ≥2 peptides/protein) were quantified. Proteins that *i)*showed continuously increasing trends as the g-force increases, *ii)*are from the intracellular compartment, and *iii)*were reproducibly detected in multiple independent experiments were selected as cell-rupture

markers. Finally, a panel of six markers was used to quantitatively determine intracellular contamination in the IF: *Ppr1b, Glcne, Uap1, Sias, Mvd1,* and *Sta10*. By monitoring these markers, the optimal g-force was determined at 661g, which provided the highest yields with minimal intracellular contamination. The high quality of IF collected at this condition was validated in separate sets of healthy mice(n=5) and tumor-bearing mice(n=5). In summary, a panel of cell-rupture markers was developed to quantify IF contamination and thereby identify optimal protocols for contamination-free IF preparation with high recovery. The novel QC strategy is valuable in guiding the development of an optimal, fit-for-purpose protocol for high-quality IF procurement from tissues, which provides a solid foundation for reliable and accurate biomedical and pharmaceutical investigations in the IF.

P17.13 | A rapid and easily deployed approach for subcellular proteomics

Frank McCarthy[1] (frank.mccarthy@czbiohub.org), Kibeom Kim[1], Daniel Itzhak[1], Josh Elias[1], Manuel Leonetti[1]
[1]*CZ Biohub, San Francisco, CA*

Protein functions can dynamically vary as they move between different organelles within a cell. Conversely, a protein can determine an organelle's function. Typical whole cell proteome experiments can miss this important dimension. Spatial proteomics involves the separation and enrichment of organelles prior to mass spectrometry-based protein measurement. With it, we can track transient protein movement between subcellular compartments in perturbed cells guided by "ground truth" marker proteins. This approach can help us achieve better insight into molecular responses and subcellular processes that may have little impact on global protein expression.
Subcellular enrichment has historically been performed by density-based differential centrifugation (DC) fractionation. However, membrane-bound organelles' similar densities can prevent their separation between centrifugal fractions. Here, we overcome this shortcoming with an aqueous two-phase separation method using Dextran and Polyethylene Glycol (DEXPEG) along with two ionic salts, potassium phosphate and potassium chloride. The ionic salts' different chemical affinity for the two polymers creates an electrical potential that more effectively partitions organelles. We found DEXPEG enrichment of major organelles was superior to DC, most notably the plasma membrane (70% vs 25%) and endoplasmic reticulum (50% vs 30%). Importantly, DEXPEG reduces the need for extensive fractionation and ultracentrifugation, improving overall experiment speed and throughput.
We applied this system to two different cell lines (HEK293 and GRANTA) to demonstrate its versatility and robustness. We further demonstrate that label-free quantitation (LFQ) can maintain quantitative accuracy and increase proteomic depth relative to SILAC labeling. Finally, we demonstrate that high throughput instrumentation and data acquisition using PASEF-DIA on a Bruker timsTOF SCP reduces data acquisition from 5 days (DDA on Thermo Fusion Lumos) down to 28 hours without any substantial loss in sensitivity.
The DEXPEG procedure is an improved addition to the spatial proteomic toolbox. We recommend its use particularly for perturbations involving the plasma membrane.

P17.14 | Quantitative Proteomic Analysis of Sputum from 'Responder' and 'Non-responder' Individuals Exposed to Woodsmoke Pollutants

Allie Mills[1] (allie.mills@unc.edu), Angie Mordant[1], Laura Herring[1], Neil Alexis[1], Heather Wells[1], David Peden[1], Julia Rager[1], Ilona Jaspers[1]
[1]*University of North Carolina - Chapel Hill, Chapel Hill*

Woodsmoke (WS) exposure happens in a variety of settings including indoor spaces that use wood cook stoves, ambient exposures from residential wood burning, or wildfires. However, WS impact on lung health is not clear. In certain individuals, acute exposure to WS results in robust airway inflammatory response ("responders"), while in others, minimal inflammatory response occurs ("non-responders"). Responders are defined as having a greater than 10 percentage point increase in sputum neutrophils and elevated pro-inflammatory cytokines 4 and 24hr following WS exposure. The SMOKESCREEN study is a controlled human exposure study in which sputum samples are collected from adult participants pre- and 24hr post-exposure to 500ug/m^3 of WS particles. In an effort to uncover proteins predictive of responder status, we analyzed data from 54 sputum samples, corresponding to 27 individuals pre- and post- WS exposure to identify proteomic differences between responders and non-responders. Participants included male and female responders and non-responders. Sputum samples were prepared using S-trap (Protifi) and analyzed by LC-MS/MS Data Independent Acquisition (DIA) on an Ultimate3000-Exploris480 (Thermo). Data were analyzed using Spectronaut (Biognosys) and ~20,000 peptides corresponding to >2,200 human proteins were identified. Among these proteins, hundreds significantly changed upon WS exposure (pairwise t-tests, FDR-corrected p-value < 0.05, absolute Log2 fold change of 0.6). Despite high inter-individual variability of these samples, we observed systematic differences in the proteomes of responders and non-responders; as well as sex-dependent differences. Interestingly, analyses showed few differentially expressed proteins at baseline, but many proteins were differentially expressed between groups after WS exposure. Additionally, we performed metaproteomic analysis to examine potential differences in the sputum microbiota upon WS exposure, and the relationship between responder status and microbiota composition. These data will help clarify the impact of WS exposure on individuals, as well as help pinpoint biomarkers associated with the responder and non-responder phenotypes.

P17.15 | Micro-proteomics for spatial embryology

Leena Pade[1] (lpade@umd.edu), Peter Nemes[1]
[1]University of Maryland, College Park, College Park, MD

Technologies enabling the identification of diverse types of proteins can improve our existing knowledge of mechanisms underlying tissue patterning and normal development. We studied two neighboring regions in an embryo that later form neural and mesodermal tissues to understand proteomic differences involved in establishing spatially distinct tissue lineages. The limited size of the *Xenopus* embryonic tissues and presence of abundant yolk proteins, however, challenge the detectable proteome coverage. To improve our detection sensitivity, we developed a microscale yolk depletion device capable of isolating yolk platelets from low volumes of tissue lysates (≤ 10 µL), while minimizing protein losses during processing and precluding the need for pooling large tissue amounts. In parallel, we developed a liquid chromatography (LC) mass spectrometry approach to measure the proteome in a discovery setting. We maximized protein identifications by employing a new-generation micro-pillar array column (µPAC), tailored in performance for the separation rate, solvent gradient, and column temperature. Next, we combined our microscale yolk depletion device with high-pH fractionation, followed by low-pH µPAC LC to process and analyze the embryonic tissues. This integrated approach enabled ~8,000 protein identifications from a few micrograms of *Xenopus* tissues, providing 4-fold deeper coverage compared to our previous studies. Comparison of the proteomic profiles of the two types of tissues showed distinct enrichment in energy mechanism. For example, we found the enzymes involved in glycolysis and pyruvate metabolism upregulated in mesodermal-fated tissue, whereas citric acid cycle and fatty acid oxidation enzymes were enriched in the neural-fated tissue. These encouraging results have led us to develop orthogonal functional assays to assess the role of energy cycles in tissue induction and differentiation. Overall, the techniques developed for this study can be tailored to analyze other biological tissues while the in-depth proteomics data can enable discovery of important markers for normal embryonic development.

P17.16 | Quantitative deconvolution proteomics enables accurate protein quantification in patient-derived xenografts

Shuo Qian[1] (sqian@buffalo.edu), Shichen Shen[2], Min Ma[1], Shihan Huo[1], Jun Qu[1]
[1]University at Buffalo, Buffalo, NY, [2]SUNY at Buffalo, Buffalo, NY

Comparing to the injection models, Patient-derived xenograft models(PDXs) more faithfully recapitulate tumor-stroma interactions in cancer but conventional immunoassays and LC-MS is largely inadequate for highly-specific investigation of tumor and stromal proteins, which are species-specific in PDXs. Here, we describe a quantitative deconvolution proteomics approach embedded in IonStar, quantitative deconvolution proteomics IonStar(QDP-IonStar) that can unambiguously and accurately quantify the tumor(human-derived) and stromal(mouse-derived) proteins in PDX samples, enabling unbiased investigation of tumor and stromal proteomes with excellent quantitative reproducibility. One shortcoming of the deconvoluted quantitative proteomics is suboptimal accuracy and precision. Specifically, an informatic tool for specific and accurate multi-species proteomics data processing is lacking, and the variable stroma and cancer cell contents across individual PDX tumors causes remarkable quantification bias. To address these challenges, we developed the QDP data processing module, which has two major functions: i) a protein/peptide ID rematching algorithm to achieve accurate assignment of species information for identified proteins/peptides; ii) a species-specific normalization method to markedly reduce bias caused by the variability in tumor/stroma contents among different PDX tumors. This module not only works with IonStar, but also is compatible with the peptide outputs from different proteomics software packages such as Proteome Discoverer, MaxQuant, and FragPipe. The performance of this module was validated by two benchmark experiments with concocted proteomics. It was observed that both the quantitative accuracy and precision were prominently improved after QDP processing. QDP was then applied to the investigation of a pancreatic cancer PDX model treated with nab-paclitaxel (n=24 animals). Comparing with the conventional data processing procedures, QDP achieved more accurate detection of biological processes associated with well-known drug effects, and obtained novel biological insights. In summary, QDP is a promising tool to accurately deconvolute the species-specific proteome changes, which could greatly facilitate the study of cancer-stroma interaction in studies involving PDX models.

P17.17 | Organism-wide Secretome Mapping Uncovers Pathways of Tissue Crosstalk in Exercise

Nicholas M. Riley[1] (riley.nicholasm@gmail.com), Wei Wei[1], Xuchao Lyu[1], Carolyn Bertozzi[1], Jonathan Long[1]
[1]Stanford University, Stanford, CA

Introduction: Exercise is a powerful physiologic stimulus that provides benefits to multiple organ systems and confers protection against disease. These effects are in part mediated by blood-borne factors that mediate tissue crosstalk and function as molecular effectors of physical activity. To globally understand how physical activity reshapes cellular secretomes, here we use a proximity

biotinylation approach to profile cell type-specific secretomes following treadmill running in mice. This organism-wide, 21-cell type, 10-tissue secretome atlas reveals complex, bidirectional, and cell type-specific regulation of secreted proteins following exercise training.
Methods: An engineered biotinylation enzyme TurboID was delivered into the secretory pathway of cells via adeno-associated virus (AAV) transduction. Cell type-specific labeling is achieved genetically because the expression of the TurboID is restricted to those cells expressing cre recombinase. Biotinylated and secreted plasma proteins can then be purified directly from blood plasma using streptavidin beads and analyzed by LC-MS/MS using data-independent acquisition with DIA-only chromatogram libraries generated through gas-phase fractionation.
Results: In total across all samples (N = 3 mice/condition x 2 conditions x 21 genotypes), we detect 1,272 unique cell type-protein pairs with ? 2 peptides detected in all 3 replicates of at least one condition. Exercise significantly altered 256 cell type-protein pairs (20.1% of the entire dataset, adjusted P-value < 0.05). We identify a gradient of secretome responses across cell types, with secretomes from Pdgfra+ being one of the most exercise-responsive in the entire dataset. Peptide-level correlation analysis uncovers exercise regulation of cell type-specific secreted proteoforms. Finally, we show that exercise-inducible, liver-derived CES2 proteins modulate systemic energy metabolism and suppress obesity in high fat diet-fed mouse models.
Conclusions: Together, our studies map exercise-regulated cell types and secreted proteins and illuminate the dynamic remodeling of cell and tissue crosstalk by physical activity.

P17.18 | High Resolution DIA: A Workflow for Highly Accurate Relative Label-Free Quantification of Microbial Proteins in Complex Cell Lysates

Julia Kraegenbring[1], Tabiwang Arrey[1], Jeff Op de Beeck[3], Natalie Van Landuyt[4], **Aaron M. Robitaille[2] (aaron.robitaille@thermofisher.com)**, Alexander M. Harder[1]
[1]Thermo Fisher Scientific, Bremen, Germany, [2]Thermo Fisher Scientific, Santa Clara, CA, [3]Thermo Fisher Scientific, Gent, Oost-Vlaanderen, Belgium, [4]Thermo Fisher Scientific, Gent, Belgium

Introduction: Relative quantification of proteins in complex samples raises a demand for high sensitivity and reproducibility throughout large sample sets to gain meaningful insights on biological processes. Data-independent analysis has emerged as a powerful technique enabling quantification of thousands of proteins because it avoids the intensity bias and missing value problem that typically limit data-dependent methods. DIA in principle interrogates all peptides that are present in a sample and therefore is especially suitable for high-throughput and large-cohort studies.
Methods: Different microbial proteomes were spiked into a human proteome background at different ratios, yielding two proteome and three proteome mixtures with varying total protein amounts. Samples were separated on a 50 cm µPAC™ HPLC columns in direct injection setup on a Vanquish™ Neo system under nano-flow conditions. DIA experiments were run on an Orbitrap Exploris™ 240 mass spectrometer. Data was analyzed by Spectronaut™ 16 using a library-free approach.
Results: Using micropillar array-based column technology under nano-flow conditions for separation of peptides gives optimal peak shapes and intensities reproducibly over a long-term acquired data set with minimal performance loss. In connection with the high resolution DIA methodology, this enables for wide proteome coverage in two- and three-proteome mixtures as well as quantification accuracy below 10 % at high sample throughput. The uncomplicated and easily implemented library-free data analysis yields similar performance as low-key library-based approaches.
Conclusions: Micropillar array-based separation technology and high-resolution data-independent analysis enable for a wide proteome coverage at high throughput, while maintaining excellent quantification accuracy of relative protein ratios in complex cell lysates.

P17.19 | A modified Orbitrap Tribrid MS enables simultaneous manipulation of three ion populations to improve sample throughput, coverage, and sensitivity

Jingjing Huang[1], David Bergen[1], Yuchen He[2], **Amanda E. Lee[1] (amanda.lee3@thermofisher.com)**, Aaron M. Robitaille[3], Joshua Coon[2], Vlad Zabrouskov[1]
[1]Thermo Fisher Scientific, San Jose, CA, [2]University of Wisconsin-Madison, Madison, WI, [3]Thermo Fisher Scientific, Santa Clara, CA

To better understand the function of proteins in complex biological systems, it is necessary to measure changes in protein abundance across temporal and biological conditions. Unlike affinity-based approaches that are dependent on the specificity of consumable reagents, mass spectrometry-based proteomics can measure the proteoform diversity of sequence variations, splice isoforms, and post-translational modifications (PTMs) with high sensitivity and selectivity. Here, we evaluated a modified Orbitrap Tribrid MS that enables simultaneous manipulation of three ion populations to improve sample throughput, coverage, and sensitivity. Specifically, we analyzed new hardware including dual ion routing multipoles (IRMs) that increase proteome coverage by detecting more precursors in less time, and a modified ion funnel designed to capture ions under gentle conditions to improve the transfer of labile compounds including phosphopeptides. To analyze the potential benefits of a dual IRM design, we analyzed 1µg of human cell digest using 30min gradients and wide-window acquisition strategy. Here, we found that combined with a faster scanning

Orbitrap acquisition of 45 Hz, the numbers of unique peptides could be increased by 34%, or similar coverage could be obtained 33% faster, when compared to previous Orbitrap Tribrid MS instrumentation. Secondly, we determined if softer ionization and transmission could improve the detection of PTMs including phosphorylation. We analyzed 0.5 µg of phophopeptides enriched from a human cell line using a 90 min gradient and observed a 18% improve in the numbers of unique phosphopeptides identified, and a corresponding 25% improve in the modification site localization. The improvement in modification site localization could be explained by an improvement in phosphopeptide Xcorr resulting from higher spectrum quality. In summary, we observed that the modified Orbitrap Tribrid MS, recently introduced as the Thermo Scientific™ Orbitrap™ Ascend Tribrid MS, enables improved proteome coverage in less time than previous instrument.

P17.20 | A new StageTips method based on an innovative sorbent for fast and efficient peptide fractionation in proteomic studies

Kaynoush Naraghi[1] (kaynoush.naraghi@affinisep.com), Mana Shafei[2], Michel Arotçarena[3], Florine Hallez[3], Cerina Chhuon[4], Chiara Guerrera[4], Sami Bayoudh
[1]Affinisep, Le Houlme, France, [2]AFFINISEP USA LLC, Miami, FL, [3]AFFINISEP, Le Houlme, Seine-Maritime, France, [4]INSERM – US24 SFR Necker Proteome, Paris, France

Peptide fractionation for full proteome characterization is very challenging, especially in the case of complex samples. The objective of this study was therefore to develop a simplified procedure for the efficient and fast fractionation of peptides at basic pH, to contribute to the simplification of peptide separation and analysis. A new reversed-phase sorbent, based on small sorbent particles tightly embedded in a monolithic membrane packed in SPE StageTips, was used for the fractionation of peptides resulting from the enzymatic proteolysis of HEK293 cell lysate, and the results were compared to a reference commercial fractionation kit.
Eight fractions were performed on both the commercial column and the SPE StageTips, with an acetonitrile gradient, and each fraction was then evaporated to dryness before being resuspended in an appropriate solvent for nanoLC-MS/MS analysis.
If the total number of proteins identified and the percentage of peptides eluting in only one fraction (50%) were similar for both sorbents, with a good distribution of peptides over the eight fractions, it appeared that the fractionation on the new sorbent presented several advantages compared to the reference kit since it can be stored dry at room temperature while the commercial columns have to be stored at 4°C in a storage buffer. Moreover, due to the SPE StageTips format, the time required for the evaporation of each fraction is almost halved compared to the commercial columns.
Thus, the new sorbent appears as a promising solution for the fractionation of complex samples or the generation of spectral libraries, since it leads to an increase of more than 25% in the number of proteins identified, compared to unfractionated samples. Finally, this new sorbent offers flexibility of format and capacity, since it is also available as spin columns, for high amounts of peptides, or as 96 wellplates, for high throughput experiments.

P17.21 | Deep Proteome Coverage in High Throughput diaPASEF for Complex Arabidopsis Thaliana Samples

Savannah Snyder (Savannah.Snyder@bruker.com), Bruker Scientific, Cuyahoga Falls, OH

The model plant *Arabidopsis thaliana* has geographically and phenotypically diverse accessions, some of which now have well characterised genomes. These accessions enable the exploration of the influence of genetic variation on protein expression. *A. thaliana* has a complex genome with over 25,000 protein coding genes, and many more alternative transcript variations. The *A. thaliana* proteome is very complex, with some paralogous genes producing similar peptides. It is therefore imperative to acquire sufficient depth of peptide coverage to reveal the impact of genetic and transcript variation on protein abundance. The new timsTOF HT allows measurement of samples over a higher dynamic range and affords an opportunity for in-depth coverage of this complex proteome at high throughput. Here, using an optimised diaPASEF method on the timsTOF HT, we measured two different *A. thaliana* accessions, *Col-0* and *Can-0* processed without pre-fractionation or depletion of abundant proteins, such as Rubisco. First, the sample loading volume was optimised to fit the increased ion capacity of timsTOF HT. Next, using py_diAID package , precursor isolation windows in diaPASEF were optimised. Finally, the *A. thaliana* samples were measured by an optimised method with 30 min LC gradient, resulting in an average 122,743 precursors and 10,098 protein groups identified at 1% FDR, respectively. Median coefficient of variation (CV) of the protein groups was between 4.7–5.0%, in addition, 8,441–8,869 protein groups had less than 10% CV. Our results demonstrate a deep coverage of complex proteomes with rapid single-shot run, enable large-scale genetic studies in *A. thaliana*.

Skowronek P et al. 2022. *Mol Cell Proteomics.* 21(9):100279.

P17.22 | Proteomics to Understand Early Neural Tissue Formation During Embryonic Development

Kaitlyn Stepler[1] (kstepler@umd.edu), Jie Li[1], Peter Nemes[1]
[1]University of Maryland-College Park, College Park, MD

Cell fates of early stage blastomeres have been mapped to distinct tissues of the *Xenopus laevis* model organism, which is well-suited to study tissue differentiation during embryonic development due to its short developmental timescale, ease of acquiring embryos, and genetic homology to humans. However, the temporal molecular changes underlying the evolution of different cell lineages remain unknown—particularly in early developmental stages, when cells are too large (i.e., > 50 µm) to be collected via fluorescence-activated cell sorting (FACS). To address this need, we recently developed a novel method to label cell lineages with fluorescent magnetic beads and magnetically sort descendant clones: magnetic cargo (MagCar). We first demonstrated the successful use of MagCar to isolate neural-fated cells from neurula-stage *Xenopus* embryos and validated that MagCar sample collection was compatible with downstream bottom-up proteomics analysis. To evaluate the performance of MagCar, we compared MagCar-sorted cells to those collected by existing methods such as FACS. Though similar numbers of proteins were quantified between MagCar and FACS, FACS resulted in more widespread proteomic disruption than MagCar, with almost one-third of the quantified proteins having significant abundance differences compared to dissected controls. Thus, MagCar lessens the proteomic impact of sample collection and allows more accurate measurements of the true biological state of the cells. Next, we evaluated the feasibility of MagCar for collecting neural-fated cells at key developmental stages and found that MagCar could successfully isolate these cells of varied sizes (10-150 µm) while maintaining high purity and viability. Therefore, we are now applying MagCar with TMT quantitative proteomics to study the temporal evolution of neural-fated tissue throughout embryonic development. Combining MagCar with subsequent proteomics analysis allows us to characterize temporal changes in protein levels in cells with different tissue fates and will thus improve understanding of evolution of tissue-specific cells during embryonic development.

P17.23 | Assessment of pQTL method performance reveals optimal proteogenomic approach to assess the impact of genetic variation on plasma protein levels

Guhan R. Venkataraman[1] (gvenkataraman@seer.bio), Harendra Guturu[2], Ryan W. Benz[2], Khatereh Motamedchaboki[3], Matthijs B. De Geus[4], Sudeshna Das[4], Pia Kivisäkk[4], Steven E. Arnold[4], Serafim Batzoglou[3], Asim Siddiqui[5]
[1]Seer, Inc., Cupertino, CA, [2]Seer, Inc., Redwood City, CA, [3]Seer Bio, Redwood City, CA, [4]Massachusetts General Hospital, Cambridge, MA, [5]Seer, Inc, Redwood City, CA

Quantitative trait loci (QTL) analysis is a useful tool for understanding the genetic etiologies of molecular mechanisms underlying human health and disease. While a vast body of expression QTL (eQTL) studies have provided important insights into the relationship between genetic variation and gene expression, studying the impact of genetic variation on protein levels could provide greater insights into complex human biology. Recent advancements in proteomics – unbiased, deep, and scalable assessment of the plasma proteome using physicochemically distinct nanoparticles coupled with liquid chromatography-mass spectrometry (LCMS)[1] – have enabled high-resolution protein QTL (pQTL) analysis. While proteogenomics approaches have potential for a better understanding of human health through the identification of new disease biomarkers and drug targets or the development of new diagnostic tools to improve disease prediction, a comprehensive evaluation of existing computational tools employed to perform pQTL analysis should be performed to ensure reliable and sensitive results. Here, we used intensities from 5,058 proteins (acquired Proteograph™ workflow and LCMS measurements) and genotyping array data to compute pQTL associations in 184 plasma samples. We assessed the impact of several genetic association tools (BOLT, REGENIE, and PLINK), protein and genotype imputation strategies, and protein quantification methods on computational speed and pQTL sensitivity. We also measured how these parameters affect the number and spread of *cis*- and *trans*-pQTLs throughout the genome and validated these results through 1) cross-referencing of pQTL and eQTL databases; and 2) functional enrichment analyses. Evaluation of the various association techniques yielded a method that is twice as sensitive and fast as the other methods. Overall, this study provides a comprehensive assessment of pQTL mapping tools and provides recommendations for future pQTL studies.

[1]Blume, John E., et al. "Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona." *Nature communications* 11.1 (2020): 1-14.

P17.24 | Proteome Turnover Measurements Identify What and How ApoE Alleles Modulate Proteostasis

Nathan R. Zuniga Pina[1] **(natez@byu.edu)**, Noah Moran[2], Noah Earls[2], Ethan Smith Smith[1], John C. Price[1]
[1]Brigham Young University, Provo, UT, [2]JC Price Lab, Provo, UT

Apolipoprotein E (ApoE) facilitates lipid homeostasis, and years of scientific evidence suggest different ApoE isoforms modulate the risk for disease onset. At the proteome level, we can explain disease as a loss of individual and global protein homeostasis. Our long-term goal is to elucidate how different risk factors, such as certain ApoE alleles, dysregulate proteostasis which eventually manifests in pathological changes. The model of proteostasis posits that the rate of synthesis and degradation plays a role in global proteome maintenance. We use LC-MS proteomics to elucidate how the expression of different ApoE alleles impact proteostasis in the brains of transgenic ApoE mice. In addition to traditional quantitative proteomics, we track time dependent incorporation of heavy isotopes to measure how protein turnover rates change as a function of ApoE allele expression. We have successfully quantified thousands of proteins in cytosolic and membrane fractions of the brain. Furthermore, we have also calculated turnover measurements to provide a holistic model of ApoE related changes to individual and global proteostasis. By incorporating turnover rate measurements, we have identified how synthesis and degradation play a role in driving the changes we observe in concentration measurements. From our data we have quantified how pathways diverge as a function of ApoE allele expression. These initial findings may explain the role of different ApoE alleles as a risk factor for the onset of disease. Our proteomics workflow can be implemented by anyone who aims to understand the mechanisms that drive changes in proteostasis.

P18: Single-Cell Proteomics

P18.01 | Consequences of genomic instability on single cell variability during aging

Ronald Cutler[1] (ronald.cutler@einsteinmed.edu), Simone Sidoli[1]
[1]*Albert Einstein College of Medicine, Bronx, NY*

A long-standing theory posits that the primary cause of aging is the stochastic accumulation of *de novo* somatic mutations. However, conclusive evidence of how these contribute to aging has remained unattainable, until recently, due to limited resolution at the level of individual somatic genomes, i.e. single cells. In this work, we used emerging single cell proteomic and transcriptomic technologies to investigate the mechanism by which somatic mutations might contribute to cellular dysfunction seen during aging. Human lung fibroblast were exposed to gamma radiation (induces double strand breaks) or N-ethyl-N-nitrosourea (ENU) (induces single base substitutions) to induce an accumulation of somatic mutations followed by a recovery period. An increase in somatic mutations was confirmed by single-molecule mutation sequencing. Structural variant mutations were enriched in cells exposed to gamma radiation, while single base substitution mutations were enriched in cells exposed to ENU. Analysis of single cell transcriptomes and proteomes revealed that mutation accumulation correlated with a global increase in cell-to-cell variability and transcriptional or proteomic noise. Using an approach to determine stability of underlying gene/protein regulatory networks, a decrease in in network stability was observed at both the transcriptional and proteomic levels. These results suggest that the effect of stochastic accumulation of somatic mutations increases noise of transcript abundances as well as protein abundances between cells and therefore increases cell-to-cell variability. This might be due to destabilization of underlying gene regulatory networks as demonstrated by the GCL analysis. Future work will be directed towards examining the dysregulation of transcripts and their corresponding proteins.

P18.03 | Automated Container-less Cell Processing Method for Single-cell Proteomics

Cory J. Matsumoto[1] (cmatsu5@uic.edu), Xinhao Shao[1], Marko Bogosavljevic[1], Yu Gao[1]
[1]*University of Illinois at Chicago, Chicago, IL*

Single-cell proteomics stands to reveal aspects of cell heterogeneity and perturbation at a phenotypically direct level, enabling the identification of novel sub-populations and cell development. However, with only picograms of protein in a cell, one of the major challenges is the considerable sample loss from surface binding during sample preparation. These surfaces can include pipette tips and sample containers, which not only result in sample loss, but can also introduce errors, undermining the ability to detect true heterogenous cellular events. To address these issues, we created a containerless platform for processing the cell without any contact surface using acoustic levitation. Our system uses an array of ultrasonic transducers to levitate samples through sound pressure, with added controls for temperature, humidity, visualization, and sample addition. Incubation of protein and peptide standards form HEK293T cells, revealed virtually no loss at 200-10ng amounts when processed in our levitation platform and significant loss when processed in a PCR tube. The amount of sample loss increased with decreasing sample amounts for the tube, with only 77% and 66% recovered for peptides and protein respectively at 10ng.

To compare whole cell processing with levitation vs. PCR tube, SCoPE2 was used at a 50 and single-cell level. At both levels, levitation processed samples displayed greater protein and peptide intensity (about 2x greater in levitation samples) as well as greater identifications when using a signal to noise ratio cut-off. Further analysis comparing peptide intensities revealed that the levitation platform displayed a significant positive correlation in recovering the longer and more hydrophobic peptides compared to tube processing.
Furthermore, we modified this system to add reagents and cells through an automated sample addition arm. When combined with the ability to process up to five samples in parallel and create multiple systems, we can feasibly process 100s of samples per day.

P18.04 | Pharmacokinetic implications of alterations in the proteomes of single human cells

Ben Orsburn (borsbur1@jhmi.edu), The Johns Hopkins Medical School, Baltimore, MD

Single cell proteomics can provide insight in the level of heterogeneity in response to drug treatment and toxicology models. When drug responses are observed in a bulk cell homogenate sampling treated single cells dosed under the same conditions can provide quantitative estimates on the percentage of cells that are responding in that manner. Drug accumulation and activity, however, is not a singular event. Pharmacokinetics is multivariate and is affected by influx, efflux, metabolism and protein turnover, among other variables. To better understand single cell heterogeneity following drug treatment, we have developed methods to simultaneously

monitor drug accumulation, and in some cases, activity simultaneously with proteomic measurements. While not all compounds or metabolites are amenable to the same LCMS parameters necessary to quantify peptides, when they are, we find that these can provide new insights into the differences present in cellular populations.

P18.05 | Integrating single cell transcriptomics and proteomics data in a visual framework

Ben Orsburn (borsbur1@jhmi.edu), The Johns Hopkins Medical School, Baltimore, MD

Single cell transcriptomics is a mature technology with the ability to analyze thousands or tens of thousands of cells. Unfortunately, transcript abundance is largely meaningless information. One of the few cases where transcript abundance has any value is in relatively short exposure to drugs when increasing amounts of protein are needed to respond to drug induced effects. Single cell proteomics, while vastly superior to transcriptomics, is currently limited in relative throughput. As such, the direct combination of single cell transcriptomics data and proteomics data from cells treated with the same compound can be a valuable way to understand a cellular population. This poster will describe the use of an open source program used primarily by astronomers and cartographers to simultaneously evaluate the transcripts and proteomes of drug treated single cells.GlueViz single cell is publicly available. This poster is a submission for the #MassSpecTwitter US HUPO 2023 Principal Investigator poster challenge. The presenter of this poster personally cultured, sorted, digested, labeled and analyzed every sample on this poster and developed this data analysis solution.

P18.06 | Single-Cell Proteomic Analysis of Mouse Aorta

Sarah J. Parker[1] (sarah.parker@cshs.org), Simion Kreimer[1], Jesse Meyer[1], Jennifer Van Eyk[1], Aleksandra Binek[1], Alexandre Hutton[1], Ali Haghani[1]
[1]Cedars-Sinai Medical Center, Los Angeles, CA

Introduction: The Aorta is composed of heterogenous cell types including endothelial, smooth muscle (SMC), fibroblasts, adipocytes, nerve cells, and several different immune cell populations. Single-cell transcriptomic studies have demonstrated important shifts in the proportional composition and the phenotypes within common cellular types that dominate between healthy and disease states. There are no studies interrogating the proteome of the aorta with single-cell resolution. Here, we report our first application of emerging single-cell approaches for proteomic analysis to the study of dissociated mouse aorta.
Methods: Mouse aorta cells were sorted using a CellenONE instrument into 384-well plates and digested with trypsin in a final volume of 400nL. Digested cells were analyzed by DIA-PASEF on a Bruker single-cell timsTOF. DDA-PASEF from 50 cells was used to create a preliminary spectral library. A series of library-free (DIA-NN), DDA-based library, and merged library approaches were merged to maximize proteomic coverage and representation of known cell types present in aortic tissue. Single cell proteomes were dimension reduced with UMAP, and cell types were discovered from the reduced data with Leiden unsupervised clustering based on quantitative protein expression.
Preliminary results: We identified up to 1859 proteins per cell, with a median count of 507, using a custom library built from sample-specific and public cell type proteomes. Leiden clustering of dimension reduced (UMAP) single cell proteomes indicated at least 9 unique clusters of putative cell types were present in the dataset. Intensity profile of cell type specific marker proteins indicate that SMC (myh11, tagln, smtn), endothelial (CD31, nos3), fibroblasts (cygb, prg4), adipocytes (adipoq, rpl7), and at least two different immune populations (coro1a, itga4, lsp1) cells were all now detectable in the dataset. Interestingly, 3 unique smooth muscle cell phenotypes were identified based on proteome expression. Phenotypic plasticity within SMCs is a known biological phenomenon with important functional implications.

P18.07 | Enhancing single-cell proteomics through tailored DIA and micropillar array-based chromatography

Valdemaras Petrosius[1] (valpet@dtu.dk), Pedro A. Fernández[1], Nil Üresin[2], Teeradon Phlairaharn[3], Jeff Op de Beeck[4], Erwin M. Schoof[1], Benjamin Benjamin Furtwängler[2]
[1]Technical University of Denmark, Copenhagen, Copenhagen, Denmark, [2]University of Copenhagen, Copenhagen, Copenhagen, Denmark, [3]The Ohio State University, Columbus, OH, [4]Thermo Fisher Scientific, Gent, Oost-Vlaanderen, Belgium

Single-cell resolution analysis of complex biological tissue is fundamental to capture cell state heterogeneity and distinct cellular signaling patterns that remain obscure with population based techniques. However, the limited amount of material encapsulated in a single cell raises significant technical challenges to proteome profiling. Due to extensive optimization efforts mass spectrometry-based single-cell proteomics (scp-MS) has emerged as a powerful tool facilitating proteome profiling from ultra-low amounts of input, albeit further development is needed to realize its full potential. Accordingly, we carried out comprehensive analysis of orbitrap based data independent acquisition (DIA) for limited material proteomics. Notably, we found a fundamental difference between

optimal DIA methods for high and low-load samples. We further improved our low-input DIA method by relying on high-resolution MS1 quantification more efficiently utilizing available mass analyzer time. With our ultra-low input tailored DIA method, we are able accommodate the long injection times and high resolution, while keeping the scan cycle time low enough to ensure robust quantification. Finally, we further enhance our workflow by combining WISH-DIA with the latest chromatographic and computational advances to and conclude with showcase our developed workflow to profile single-cell proteomes.

P18.08 | Spectral Similarity Comparisons Improve the Accuracy of Match Between Runs in Quantitative Single Cell Proteomics

Alexander Solivais[1] (solivais@wisc.edu), Michael R. Shortreed[1], Samuel H. Payne[2], Lloyd M. Smith[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Brigham Young University, Provo, UT

Single cell proteomics (SCP) provides an in-depth biological portrait by quantifying thousands of proteins and revealing their post-translational modifications. However, SCP pushes the detection limits of mass spectrometry, making accurate quantification difficult. A strategy often used for label free quantitation is 'match-between-runs' (MBR). However, applying MBR to SCP datasets leads to significant inaccuracies. The nature of SCP is such that almost every peptide that is observed in an MS1 spectrum is selected for fragmentation. However, low peptide abundance results in lower quality spectra that are missed in the original search. We developed an algorithm that enhances MBR results for SCP data by locating and scoring MS2 spectra corresponding to MBR events using spectral similarity.
Sample tandem mass-spectrometry data obtained from 0.2ng injections of HeLa protein digest were downloaded from MassIVE (MSV000087524). These data were searched using MetaMorpheus and peptide abundance was quantified via label free quantification. For peptides quantified via MBR, a second search was performed. MS2 spectra with matching retention times and parent masses were searched against a spectral library assembled from top scoring spectra identified in the same dataset. Positive and negative score distributions were generated by calculating the spectral angle between the spectra of homologous and non-homologous peptides independently identified in different runs. For the positive score distribution, 99% of the spectral contrast angles were greater than 0.54. For the negative distribution, 99% of the angles were below 0.42.
We recovered 2636 MBR spectra from six runs. 1255 spectra had a spectral contrast angle consistent with the true positive distribution. 1200 were found to have a spectral contrast angle consistent with the negative distribution, implying that they were falsely identified. Recovering these spectra and calculating their spectral contrast angles relative to a spectral library enables us to discern accurate and inaccurate identifications in MBR quantification.

P18.09 | Developing high-throughput nano LC-MS methods for sample-limited proteomics

Alec Valenta[1] (alec.valenta@thermofisher.com), Christopher Pynn[2], Runsheng Zheng[3], Karl Mechtler[4], Manuel Matzinger[4], Rupert Mayer[4]
[1]Thermo Fisher Scientific, Somerset, NJ, [2]Thermo Fisher Scientific, Germering, Bayern, Germany, [3]Thermo Fisher Scientific, Gemering, Bayern, Germany, [4]Research Institute of Molecular Pathology, Vienna, Wien, Austria

Low-flow liquid chromatography (LC) coupled to electrospray ionization-based mass spectrometry (MS) techniques have the capacity to probe extremely limited sample amounts. The increase in electrospray ionization (ESI) efficiency required is achieved by adopting narrow separation columns and reducing the LC flow rates to the "ultra-nano" (? 100nL/min) range. Several aspects of ultra-low sample quantity analysis must be considered in the creation of robust and reproducible methods for this type of application. First, flow rate must be optimized for both sensitivity and throughput. Second, the LC platform must permit efficient sample analysis without wasting valuable MS acquisition time. Third, MS acquisition parameters must be optimized for the relatively low signal intensity observed from small sample quantities.

Here we describe a standardized LC separation setup together with 5 novel methods for balancing the sensitivity, throughput, and reproducibility required for routine analytics.
Experiments were performed on a Vanquish Neo UHPLC system coupled to an Orbitrap Exploris 480 mass-spectrometer. HeLa protein digest was separated on a 50 µm I.D. column at a flow rate of 100 nL/min. Contrasting data acquisition strategies, *i.e.*, data-dependent acquisition (DDA) and data-independent acquisition (DIA), were also compared for their impact on method performance. Overall, we developed five ultra-low nano-flow LC-MS methods with gradients from 10 to 50 min, providing sample throughput of 24, 36, 40, 60, and 72 samples per 24 hours (up to 85% MS utilization). Using 250 pg diluted HeLa digest and the 10-min LC gradient (72 samples/day), ca. 800 protein groups were identified in DDA (Sequest), >1,800 protein groups in DDA (CHIMERYS), and ca. 2,100 – 2,700 protein groups in DIA (SN16 & DIA-NN18).

P18.10 | Combined Single Neuron Patch-Clamp/Mass Spectrometry (PatchC-MS) Analyses

Jolene K. Diedrich, Matt Albertolle, Nima Dolatabadi, Swagata Ghatak, Maria Talantova, Stuart A. Lipton, Larry Rodriguez, Marisa Roberto, **John R. Yates (jyates@scripps.edu)**, The Scripps Research Institute, San Diego, CA

As interest in single-cell analysis increases, performing single cell MS still remains a challenge. Herein we demonstrate patch-clamp electrophysiological recordings of single human iPSC-derived neurons followed by mass spectrometry analysis of the same cell. Human induced pluripotent stem cell (hiPSC)-derived cerebrocortical neurons are evaluated electrophysiologically by whole-cell recordings with a patch electrode capillary. The neuron is then aspirated into the capillary and expelled into a microtube. A simple digestion protocol is performed, and samples are analyzed by mass spectrometry. The single-cell digests are separated by nanoflow UPLC coupled to a Bruker timsTOF or a Thermo Eclipse, both operating in data dependent modes. Whole-cell recordings were performed on Alzheimer's disease (AD) and isogenic, gene-corrected control (wild-type/WT) hiPSC-derived cerebrocortical neurons. WT neurons of interest were chosen based on their ability to fire action potentials, manifest voltage-gated sodium and potassium currents, and neurotransmitter-mediated postsynaptic currents. We have previously published that AD hiPSC neurons, like those in human AD brain, exhibit enhanced spontaneous action potential frequency, increased voltage gated sodium currents, and increased excitatory postsynaptic current frequency compared to WT neurons (Ghatak et al., eLIFE, 2019). We selected these AD neurons to compare to WT controls for further proteomic analysis. MS data analysis was performed with ProLuCID, Byonic and MSFragger. When injecting half of the contents of a single digested neuron, we were able to identify between 400-2000 proteins per sample. Advances in this methodology are used to perform patch clamping and proteomics analysis on neurons from brain tissue slices. We performed single-cell patch-clamp electrophysiology combined with mass spectrometry proteomic analysis.

P18.11 | Hands-free processing of single cells using the cellenONE and the timsTOF SCP in a label-free proteomics approach

Matthew Willetts[1] (matt.willetts@bruker.com), Chrsitoph Krisp[2], David Hartlmayr[3], Anjali Seth[3], Guilhem Tourniaire[3], Thorsten Ledertheil[4], Jean-Francois Greisch[4], Markus Lubeck[5], Gary Kruppa[6]
[1]Bruker, Billerica, MA, [2]Bruker, Bremen, Bremen, Germany, [3]Cellenion, Lyon, France, [4]Bruker, Bremen, Germany, [5]Bruker Daltonik GmbH, Bremen, Bremen, Germany, [6]Bruker Daltonics, Billerica, MA

Introduction: Single cell proteomics is a rapidly developing field with the potential to make important contributions to the understanding of cellular heterogeneity. Single cell protein extraction, minimal exposure of samples to surfaces and optimal storage and transfer conditions are crucial for loss-less single cell proteome analyses. Recent enhancements in trapped ion mobility spectrometry (TIMS) coupled to fast and sensitive mass spectrometry established in the timsTOF SCP paired with automated single cell sorting and sample preparation realized with the cellenONE® platform allows for sensitive proteome analyses at the single cell level.
Methods: Single HeLa cells were sorted into the label-free proteoCHIP®, directly lysed, and proteins digested at 50°C with high humidity on deck using the cellenONE platform. The label-free proteoCHIP with tryptic peptides was placed into the nanoElute series autosampler, and peptides were injected onto a 25 cm x 75 µm Aurora C18 column (IonOpticks) and eluted into a timsTOF SCP. dia-PASEF® mode was used and analyzed with TIMS-DIA-NN on PaSER® using a library generated from a deeply fractionated human cell line.
Results: Sample pick-up directly from the label-free proteoCHIP was assessed with HeLa lysate digests (Pierce) showing excellent reproducibility at various concentrations. Injections of 1 ng of HeLa peptides on column (1 µL in well) resulted in 23000 peptides from 3600 proteins which were matched by 1 ng HeLa peptides injected from a vial (1 ng/µL). HeLa cells were directly sorted and prepared in the proteoCHIP and identified >n 2000 proteins per single cell. Quantitative comparison of the single cells demonstrated good reproducibility with some variations dependent on cell size, elongation, and cell cycle stage.
Conclusion: A fully hands-free and label-free analysis workflow reproducibly identifies>2000 proteins from single HeLa cells, using the CellenONE platform with the lable-free proteoCHIP and the timsTOF

P19: Structural Proteomics and Native MS

P19.01 | Utilizing Proteome-wide Quantification of Protein Folding Stability to Provide Insight into Proteome Maintenance and Empower Novel Biomarker Discovery

Chad D. Hyer[1] (hyer.chad@gmail.com), Michael J. Zachrison[1], Lavender Lin[1], Leena M. Patil[1], John C. Price[1]
[1]*Brigham Young University, Provo, UT*

Folded protein structure determines function. Hence, correct protein folding is essential for the mechanistic maintenance of proteostasis and overall health. Despite this, modern proteomics often approaches the study of proteostasis using concentration-based metrics. These metrics provide insight into the consequences of changes in proteostasis but fail to demonstrate the biophysical mechanisms leading to dysregulation such as changes in protein efficacy due to misfolding. Protein efficacy and misfolding are functions of protein folding stability (PFS), meaning PFS can be utilized as a metric of protein folding quality to better understand proteostasis. Therefore, MS based assays that determine PFS through partial unfolding and covalent labeling - such as the Iodine Protein Stability Assay (IPSA) - can be utilized to provide proteome-wide insights into changes in proteostasis.PFS assays are widely applicable as they can be performed under physiological conditions, have residue specific resolution, and can be altered to target any chemically modifiable residue. These assays can establish baselines for the “normal” PFS of proteins in their native states, which can then be compared to experimental conditions for outlier identification.

Here, we report on results from several novel applications of PFS assays to probing proteome maintenance. Using PDB x-ray crystallography data, AlphaFold predictions, and IPSA results from human serum, we found significant correlations between PFS and protein surface accessibility, suggesting that PFS assays can provide vital contributions to experimental protein structure determination and prediction. Using IPSA on human samples with previously determined protein turnover rates, we also identified significant correlations between turnover rates and PFS, suggesting a measurable role of PFS in biophysical proteostasis models. We also report on clinical applications of IPSA including testing drug binding effects on PFS and using IPSA as a biomarker discovery tool for identifying amyloidogenic proteoforms associated with mutations linked to hereditary transthyretin amyloidosis.

P19.02 | Radical-directed dissociation for intact protein characterization: Sequence or Structure?

Ryan Julian (ryan.julian@ucr.edu), UC Riverside, Riverside, CA

Radical-directed dissociation (RDD) is a method employing photodissociation to create an active radical at a site-specific location, which can subsequently be employed to cause fragmentation. The method is compatible with intact protein characterization, and RDD can be tuned to favor highly specific fragmentation channels or for the extraction of structural information. Protein structure can be mapped by initiating a radical at a single location, which leads to RDD at sites that are proximal in space following brief radical migration. Alternatively, fragmentation of particular bonds can be favored by creating the initial radical in a position primed to cleave a particular bond, producing large fragments that can be subsequently sequenced. Results from both types of experiments will be presented for model proteins.

P19.03 | Global analysis of protein conformational changes in a pancreatic cancer cell and stellate cell co-culture system

Haiyan Lu[1] (hlu244@wisc.edu), Yuan Liu[1], Hua Zhang[1], Xudong Shi[1], Lingjun Li[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Pancreatic cancer is a deadly disease that involves the interplay between pancreatic cancer cells (PCC) and stromal cells such as pancreatic stellate cells (PSC). The interaction between PCC and associated PSC induces secretion of multiple factors that are critical for cancer development, differentiation, and metastasis. These secreted factors bind to their receptors and induce a cascade of signaling events causing conformational changes in multiple proteins. However, protein conformational changes induced by PCC and PSC interaction s is not well studied and the mechanism is unclear. To globally analyze protein conformational changes induced by PCC and PSC interaction, we designed a co-culture system with pancreas ductal adenocarcinoma cells (PANC-1) and PSC. PANC-1 and PSCs were co-cultured for 2 days and proteins were extracted and subjected to limited proteolysis coupled with mass spectrometry (LiP-MS). Each sample was divided into two aliquots and treated with or without protease-K under native conditions, followed by tryptic digestion under denatured conditions. After desalting and drying down, the peptides were analyzed with a Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer and label free quantification was used to compare samples

treated with or without protease K to determine structure-dependent proteolytic patterns. Our results identified 35 shared proteins exhibiting structural changes between co-culture cells compared to PANC1 or PSC culture alone. The relative quantitation of 35 proteins showed that 7 proteins (including PKM, ARHGDIA, GANAB, CLTC, ACOT13, HK1, and SEC24D) were up-regulated, whereas PRDX5 was down-regulated in co-culture system compared to controls. Moreover, pathway enrichment analysis revealed that 35 shared proteins with altered proteolytic patterns were mainly involved in glycolysis, pyruvate metabolism and TCA cycle. These results suggest PSC-PANC1 co-culture may affect cancer cell energy homeostasis and metabolism pathways and may provide novel insights into cancer cell and stromal cell interactions that potentially provide targets for pancreatic cancer treatment.

P19.04 | Protein structure validation by limited proteolysis

Xinhao Shao[1] **(xshao8@uic.edu)**, Chris Grams[1], Rahul Narayanan[1], Yu Gao[1]
[1]*University of Illinois at Chicago, Chicago, IL*

Introduction: In 2021, breakthroughs in protein structure prediction using deep learning from both Google's AlphaFold and Baker lab have shown promising results for the elucidation of most human proteins. However, for proteins with unknown structures and domains, the accuracy of the prediction has not been fully validated. Here, by using a native(limited) proteolysis strategy with orthogonal approaches and a 3D protein coverage visualization tool, we present a method to validate the result from prediction and possibly provide a new dataset for structure refinement.
Methods: We performed limited digestion on native HEK293 cell lysate at various time points. Using data from PDB, AlphaFold, and Baker lab, we mapped all different peptides from the digestion at different time points to the 3D structures of more than 3,000 proteins by label-free and isotope-labeling method. These 3D protein structure time series are then analyzed in terms of computational metrics such as average cleavage site distance, cleavage site accessibility, etc. In addition, covalent protein painting was performed to assist in the structure validation orthogonally. The time-series 3D protein coverage was visualized by a web application we developed.
Preliminary Results: In total, we have collected native digestion data for more than 3,000 proteins at 12 time points. The average cleavage site distance from the centroid of the protein significantly decreases over time, showing the outer shell of the protein gets digested first. When implemented protein disordered regions from DisProt database, most of the proteins with highly disordered regions did not have consistent time-series distance trend with the rest of proteins, reasoning that those highly disordered regions are flexible and may behave differently under native digestion. A comprehensive statistical analysis showed that most proteins with low-confident predicted structures were clustered together.
Conclusion: We converted proteomics data into protein structural insights and validated deep-learning protein structure predictions using orthogonal approaches.

P20: Top-Down Proteomics and Proteoform Biology

P20.01 | HIV-1 Virion Proteoform Analysis Reveals Conserved and Novel Post-Translation Modifications Modulating Viral Function

Claire E. Boos[1] (ceboos@wsic.edu), James W. Bruce[1], Mark Scalf[1], Rachel M. Miller[2], Nathan M. Sherer[1], Lloyd M. Smith[1]
[1]University of Wisconsin-Madison, Madison, WI, [2]Streck Laboratories, Omaha, NE

Human immunodeficiency virus type 1 (HIV-1) is one of the most effective retroviral killers responsible for 13,000 deaths in United States per year, yet we still lack deep knowledge of the post-translational modifications (PTMs) modulating its function. Proteoform-level knowledge, provided by top-down proteomics, is fundamental to understanding the phenotypes underlying complex biological systems and disease regulation. Single shot LC-MS/MS top-down proteomic analysis of *in vitro* HIV-1 virions has revealed both known and novel highly conserved modifications. Of the conserved chemical modifications, myristoylation and palmitoylation of matrix protein along with phosphorylation of p6 protein was detected. Prenylation of matrix is necessary for membrane binding and transport while phosphorylation of p6 regulates viral budding, signifying key regulatory modifications are being characterized. Further, 193 proteoforms of capsid protein were identified with a high frequency of localized PTMs observed in the C-terminal domain (CTD). The CTD is an important region of HIV-1 capsid for dimerization and oligomerization processes pertinent to viral particle assembly. Among the many proteoforms observed, the presence of dehydroamino acid residues was confirmed in the capsid protein as reported by our lab in a published study (Miller, 2022). This proteoform-level characterization of HIV-1 virions can provide valuable insight into viral pathways and infectivity and inform our understanding of HIV-1 maturation. Observed PTMs can suggest new targets for drug discovery in HIV-1 therapeutics. Future work will employ size-based enrichment, multiple dissociation type fragmentation, and targeted analysis of putative proteoforms. Additionally, multi-protease bottom-up data will be used to inform proteoform identifications in top-down for a deeper coverage of the HIV-1 virion proteome.

P20.02 | Extending the analysis of human proteoforms beyond the 30 kDa barrier using a tribrid Orbitrap mass spectrometer

Luca Fornelli[1] (luca.fornelli@ou.edu), Christopher Mullen[2], Jake T. Kline[1], Joshua D. Hinkle[2], Jingjing Huang[2], David Bergen[2], Kenneth R. Durbin[3], Romain Huguet[2], Graeme McAlister[2], Vlad Zabrouskov[2], John E. Syka[2]
[1]University of Oklahoma, Norman, OK, [2]Thermo Fisher Scientific, San Jose, CA, [3]Proteinaceous, Inc, Evanston, IL

Despite the theoretical advantages offered by the gas-phase interrogation of whole proteoforms, the application of top-down proteomics is usually limited to the analysis of small proteins (<30 kDa), negatively impacting the practical utility of this proteomic approach for both basic and translational research. However, the latest generations of tribrid Orbitrap mass spectrometers offer the possibility of simplifying mass spectra using proton transfer charge reduction (PTCR), enabling the high-throughput analysis of large proteoforms on the liquid chromatography (LC) time scale. We previously described a novel data acquisition strategy, termed targeted PTCR (tPTCR), which allowed to double the number of proteoforms identified from bacterial lysates compared to standard data-dependent acquisition. Here we presents a refined version of the tPTCR method performed on the Orbitrap Ascend to unravel the complexity of the high molecular weight human proteome. Compared to our previous study, the number of LC runs per sample is substantially reduced thanks to the possibility of recording high quality mass spectra of both intact proteoforms (via PTCR) and related product ions (generated via vibrational energy threshold activation methods) averaging only 1-2 time-domain transients (i.e., microscans). Preliminary results show that tPTCR optimized for the Orbitrap Ascend can lead to the identification of human proteoforms >65 kDa, including both soluble and transmembrane proteins. In order to improve the characterization of these large proteoforms, different ion activation techniques are being evaluated on modified Orbitrap tribrid mass spectrometers. In our first comparative study on ~50 kDa denatured human proteoforms, infrared multiphoton dissociation (IRMPD) has achieved higher average sequence coverage than either resonant or beam-type collisional dissociation (CID and HCD, respectively). Coupled to high-resolution molecular weight-based pre-fractionation, the tPTCR method can now produce hundreds of proteoform identifications from human cell lysates, a result unmatched by any traditional data acquisition strategy.

P20.03 | Characterization of Ribosomal Proteoform-ome using Top-Down Proteomics

Lisa Nakayama[1] (lnakayama@wisc.edu), Isabella T. Whitworth[1], John Pavek[1], Claire Boos[1], Austin Carr[1], Mark Scalf[1], Lloyd M. Smith[1]
[1]University of Wisconsin-Madison, Madison, WI

The ribosome is the key molecular machine in the synthesis of proteins. It is a complex of rRNA and ribosomal proteins that translates mRNA into amino acids and catalyzes the formation of peptide bonds. Because of their critical biological importance,

ribosomes are among the most highly conserved protein complexes and are present in all three domains of life. Ribosomes are well regulated by post-translation modifications (PTMs) during cell growth, proliferation, and differentiation processes where the coordination of the rate of protein synthesis is crucial. Mass spectrometry-based proteomics, specifically top-down proteomics, can be used to identify and quantify proteoforms, the unique form of a protein including the amino acid sequence and localized PTMs. Challenges in proteoform identification arise from the difficulty in working with intact proteins and from the large number of highly similar molecules being identified. A solution to these challenges is reducing the complexity of the sample through purification. We are developing a method for purification of ribosomal proteins in E. coli for the characterization of the ribosomal proteoform-ome by top-down proteomics. The number of ribosomal proteoforms identified within 1% FDR by top-down proteomics increased fivefold when we enriched for ribosomal proteins by ultracentrifugation with a sucrose cushion compared to lysate. Further fractionation of the sucrose cushion-enriched ribosomal proteins by PEPPI-MS resulted in an increase from 40 to 115 in the number of ribosomal proteoforms identified by top-down proteomics. We are further experimenting on increasing the number of ribosomal proteoforms by using ion exchange chromatography and 2D-electrophoresis.

P20.04 | Omics-Scale Discovery and Characterization of Multiply-Glycosylated Intact Glycoproteins from Biofluids with Top-Down Mass Spectrometry

Steven Patrie[1] (steven.patrie@northwestern.edu), Jiana A. Duan[1], Erika Cline, Caroline Alvarez, Shannon Berneche[2]
[1]Northwestern University, Evanston, IL, [2]University of Illinois, Urbana, IL

Today, blood glycoprotein microheterogeneity (i.e., *glycoproteoforms* (*gps*)) and the drivers of glyco-enzymology are often not faithfully represented by standard glycan- or glycopeptide-proteomics. We have created a new top-down glycoproteomics pipeline that permit unbiased discovery and characterization of intact protein glycosylation across the blood and CSF proteomes. Our work overcomes limitations of omics-scale processing of glycoproteoforms such as the high dynamic range in protein concentration and physiochemical properties, sample buffer complexity, and lack of bioinformatics resources that discriminate Fuc:Hex:GlcNAc:SA mass isomers allowing unambiguous assignment of 10-100s glycoproteoforms for each protein in an unbiased manner. We describe an evolutionary step for non-targeted glycoproteoform investigations through the online implementation of tandem microflow size-exclusion chromatography with nano-LC-MS/MS (μSEC[2]-nLC-MS/MS). In combination with OFFGEL isoelectric focusing (IEF) the workflow provides enhanced dynamic range for improved characterization of 10-100s of glycoproteoforms for a single biofluid glycoprotein. Glycoproteoform heterogeneity for IEF separated glycoproteins has been efficiently assigned by a Proteoform Network Analysis which uses machine learning to make high-fidelity Fuc:Hex:GlcNAc:SA assignments from the 100,000s of glycoproteoforms often possible for each glycoprotein, with validation and scoring of assignments by a novel graph theory approach. To date, we have found the workflow permits simultaneous detection of more abundant biofluid glycoproteins with 10s-100s of glycoproteoforms each (e.g., L-PGDS, alpha-2-HS-glycoprotein, Antithrombin-III, alpha-acid-glycoprotein, Alpha-1-antitrypsin and beta-haptoglobin) from as little as 10 μL CSF and 5 μL NPP. The use of machine learning to model pI space of separated glycoproteoforms readily discriminated common confounding events such as 2 Fuc vs 1 SA (292 vs 291 Da, respectively) as well as allows robust site-independent determination of expected N-glycans, even in the presence of O-glycosylation or LacNAcation without the need for conventional glycoproteomics methods.

P20.05 | NeuCode Chemical Labeling for MS1-Only Identification of Prostate Tumor Proteoforms

John Pavek[1] (jpavek@wisc.edu), Brian Frey[1], Dustin C. Frost[1], Ting-Jia Gu[1], Lingjun Li[1], Lloyd M. Smith[1]
[1]University of Wisconsin-Madison, Madison, WI

Proteoforms are the main actors in biological processes, and therefore efficient proteoform characterization is important for complete understanding of biological systems. Top-down mass spectrometry is currently the most widely applied method for proteoform characterization. However, its reliance upon isolation and fragmentation limits the number of proteoforms that can be identified in an LCMS experiment. To circumvent this limitation, we seek to identify proteoforms without requiring isolation and fragmentation. However, intact mass alone is often insufficient to produce confident proteoform identifications, which suggests the need for another piece of information to disambiguate identifications. We target the count of a specific amino acid as this additional disambiguating parameter. Previously, NeuCode SILAC has been used to determine proteoform lysine counts for identification, but SILAC is limited to cell culture samples. The current work incorporates NeuCode labels via chemical labeling, thereby enabling proteoform identification via amino acid count and intact mass on any protein sample. The developed labeling reagent targets cysteine residues; thus proteoform cysteine counts are determined; and proteoforms are identified via intact mass and cysteine count. The accuracy of cysteine counting via the developed method has been confirmed through analysis of a human (K562) protein extract spiked with Lysozyme (8 cysteines). To demonstrate the efficacy of the technique, we have applied the developed NeuCode chemical labeling method, along with a sample-specific proteogenomic database, for the identification of prostate tumor proteoforms. Preliminary data analysis of a subset of the data identified 62 proteoforms at a 39% FDR. Although the number of proteoforms from this proof-of-principle experiment is low relative to a deep top-down proteomics experiment, further development is possible. Improvements in instrument technology and the development of sample-specific proteoform databases, or proteoform

atlases, have the potential to enable efficient proteoform characterization without the additional analysis time and complexity of top-down mass spectrometry.

P20.06 | The epigenetic response of brown adipose tissue to cold stress: histone proteoform, RRBS, and RNAseq analysis

Nicolas L. Young[1] (nicolas.young@bcm.edu), Bethany C. Taylor[1], Alli M. Antar[1]
[1]*Baylor College of Medicine, Houston, TX*

Regulation of the thermogenic response by brown adipose tissue (BAT) is of interest in the treatment of obesity and diabetes. We hypothesized that chronic activation of BAT results in epigenetic modifications that affect the expression of genes regulating thermogenesis. Wildtype male C57BL/6J mice were housed under chronic conditions of thermoneutral temperature (28.8°C), mild cold/room temperature (22°C), or severe cold (8°C). Carbohydrate response element-binding protein (ChREBP), a transcription factor and negative regulator of BAT thermogenesis, KO mice were developed to interrogate its function. We developed methods for the isolation and quantitation of histone proteoforms from BAT. We observed distinct on/off histone signals for histone H3.2 in BAT, but not in the liver, when mice are acclimated to severe cold. Specifically, we observe a larger percentage of histone H3.2 with the strongly repressive proteoforms with K9me2 coupled in cis with K36me1 and K23ac (4.5 and 2 percent, 8°C and 28.8°C). This repression of previously active chromatin is likely an effect of decreased expression of G9a/GLP, a histone H3K9 dimethyltransferase. This is supported by consensus transcriptional regulatory networks using publicly available RNA-seq data.

With ChREBP manipulation, histone H4 proteoforms containing {K5ac,K16ac,K20me1} influenced 4 percent of the genome with an additional combined 7-fold change increase with ChREBP knockout (p=0.001, 0.05). Interestingly, the proteoform containing only K16ac significantly increased (p=0.006, 2-fold change), affecting one percent of the genome. Within BAT comparisons with H3, there were 299 proteoforms with significant (p<0.05) change and many were not present in ChREBP KO. An analysis of binary PTM combinations also implicate changes with K4me1/me2 and K9ac. This work establishes a new methodology to quantitatively study histone proteoforms in BAT, allowing for direct comparisons to decipher mechanistic changes during the thermogenic response. Our study reveals changes in epigenetic modifications in murine BAT in response to cold adaptation and ChREBP KO.

P20.07 | Revealing the Structures and Molecular Landscape of Low-Abundance Proteins by Top-Down Proteomics and Nanotechnology

David S. Roberts[1] (dsroberts@wisc.edu), Morgan W. Mann[1], Jake Melby[1], Eli J. Larson[1], Brad H. Li[1], Donguk Kim[1], Allan R. Brasier[1], Song Jin[1], Ying Ge[1], Emily Chapman[1]
[1]*University of Wisconsin-Madison, Madison, WI*

Top-down mass spectrometry (MS)-based proteomics is the premier technology for characterizing proteoforms to decipher post-translational modifications (PTMs) together with genetic variations and alternative splicing isoforms toward a proteome-wide understanding of biological functions. Many strides have been made in the past decade to enable the application of top-down proteomics for understanding basic biological functions, unraveling disease mechanisms, and discovering new biomarkers. However, major challenges related to the proteome dynamic range, proteome complexity, and establishing proteoform–function relationships persist and limit the broader application of the top-down strategy for challenging biological systems. This talk will describe how, by bridging the diverse silos of nanotechnology and MS-based proteomics, novel, unconventional, and effective solutions to addressing these challenges emerge. In one tale, this talk will detail how rationally designed surface-functionalized multivalent superparamagnetic nanoparticles (NPs) can be used as a general affinity platform to capture and enrich low-abundance proteoforms with high specificity for top-down MS applications. I will highlight the first generation use of this nanomaterials platform for the enrichment of low-abundance phosphoproteins, the recent development of a “nanoproteomics” platform to enrich cardiac troponin I (cTnI), the gold-standard biomarker for acute myocardial infarction (also known as a heart attack), from human serum with high specificity and sensitivity, and the ongoing work for targeting important endogenous receptor membrane glycoproteins. In the second tale, I will introduce new methodologies leveraging ultrahigh-resolution Fourier transform ion cyclotron resonance (FTICR)-MS, using a 12T solariX FTICR, and trapped ion mobility spectrometry (TIMS), using a timsTOF, to advance the comprehensive molecular characterization of complex native glycoproteins by top-down MS. This hybrid top-down MS approach is capable of providing detailed molecular insights to characterizing the structures and heterogeneity of complex glycoproteins, and specific examples involving variants of the SARS-CoV-2 Spike receptor-binding domain (S-RBD) protein will be illustrated.

P20.08 | Utilization of Electron Activated Dissociation (EAD) on a New QTOF Platform for Comprehensive Analysis of Histone Post-translational Modifications

Emily Zahn[1] **(zahn@wustl.edu)**, Richard M. Searfoss[1], Yixuan (Axe) Xie[1], Zongtao Lin[1], Faith M. Robison[1], Benjamin A. Garcia[1]
[1]*Washington University School of Medicine in St. Louis, St. Louis, MO,*

Histones are small proteins whose post-translational modifications (PTMs) are known to regulate gene expression patterns in an epigenetic manner. These modifications now include over 75 distinct types, and also have been localized to over 200 sites across the 5 histone proteins (H1, H2A, H2B, H3 and H4). Bottom-up mass spectrometry has played a key role in detection of novel and quantification of known histone PTMs. However, characterization of the combinations of histone modifications that occur simultaneously still remains a challenge, especially at the intact protein level. Different fragmentation approaches such as electron capture dissociation (ECD) and electron transfer dissociation (ETD) have been shown to be advantageous for sequencing longer histone polypeptides or intact histone protein. Here we describe the use of electron activated dissociation (EAD) on the new Sciex 7600 ZenoTOF for the characterization of histone PTMs. Optimization of many parameters such as the electron kinetic energy, accumulation time and reaction time will be discussed. Data with different EAD energy allowed for fragmentation of different sizes of histone polypeptides, and for high confidence identification of modified PTM isomers even at the intact protein level. Additionally, we found that high energies (10eV) which produce electron impact excitation of organic ions (EIEIO) can be used to fragment the modified sidechains of the histone peptides for more confident identification. Overall, this new type of fragmentation approach was found to be extremely useful for a comprehensive histone PTM characterization.