US HUPO 2021 VIRTUAL CONFERENCE

ABSTRACT BOOK

MARCH 8-12, 2021

US HUPO
from genes to function

WWW.USHUPO.ORG

US HUPO 2021 Virtual Conference Agenda-at-a-Glance

View full agenda details here: https://www.ushupo.org/Abstract-Book/

ALL TIMES IN CENTRAL TIME (GMT-6) | Virtual Platform Locations are Noted by (Room)

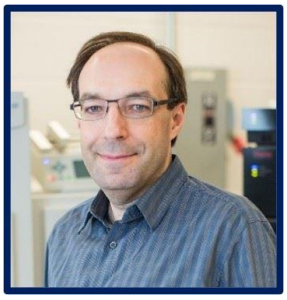
MONDAY, MARCH 8	TUESDAY, MARCH 9	WEDNESDAY, MARCH 10	THURSDAY, MARCH 11	FRIDAY, MARCH 12
9:00 AM – 9:45 AM **Morning Yoga** *(Navy Pier)* **Coffee Chat** *(Millennium Park)*	9:00 AM – 9:45 AM | Sponsor Seminars **Evosep Seminar** *(Ballroom 1)* **Promega Seminar** *(Ballroom 2)*	9:00 AM – 9:45 AM | Sponsor Seminars **Newomics Seminar** *(Ballroom 1)* **ReSyn Biosciences Seminar** *(Ballroom 2)*	9:00 AM – 9:45 AM | Sponsor Seminars **Olink Seminar** *(Ballroom 1)* **Cellenion Seminar** *(Ballroom 2)*	9:00 AM – 9:45 AM **Morning Yoga** *(Navy Pier)* **Coffee Chat** *(Millennium Park)*
10:00 AM – 10:50 AM **Opening Plenary Session - Carol Robinson** *(Auditorium)*	10:00 AM – 10:50 AM **Donald F. Hunt Distinguished Contribution in Proteomics Award** *(Auditorium)*	10:00 AM – 10:50 AM **Trainee Award Winner Presentations** *(Auditorium)*	10:00 AM – 10:50 AM **Robert J. Cotter New Investigator Award** *(Auditorium)*	10:00 AM – 10:50 AM **Gilbert S. Omenn Computational Proteomics Awards** *(Auditorium)*
11:00 AM – 12:30 PM | Parallel Sessions **Cellular Signaling & Systems Biology** *(Ballroom 1)* **Single Cell Proteomics** *(Ballroom 2)*	11:00 AM – 12:30 PM | Parallel Sessions **Cardiovascular Proteomics** *(Ballroom 1)* **Biomarkers & Precision Medicine** *(Ballroom 2)*	11:00 AM – 12:30 PM | Parallel Sessions **Epigenetics and Chromatin Biology** *(Ballroom 1)* **Technology Innovation in Multi-omics** *(Ballroom 2)*	11:00 AM – 12:30 PM | Parallel Sessions **Top-Down Proteomics and Native MS** *(Ballroom 1)* **Protein Post-Translational Modifications** *(Ballroom 2)*	11:00 AM – 12:30 PM | Parallel Sessions **Stem Cell proteomics and Chemical Proteomics in Drug Discovery** *(Ballroom 1)* **New Strategies of Interactomics in Structural Biology** *(Ballroom 2)*
12:30 PM – 1:00 PM **Networking Break in Exhibit Hall**	12:30 PM – 1:00 PM **Networking Break in Exhibit Hall**	12:30 PM – 1:00 PM **Networking Break in Exhibit Hall**	12:30 PM – 1:00 PM **Networking Break in Exhibit Hall**	12:30 PM – 1:00 PM **Networking Break in Exhibit Hall**
1:00 PM – 1:45 PM | Sponsor Seminars **Sciex Seminar** *(Auditorium)*	1:00 PM – 1:45 PM | Sponsor Seminars **Bruker Seminar** *(Ballroom 1)*	1:00 PM – 1:45 PM | Sponsor Seminars **ThermoFisher Scientific Seminar** *(Ballroom 1)* **Biognosys Seminar** *(Ballroom 2)*	1:00 PM – 1:45 PM | Sponsor Seminars **Seer Seminar** *(Auditorium)*	1:00 PM – 2:00 PM **Catherine E. Costello Lifetime Achievement in Proteomics Award and Closing Remarks** *(Auditorium)*
2:00 PM – 3:30 PM | Parallel Sessions **Proteomics of Infectious Disease** *(Ballroom 1)* **Omics of the Central Nervous System** *(Ballroom 2)*	2:00 PM – 3:30 PM **Lightning Talks – Round 1** *(Auditorium)*	2:00 PM – 3:30 PM **Lightning Talks – Round 2** *(Auditorium)*	2:00 PM – 3:30 PM | Parallel Sessions **Multi-'omics - Methods and Applications** *(Ballroom 1)* **Machine Learning and Computational Proteomics** *(Ballroom 2)*	ON-DEMAND | Sponsor Seminar **Agilent Technologies Seminar** *(Magnificent Mile Foyer)* *This session is viewable at any time during the Conference.*
3:45 PM – 5:15 PM | Evening Workshops **Interactive Career Planning and Development** *(Navy Pier)* **Top-down Proteomics: Tips and Tricks for Beginners** *(Ballroom 1)* **Biomarkers For Early Detection** *(Auditorium)*	3:45 PM – 5:15 PM **Poster Session** *(Poster Rooms)* *All presenters should be at their posters during this time.*	3:45 PM – 5:15 PM | Evening Workshops **Cardiovascular Proteomics in Translational and Clinical Research** *(Ballroom 1)* **How to Bring the Impact of Multi-omics Research to the Next Level? From Industry and Academic Perspectives** *(Ballroom 2)*	3:45 PM – 5:15 PM | Evening Workshops **Putting Humpty Dumpty Back Together Again: What Does Protein Quant Mean in Bottom-Up Proteomics?** *(Auditorium)* **Targeted Mass Spectrometric Assays for Diabetes and Obesity Research** *(Navy Pier)*	2021 VIRTUAL 17TH ANNUAL CONFERENCE US HUPO March 8–12, 2021
5:30 PM – 7:30 PM **Evening Social Time on Own** *(Magnificent Mile Foyer)*	5:30 PM – 7:30 PM **Company and Lab Pitches** *(Navy Pier)*	5:30 PM – 7:30 PM **ECR and FeMS Social Event** *(Millennium Park)*	5:30 PM – 7:30 PM **Virtual Cocktail Party (BYOB!)** *(Magnificent Mile Foyer)*	

COMMITTEES

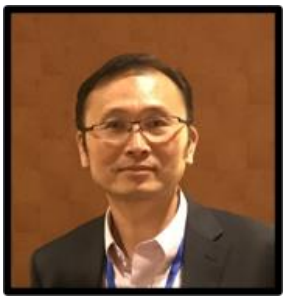
CONFERENCE COMMITTEE

Ying Ge, PhD
University of Wisconsin-Madison

Stephanie Cologna, PhD
University of Illinois-Chicago



Alexey Nesvizhskii, PhD
University of Michigan



Yingming Zhao, PhD
The University of Chicago

BOARD OF DIRECTORS

PRESIDENT
Jennifer Van Eyk, Cedars-Sinai Medical Center

PAST PRESIDENT
Ileana Cristea, Princeton University

TREASURER
Birgit Schilling, Buck Institute for Research on Aging

SECRETARY
David Arnott, Genentech

Robert Chalkley, University of California San Francisco
Joshua Coon, University of Wisconsin, Madison
Catherine Costello, Boston University
David Fenyo, New York University
Benjamin Garcia, University of Pennsylvania School of Medicine
Timothy Griffin, University of Minnesota
Rebekah Gundry, University of Nebraska Medical Center
Sonja Hess, AstraZeneca
Lan Huang, University of California, Irvine
Lingjun Li, University of Wisconsin, Madison
Joseph Loo, University of California, Los Angeles
Michael MacCoss, University of Washington
Robert Moritz, Institute for Systems Biology
David Muddiman, North Carolina State University
Peter Nemes, University of Maryland
Gil Omenn, University of Michigan, Ann Arbor
Laurie Parker, University of Minnesota
Peipei Ping, UCLA
Karin Rodland, Pacific Northwest National Laboratory
Henry Rodriguez, NIH, NCI
Devin Schweppe, University of Washington
Michael Snyder, Stanford University
Hanno Steen, Harvard University
W. Andy Tao, Purdue University
Judit Villen, University of Washington
Olga Vitek, Northeastern University
Christine Vogel, NYU
John Yates, The Scripps Research Institute

US HUPO
from genes to function

2021 VIRTUAL

SUPPORTERS

The Co-chairs would like to express their gratitude to the below sponsors for their generous financial support:

PREMIER SPONSOR

BRUKER

GOLD SPONSORS

ThermoFisher SCIENTIFIC
The world leader in serving science

BIOGNOSYS
NEXT GENERATION PROTEOMICS

Agilent Technologies

seer

CONFERENCE SPONSORS

abbvie

Genentech
A Member of the Roche Group

TRAINEE AWARD SPONSORS

seer

Molecular Omics

Buck
Live better longer.

US HUPO
from genes to function

2021 VIRTUAL

SUPPORTERS

SEMINAR SPONSORS

SCIEX

NEWOMICS

EVOSEP

ReSyn Biosciences™

Olink

Promega

The Novo Nordisk Foundation
Center for Protein Research

EXHIBITORS

Visit the following organizations in the Exhibit Hall:

Agilent Technologies
Biognosys AG
Bruker
CoAnn Technologies, LLC
Covaris
ESI Source Solutions, Inc.
Newomics
Shimadzu Scientific Instruments, Inc.
Waters

US HUPO
from genes to function

2021 VIRTUAL

MONDAY, MARCH 8

(all times in Central Time)

9:00 AM – 9:45 AM — **Morning Activities**

Morning Yoga
Navy Pier

Join us in the Navy Pier room for a beginner yoga session to gear you up for a day full of sessions! No experience required.

Coffee Chat
Millennium Park

Not a morning person? Bring your own beverage and join your fellow attendees in the Millennium Park room to catch up over coffee.

10:00 AM – 10:50 AM — **Opening Plenary Session** — **Auditorium**

Session Chair: Ying Ge, University of Wisconsin-Madison, Madison, WI

Welcome Remarks
Jennifer Van Eyk, US HUPO President, Cedars-Sinai Medical Center, Los Angeles, CA

Mass Spectrometry: From Plasma Proteins to Mitochondrial Membranes
Carol Robinson, University of Oxford, Oxford, United Kingdom

11:00 AM – 12:30 PM — **PARALLEL SESSION** — **Ballroom 2**

Single Cell Proteomics

Session Chairs: Nikolai Slavov, Northeastern University, Boston, MA and Ying Zhu, PNNL, Richland, WA

11:00 AM – 11:30 AM
Single-Cell Proteomics: Practical Considerations for Unlocking the Proteome One Cell at a Time
Ryan Kelly, Brigham Young University, Utah, United States

11:30 AM – 12:00 PM
Mass Spectrometry Gone Subcellular to Drive Developmental Neurobiology
Peter Nemes, University of Maryland, Maryland, United States

12:00 PM – 12:15 PM
01.01: Principled and Reproducible Mass Spectrometry-based Single Cell Proteomics Data Analysis
Laurent Gatto and Christophe Vanderaa

12:15 PM – 12:30 PM
01.02: Robust, very low nanoflow LC-MS methods for ultra-sensitive single-cell and phosphoproteomics
Dorte B. Bekker-Jensen, Ole B. Hørning, Andreas-David Brunner, Marvin Thielert, Jesper V. Olsen, Ole Vorm, Matthias Mann and Nicolai Bache

11:00 AM – 12:30 PM	**PARALLEL SESSION** **Cellular Signaling & Systems Biology**	**Ballroom 1**

Session Chairs: Andy Tao, Purdue University, Lafayette, IN and Matthew Champion, University of Notre Dame, Notre Dame, IN

11:00 AM – 11:30 AM
Control of Nuclear Integrity via Lamin Acetylation During Cell Cycle Progression and Viral Infections
Ileana Cristea, Princeton University, New Jersey, United States

11:30 AM – 12:00 PM
Proteomics-Based Systems Biology of Human Disease
Junming Peng, St. Jude's Hospital, Tennessee, United States

12:00 PM – 12:15 PM
02.01: Phosphorylation-driven Regulation of IFI16 Viral DNA Sensing at the Nuclear Periphery
Dawei Liu, Krystal Lum, Joshua Justice and Ileana Cristea

12:15 PM – 12:30 PM
02.02: Protein Communities Reveal New Members When Recognizing the Deeper Coding Potential of Transcripts
Sebastien Leblanc, Marie A Brunet, Jean-François Jacques, Amina M Lekehal, Alan A Cohen, Michelle S Scott and Xavier Roucou

12:30 PM – 1:00 PM	**Networking Break**	**Exhibit Hall**

Visit the virtual exhibit hall to connect with our exhibitors.

1:00 PM – 1:45 PM	**Sponsored Seminar**	**Auditorium**

Sponsored By:

SCIEX

Site-specific analysis of N-glycosylation of the biotherapeutics by CESI-MS (CE-MS)
Miloslav Sanda, Georgetown University, Washington, D.C., United States

2:00 PM – 3:30 PM	**PARALLEL SESSION** **Omics of the Central Nervous System**	**Ballroom 2**

Session Chairs: Amanda Myers, University of Miami, Coral Gables, FL and Ling Hao, George Washington University, Washington, DC

2:00 PM – 2:30 PM	**Multi-dimensional 'Omics of Neurodegenerative Diseases** Michael Ward, National Institute of Health, Maryland, United States
2:30 PM – 3:00 PM	**Comprehensive Proteomics and Lipidomics Strategies to Advance Disparities Research in Alzheimer's Disease** Rena Robinson, Vanderbilt University, Tennessee, United States
3:00 PM – 3:15 PM	**03.01: Deconvoluting the Effects of Age and Disease on Protein Aggregation in Mouse Models** Cristen Molzahn, Erich Kuechler, Irina Zemlyankina, Lorenz Nierves, Aly Karsan, Philipp Lange and Thibault Mayor
3:15 PM – 3:30 PM	**03.02: A dynamic mouse peptidome landscape reveals probiotic modulation of the gut-brain axis** Chenxi Jia

PARALLEL SESSION
Proteomics of Infectious Disease

2:00 PM – 3:30 PM — **Ballroom 1**

Session Chairs: Aleksandra Nita-Lazar, NIH, Bethesda, MD and Yi Shi, University of Pittsburgh, Pittsburgh, PA

2:00 PM – 2:30 PM	**Host Factor Targeted Drug Discovery for SARS-CoV-2 Through an International Collaboration** Nevan Krogan, University of California San Francisco, California, United States
2:30 PM – 3:00 PM	**A Neighboring View of the SARS-CoV-2 Proximal Interactome** Anne-Claude Gingras, University of Toronto, Canada
3:00 PM – 3:15 PM	**O04.01: Versatile and Multivalent Nanobodies Efficiently Neutralize SARS-CoV-2** Yufei Xiang, Sham Nambulli, Zhengyun Xiao, Heng Liu, Zhe Sang, W. Paul Duprex, Dina Schneidman-Duhovny, Cheng Zhang and Yi Shi
3:15 PM – 3:30 PM	**O04.02: Defining Mechanisms Underlying Virus Regulation of Mitochondrial Bioenergetics During Infection** Cora Betsinger, Joel Federspiel, Connor Jankowski, William Hofstadter and Ileana Cristea

EVENING WORKSHOP
Interactive Career Planning and Development

3:45 PM – 5:15 PM — **Navy Pier**

Presented By: Laurie Parker, Associate Professor, University of Minnesota, MN

9:00 AM – 9:45 AM | **Sponsored Seminar** | **Ballroom 1**

Sponsored By:

Promega

The Novo Nordisk Foundation Center for Protein Research

ProAlanase is an Effective Alternative to Trypsin for Proteomics Applications and Disulfide Bond Mapping
Diana Samodova, Novo Nordisk Foundation Center for Basic Metabolic Research, Denmark

9:00 AM – 9:45 AM | **Sponsored Seminar** | **Ballroom 2**

Sponsored By:

EVOSEP

Ultra Sensitive and High Throughput Proteomics on the Standardized Evosep One Platform
Jesper Olsen, Group Leader at Olsen Group, University of Copenhagen

Erwin M. Schoof, Associate Professor in Single-Cell Proteomics In Disease Biology / Head Of DTU Proteomics Core, DTU – Technical University Of Denmark

10:00 AM – 10:50 AM | **Award Plenary Session** | **Auditorium**

Session Chair: John Yates, The Scripps Research Institute, LaJolla, CA, United States

Donald F. Hunt Distinguished Contribution in Proteomics Award Presentation
Peipei Ping, University of California Los Angeles, California, United States

11:00 AM – 12:30 PM | **PARALLEL SESSION** | **Ballroom 2**

Biomarkers & Precision Medicine

Session Chairs: Joshua LaBaer, Arizona State University, Tempe, AZ and Yansheng Liu, Yale University, New Haven, CT

11:00 AM – 11:30 AM

Beyond Genomics; Proteogenomic Insights Into Cancer Biology
Shankha Satpathy, Broad Institute, Massachusetts, United States

11:30 AM – 12:00 PM	**To Understand One, We Must Study Many - The Need for High-throughout Proteomics and Good Cohort Design** Hanno Steen, Harvard Medical School, Massachusetts, United States
12:00 PM – 12:15 PM	**05.01: Comparative Protein Interaction Analysis of Genes Altered in Head and Neck Cancer Instructs the Selection of Therapeutic Targets** Danielle Swaney, Dana Steffen, Zhiyong Wang, Jisoo Park, Yusuke Goto, Margaret Soucheray, Neil Bhola, Kyumin Kim, Fan Zheng, Yan Zeng, Rachel O□Keefe, John Von Dollen, Mehdi Bouhaddou, Jason Kreisberg, Minkyu Kim, Daniel Johnson, Natalia Jura, Jennifer Grandis, J. Silvio Gutkind, Trey Ideker and Nevan Krogan
12:15 PM – 12:30 PM	**05.02: Proteogenomic analysis unveils the HLA Class I presented immunopeptidome in melanoma and EGFR mutant lung adenocarcinoma** Yue Andy Qi, Tapan Maity, Constance Cultraro, Vikram Misra, Xu Zhang, Catherine Ade, Shaojian Gao, David Milewski, Khoa Nguyen, Mohammad Ebrahimabadi, Ken-ichi Hanada, Javed Khan, Cenk Sahinalp, James Yang and Udayan Guha

11:00 AM – 12:30 PM — PARALLEL SESSION — Ballroom 1

Cardiovascular Proteomics

Session Chairs: Edward Lau, University of Colorado School of Medicine, Aurora, CO and Melanie White, University of Sydney, Sydney, Australia

11:00 AM – 11:30 AM	**Identifying Translated Alternative Protein Isoforms** Maggie Lam, University of Colorado Denver, Colorado, United States
11:30 AM – 12:00 PM	**The Extracellular Matrix Proteome in Cardiac Physiology** Merry Lindsey, University of Nebraska Medical Center, Nebraska, United States
12:00 PM – 12:15 PM	**06.01: Induced Pluripotent Stem Cell-derived Engineered Cardiac Tissue Model of Hypertrophic Cardiomyopathy Enabled by Integrated Functional Assessments and Top-down Proteomics** Jake Melby, Willem de Lange, Jianhua Zhang, David Roberts, Stanford Mitchell, Trisha Tucholski, Gina Kim, Andreas Kyrvasilis, Sean McIlwain, Timothy Kamp, J. Carter Ralphe and Ying Ge
12:15 PM – 12:30 PM	**06.02: Distinctive Mechanisms in Hypertrophy and Heart Failure** Aleksandra Binek, Justyna Fert-Bober, Daniel Soetkamp, Simion Kreimer, Alejandro Rivas, Anna Pyzel, Amy Bradshaw, Michael Zile and Jennifer Van Eyk

12:30 PM – 1:00 PM — Networking Break — Exhibit Hall

Visit the virtual exhibit hall to connect with our exhibitors.

1:00 PM – 2:00 PM | **Sponsored Seminar** | **Ballroom 1**

Sponsored By:

BRUKER

PASEF in Proteomics of Health and Disease
Christopher Adams, Bruker Daltonics, California, United States

Kirk Hansen, University of Colorado, Colorado, United States

2:00 PM – 3:30 PM | **Lightning Talks – Round 1** | **Auditorium**

Session Chair: Alexey Nesvizhskii, University of Michigan, Ann Arbor, Michigan, United States

P10.01 **A single-pot ultrasensitive heterogeneity analysis of viral spike N-glycoforms**
Sabyasachi Baboo, Jolene Diedrich, Salvador Martínez-Bartolomé, Xiaoning Wang, Torben Schiffner, William Schief, James Paulson and John Yates

P11.01 **Three-Dimensional (3D) Imaging of Lipids in Skin Tissues with Infrared Matrix-assisted Laser Desorption Electrospray Ionization (MALDESI) Mass Spectrometry**
Honxgia Bai, Keith Linder and David Muddiman

P01.02 **TurnoveR: A Skyline External Tool for the Analysis of Protein Turnover from Metabolic Labeling Studies**
Nathan Basisty, Cameron Wehrfritz, Nicholas Shulman, Alexandra Marsh, Michael MacCoss, Brendan MacLean and Birgit Schilling

P06.01 **Proteo-Metabolomic Study of the Spemann's Organizer In the Vertebrate (Frog) Embryo**
Aparna B. Baxi, Vi M. Quach, Jie Li and Peter Nemes

P12.01 **A Novel Multi-omics Strategy for Human Pluripotent Stem Cell-derived Cardiomyocytes**
Elizabeth F. Bayne, Aaron Simmons, David S. Roberts, Yanlong Zhu, Timothy J. Aballo, Sean Palecek and Ying Ge

P04.01 **Selective Pulse Chase-SILAC Labeling of Three-Dimensional Multicellular Spheroids for Global Proteome Analysis**
Nicole Beller, Jessica Lukowski and Amanda Hummon

P05.03 **Quantitative Analysis Of Cell Type and Chamber Specific Cardiomyocytes**
Chase Castro, Melinda Wojtkiewicz and Rebekah Gundry

P12.02 **Confirmatory Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma**
David Clark and Yize Li

P14.01 **Spatiotemporal virus-host interactions viewed through the lens of Inter-ViSTA**
Katelyn Cook, Joel Federspiel, Michelle Kennedy, Samvida Venkatesh, Clayton Otter, William Hofstadter and Ileana Cristea

P19.01 **Quantitative Top-down Thermal Proteome Profiling of E. coli lysate and Standard Proteins**
Kellye Cupp-Sutton, Thomas Welborn, Ji Won Kang and Si Wu

P16.01 **Comprehensive Stoichiometric Characterization of the Thiol Redox Proteome**
Tong Zhang, Nicholas Day, Matthew Gaffrey, Rui Zhao, Garett Grant, Tujin Shi, Thomas Fillmore, George Rodney, My Helms and Wei-Jun Qian

P08.02 **Peptide Correlation Analysis (PeCorA) Reveals Differential Proteoform Regulation**
Maria Dermit and Jesse Meyer

P01.05 **Proteograph's multi nanoparticle proteins coronas enable deep plasma proteomics studies at scale with unmatched sensitivity in combination with trapped ion mobility**
Shadi Ferdosi, Tristan Brown, Patrick Everley, Mike Figa, Matthew McLean, Eltaher Elgierari, Xiaoyan Zhao, Veder Garcia, Tianyu Wang, Kateryna Riedesel, Jessica Chu, Martin Goldberg, Asim Siddiqui, Juan Cruz Cuevas, John Blume, Daniel Hornburg and Omid Fa

P16.02 **Multiplexed Quantification of Insulin and C-peptide Without the Use of Antibodies**
North Foulon, Elisha Goonatilleke, Michael MacCoss, Michelle Emrick and Andrew Hoofnagle

P16.03 **Proteomic Evaluation of the Dynamic Protein Turnover in Human iPSC-derived Neurons**
Ashley Frankenfield, Saadia Hasan, Michael Ward and Ling Hao

P13.02 **Relating Metabolic Control with Lysine Acylation in an Exceptional Metabolically Versatile Bacterium**
Janine Y. Fu, Robert P. Gunsalus, Rachel R. Ogorzalek Loo and Joseph A. Loo

P01.07 **A Proteomic sample preparation for mass spectrometry using an automated workstation**
Qin Fu, Johnson Casey, Danica-Mae Manalo and Jennifer Van Eyk

P19.02 **A novel tissue extraction technique combined with high-field asymmetric waveform ion mobility in top-down proteomics provides new insights into Alzheimer's disease**
James Fulcher, Aman Makaju, Ronald Moore, Mowei Zhou, David Bennett, Phil De Jager and Vladislav Petyuk

P08.06 **Predicting novel interactions between kinases and substrates using graph representation learning**
Sachin Gavali, Karen Ross, Julie Cowart, Chuming Chen and Cathy Wu

P12.04 **OpenCMS: a tool to build personalized proteogenomics databases and detect unannotated proteins and possible biomarkers.**
Noé Guilloy, Sébastien Leblanc, Jean-François Jacques, Marie Brunet, Claude Perreault and Xavier Roucou

P19.03 **Optimization of Protein-Level Tandem Mass Tag (TMT) Labeling in Complex Sample with Top-Down Proteomics**
Yanting Guo, Dahang Yu, Kellye A. Cupp-Sutton, Xiaowen Liu and Si Wu

P20.03 **Quantitative proteomics and phosphoproteomics of urinary extracellular vesicles define diagnostic and prognostic biosignatures for Parkinson's Disease**
Marco Hadisurya, Li Li, Xiaofeng Wu, Roy N. Alcalay, Shalini Padmanabhan, Anton Iliuk and W. Andy Tao

P03.03 **Label-free Mass Spectrometry Identification of Vascular Integrity Markers in Major Trauma**
Holly Hinson, Jon Jacobs, Shannon McWeeney, Tujin Shi, Kendall Martin and Karin Rodland

P05.05 **Cardiac sex disparities are established via post-transcriptional regulation**
Josiah Hutton, Wei Shi, Xinlei Sheng, Kerry Dorr, Haley Davies, Tia Andrade, Todd Grecco, Yutaka Hashimoto, Joel Federspiel, Zachary Robbe, Xuqi Chen, Arthur Arnold, Ileana Cristea and Frank Conlon

P06.04 **PARP1 inhibitors trigger innate immunity via PARP1 trapping-induced DNA damage response**
Chiho Kim, Xu-Dong Wang and Yonghao Yu

P01.12 **Photocleavable Surfactant-Enabled Extracellular Matrix Proteomics**
Samantha Knott, Kyle Brown, Harini Josyer, David Inman, Suzanne Ponik, Andreas Friedl and Ying Ge

P16.04 **MSstatsPTM: an R/Bioconductor software for detecting quantitative changes in post-translational modifications**
Devon Kohler, Tsung-Heng Tsai, Ting Huang, Erik Verschueren, Trent Hinkle, Meena Choi and Olga Vitek

P09.03 **OpenTIMS, TimsPy, and TimsR: Quick and Friendly Access to timsTOF Pro Data**
Mateusz Krzysztof Łącki, Michał Piotr Startek, Sven Brehmer, Christina Bell, Ute Distler and Stefan Tenzer

P02.04 **Proteomic Investigation of the I1061T Point Mutation Mouse Model of Niemann-Pick Type C**
Wenping Li, Melissa Pergande and Stephanie Cologna

3:30 PM – 5:15 PM **Poster Session** **Poster Rooms**

All posters will be presented in the virtual poster hall during this time.

5:30 PM – 7:30 PM **EVENING WORKSHOP**
Company and Lab Pitches **Navy Pier**

WEDNESDAY, MARCH 10

(all times in Central Time)

9:00 AM – 9:45 AM | **Sponsored Seminar** | **Ballroom 2**

Sponsored By:

ReSyn Biosciences™

Proximity Mapping of Leishmania Kinetochore Kinases
Vincent Geoghegan, The University of York, England, United Kingdom

GlycoDIA: Gene Editing and Advanced Mass Spectrometry to Bridge the Next Generation O-glycoproteomics
Sergey Vakhrushev, University of Copenhagen, Denmark

9:00 AM – 9:45 AM | **Sponsored Seminar** | **Ballroom 1**

Sponsored By:

NEWOMICS

A New LC-MS Platform for Clinical Proteomics
Na Pi Parra, Senior Director of Marketing, Newomics, CA, United States

10:00 AM – 10:50 AM | **Trainee Award Winner Presentations** | **Auditorium**

Session Chairs: Nathan Basisty, Buck Institute for Research on Aging, San Francisco, CA, United States and Lindsay K Pino, University of Pennsylvania, Philadelphia. PA, United States

10:00 AM – 10:12 AM
Top-Down Nanoproteomics Enables Comprehensive Analysis of Low-Abundance Cardiac Troponin I Proteoforms from Human Serum
David S. Roberts, Timothy N. Tiambeng, Kyle A. Brown, Yanlong Zhu, Bifan Chen, Zhijie Wu, Stanford D. Mitchell, Tania M. Guardado-Alvarez, Song Jin, Ying Ge

10:12 AM – 10:24 AM
Optimization of Pervanadate BOOST Channel in TMT Multiplexing Delivers Deeper Quantitative Characterization of the Tyrosine Phosphoproteome
Xien Yu Chua, Ricky Edmondson, Arthur Salomon

10:24 AM – 10:36 AM
caAtlas: an Immunopeptidome Atlas of Human Cancer
Xinpei Yi, Yuxing Liao, Bo Wen, Kai Li, Yongchao Dou, Sara Savage, Bing Zhang

10:36 AM – 10:48 AM
A high dimensional map of the budding yeast phosphorylation-dependent signaling network
Mario Leutert, Noelle Fukuda, Judit Villen

11:00 AM – 12:30 PM	**PARALLEL SESSION** **Technology Innovation in Multi-omics**	**Ballroom 2**

Session Chairs: David Muddiman, North Carolina State University, Raleigh, NC and Yu Tian, Abbvie, Inc, Worcester, MA

11:00 AM – 11:30 AM
Structural Biology on the Proteome-Wide Scale: An In-Cell Mass Spectrometry-Based Protein Footprinting Method
Lisa Jones, University of Maryland, Maryland, United States

11:30 AM – 12:00 PM
Multiplex Chemical Tags for High-Throughput Comparative Multiomic Analyses of Cerebrospinal Fluids in Alzheimer's Disease
Lingjun Li, University of Wisconsin Madison, Wisconsin, United States

12:00 PM – 12:15 PM
07.01: A Multi-Omic Investigation into the Role of the APOE Genotype in Alzheimer's Disease
Erin Baker, Melanie Odenkirk, Kristin Burnum-Johnson, Anna Lipton, Karen Butler, Jaclyn Kalmar, Brendan MacLean, Michael MacCoss, Edward Fox, Thomas Montine, David Muddiman and Blaine Roberts

12:15 PM – 12:30 PM
07.02: Uncovering Isoform Diversity of the Human Proteome by Long-Read Proteogenomics
Rachel Miller, Simi Kaur, Ben Jordan, Rob Millikin, Kyndalanne Pike, Isabella Whitworth, Michael Shortreed, Christina Chatzipantsiou, Raymond LeClair, Anne Deslattes Mays, Lloyd Smith and Gloria Sheynkman

11:00 AM – 12:30 PM	**PARALLEL SESSION** **Epigenetics and Chromatin Biology**	**Ballroom 1**

Session Chairs: Jun Qin, Baylor College of Medicine, Waco, TX and Michael Freitas, The Ohio State University, Columbus, OH

11:00 AM – 11:30 AM
Quantitative Proteomics for Understanding Genome Mutations that Affect Human Health
Ben Garcia, University of Pennsylvania, Pennsylvania, United States

11:30 AM – 12:00 PM
Functional Characterization of Lysine Lactylation and Hydroxybutyrylation
He Huang, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, China

12:00 PM – 12:15 PM
08.01: Elucidating the Mechanism of Germline Histone H3.3 Mutations on Neurodevelopment
Khadija D. Wilson, Elizabeth Porter, Peter Klein, Elizabeth Bhoj and Benjamin A. Garcia

12:15 PM – 12:30 PM

08.02: Improving dimethyl labeling for its application to histone lysine methylation quantitative analysis
Junfeng Huang, Ngai Ting Chan, Zexin Zhu, Shengjie Zhang, Wei Xu and Lingjun Li

---

12:30 PM – 1:00 PM | **Networking Break** | **Exhibit Hall**

Visit the virtual exhibit hall to connect with our exhibitors.

1:00 PM – 1:45 PM | **Sponsored Seminar** | **Ballroom 1**

Sponsored By:

ThermoFisher SCIENTIFIC
The world leader in serving science

Improved Sensitivity and Depth of Proteomic and PTM Profiling Using Ultra-low Flow Separation Techniques and FAIMS Technology Interfaced with Advanced Mass Spectrometry
Alexander R. Ivanov, Northeastern University, Massachusetts, United States

---

1:00 PM – 1:45 PM | **Sponsored Seminar** | **Ballroom 2**

Sponsored By:

BIOGNOSYS
NEXT GENERATION PROTEOMICS

Simplify your Targeted Proteomics Analysis with SpectroDive
Tejas Gandhi, Biognosys
Simone Di Sanzo, Leibniz Institute on Aging
Maximilian Helf, Biognosys

2:00 PM – 3:30 PM | **Lightning Talks – Round 2** | **Auditorium**

Session Chair: Stephanie Cologna, University of Illinois at Chicago, Chicago, IL, United States

P18.01 **Towards Lossless Structural Mass Spectrometry (LSMS): All Ion Unfolding Ion Mobility-Mass Spectrometry for Improved Structural Characterization of Glycoproteins**
Gongyu Li and Lingjun Li

P16.05 **Proteome-wide and Matrisome-specific Alterations during Human Pancreas Development and Maturation Revealed by 12-plex Isobaric DiLeu Tags**

Zihui Li, Daniel Tremmel, Fengfei Ma, Qinying Yu, Min Ma, Daniel Delafield, Yatao Shi, Bin Wang, Samantha Mitchell, Austin Feeney, Vansh Jain, Sara Sackett, Jon Odorico and Lingjun Li

P08.08 **Improving power while controlling the false discovery rate when only a subset of peptides are relevant**
Andy Lin, Deanna Plubell, Uri Keich and William Noble

P16.08 **Label-Free Proteomics Performance with New Orbitrap Exploris 480 mass spectrometer with Single-Cell Sensitivity**
Daniel Lopez-Ferrer, Aaron S. Gajadhar, Aman Makaju, Aaron M. Robitaille, Julia Kraegenbring, Tabiwang N. Arrey, Michelle Dubuke, Emily Chen and Alexander Harder

P10.05 **Novel mucin enrichment strategy reveals molecular signatures of cancer in cellular systems and ovarian cancer patient ascites fluid**
Stacy Malaker, Nicholas Riley, Judy Shon, Kayvon Pedram and Carolyn Bertozzi

P15.05 **Discovery of RSV-induced BRD4 Protein Interactions using native immunoprecipitation and PASEF mass spectrometry.**
Morgan Mann, David Roberts, Ying Ge and Allan Brasier

P09.04 **MASH Explorer, a Universal, Comprehensive, and User-friendly Software Environment for Top-down Proteomics**
Sean J. McIlwain, David S. Roberts, Jake A. Melby, Kent Wenger, Eli J. Larson, Zhijie Wu, Anna Janicek, Molly Wetzel, Yiran Yan, Elizabeth F. Bayne, Kyndalanne A. Pike, Xiaowen Liu, Ruixiang Sun, Irene M. Ong and Ying Ge

P17.07 **Deep proteomic profiling of human spinal motor neurons using nanoPOTS reveals single-cell protein dynamics in ALS disease**
Santosh Misal, Amanda Guise, Thy Truong, Yiran Liang, Samuel Payne, Edward Plowey and Ryan Kelly

P18.02 **Effects of Charge versus Chemical Nature of Divalent Metal Ions on Alpha-Synuclein Structure**
Rani Moons, Albert Konijnenberg, Anne-Marie Lambeir and Frank Sobott

P01.15 **Global Phosphoproteome Analysis Using High-field Asymmetric Waveform Ion Mobility Spectrometry on a Hybrid Orbitrap Mass Spectrometer**
Laura Muehlbauer, Alexander Hebert, Michael Westphall, Evgenia Shishkova and Joshua Coon

P08.09 **FragPipe: A Graphical Interface for Fast and Flexible Proteomics Data Analysis**
Alexey Nesvizhskii, Dmitry Avtonomov, Fengchao Yu, Guo Ci Teo, Felipe Da Veiga Leprevost, Sarah Haynes, Hui-Yin Chang, Daniel Geiszler and Daniel Polasky

P01.16 **Preserving Trypsin Stability for Accelerated Digestion in the ProTrap XG**
Jessica Nickerson and Alan Doucette

P08.12 **Combining the Accurate Mass and Time Tag and Ion Mobility to Find Missing Peptides**
Robin Park, Sven Brehmer, Titus Jung, Patrick Garrett, Vijayaraja Gnanasambandan, Marc-Antoine Beauvais, Casimir Bamberger, Hyunsoo Kim, Christopher Adams, Dennis Trede, John Yates and Rohan Thakur

P02.05 **Dynamic changes of exosome proteins and post-translational modifications during cellular senescence and aging.**
Sandip Kumar Patel, Jacob Rose, Nathan Basisty, Lindsay Pino, Roland Bruderer, Lukas Reiter, Judith Campisi and Birgit Schilling

P02.06 **Brain Myelin Proteomics of Neurodegenerative Disorder Niemann-Pick type C1 Mouse Model**
Chandimal Pathmasiri and Stephanie Cologna

P04.09 **Oncogenic signaling drives histone lysine lactylation via the Warburg effect and MOF1 lactyltransferase activity**
Di Zhang, Lu Yang, Ruilong Liu, Jinjun Gao, Kevin Huang, Ziyuan Li, He Huang and Yingming Zhao

P16.10 **Investigating bortezomib-induced protein degradation with pulseSILAC-DIA**
Lindsay Pino, Josue Baeza, Richard Lauman, Birgit Schilling and Benjamin Garcia

P10.06 **An Improved Search Method for Glycosylation and Other Labile Post-translational Modifications using MSFragger-Glyco**
Daniel Polasky, Fengchao Yu, Guo Ci Teo and Alexey Nesvizhskii

P15.06 **A rigorous evaluation of optimal peptide targets for MS-based clinical diagnostics of Coronavirus Disease 2019 (COVID-19).**
Andrew Rajczewski, Subina Mehta, Dinh Duy An Nguyen, Björn Grüning, James Johnson, Thomas McGowan, Timothy Griffin and Pratik Jagtap

P06.06 **Cell Type-Selective Secretome Profiling In Vivo**
Nicholas Riley, Wei Wei, Andrew Yang, Joon Kim, Stephanie Terrell, Veronica Li, Marta Garcia-Contreras, Carolyn Bertozzi and Jonathan Long

P05.07 **The Effects of Fish Oil Supplementation on High-density Lipoprotein Efflux Capacity, Proteome, Lipidome and Inflammatory Effects**
Sara Rosario, Paul Mueller, Melissa Heard, Paige Bergstrom, Rayna Gasik and Nathalie Pamir

P13.09 **OxRAC Unveils Novel Role of LARP1 Along the Photosynthetic TOR Pathway**
Amanda Smythers, Sarah Lowery, Saher Mubarek, Hailey Lewis, Danielle Slone, Derrick Kolling and Leslie Hicks

P13.12 **Unraveling the Histone Code and Its Readers on Decondensed Heterochromatin**
Stephanie Stransky, Jennifer Aguilan, Edward Nieves and Simone Sidoli

P09.05 **PINE: An Automation Tool to Extract and Visualize Protein-Centric Functional Networks**
Niveda Sundararaman, James Go, Aaron E Robinson, José M Mato, Shelly C Lu, Jennifer Van Eyk and Vidya Venkatraman

P04.12 **The Proteome of the Microbiome in Peripheral Fluids: Application to Breast Cancer Detection**
Ngoc Vuong, Rayan Alhammad, Weidong Zhou, Neha Shetty, Jennifer May, Alessandra Luchini, Claudius Mueller, Virginia Espina, Lance Liotta and Richard Hoefer

P17.09 **Proteomic Profiling of a Metabolically Defined Mammary Stem Cell Enriched Population**
Matthew Waas, Pirashaanthy Tharmapalan, Rama Khokha and Thomas Kislinger

P04.13 **Combined Surface Proteome and Phosphoproteomics Reveals Novel Therapeutic Targets for Hepatocellular Carcinoma**
Shao Huan Samuel Weng, Hsiang-En Hsu, Junmin Wang, Raghothama Chaerkady, Lisa Cazares, Elaine Hurt and Sonja Hess

P08.22 **Features of human disease-associated mutations related to structure, stability, and binding**
Jaie Woodard, Chengxin Zhang and Yang Zhang

P20.09 **Qualitative and quantitative proteomic and metaproteomic analyses of normal human urine sediment**
XiaoLian Xiao, HaiDan Sun, Chen Shao and Wei Sun

P10.08 **O-live: An Interactive O-linked Glycosylation Database**
Yuanwei Xu and Hui Zhang

P08.23 **False Discovery Rate-Controlled Match-Between-Runs Enables Accurate and Sensitive Label Free Quantification**
Fengchao Yu, Sarah Haynes and Alexey Nesvizhskii

3:45 PM – 5:15 PM

EVENING WORKSHOP
Cardiovascular Proteomics in Translational and Clinical Research

Ballroom 1

Presented By: Ron Heeren, Maastricht University, Netherlands
Joshua LaBaer, Arizona State University, Arizona, United States
Edward Lau, University of Colorado School of Medicine, Colorado, United States
Sarah Parker, Cedars-Sinai Medical Center, California, United States
Jonathan Kirk, Loyola University, California, United States
Lisandra de Castro Bras, East Carolina University, North Carolina, United States
Suneel Apte, Cleveland Clinic Lerner Institute

3:45 PM – 5:15 PM	**EVENING WORKSHOP** **How to Bring the Impact of Multi-omics Research to the Next Level? From Industry and Academic Perspectives**	**Ballroom 2**

Presented By: J. Wade Davis, AbbVie, Inc., Illinois, United States
Yu Tian, AbbVie, Inc., Illinois, United States
Christopher E. Mason, Weill Cornell Medicine, New York, United States
Thomas Conrads, Inova Health System, Virginia, United States

5:30 PM – 7:30 PM	**SOCIAL ACTIVITY** **US HUPO ECR and FeMS**	**Millennium Park**

Early Career trainees are invited to join this social happy hour, co-hosted by FeMS and US HUPO ECR Committee.

THURSDAY, MARCH 11

(all times in Central Time)

9:00 am – 9:45 am	**Sponsored Seminar**	**Ballroom 2**

Sponsored By:

cellenion
A CELLINK COMPANY

cellenONE for Single Cell Proteomics - Introduction & User Experience

9:00 am – 9:45 am	**Sponsored Seminar**	**Ballroom 1**

Sponsored By:

Olink

Plasma Proteomics and Data Tools for Elucidating Signals in COVID-19 Severity
Cindy Lawley, Olink

10:00 AM – 10:50 AM	**Award Plenary Session**	**Auditorium**

Session Chair: Birgit Schilling, The Buck Institute for Research on Aging, Novato, CA, United States

Robert J. Cotter New Investigator Award Presentation

Martin Wuhr, Princeton University, New Jersey, United States

11:00 AM – 12:30 PM	**PARALLEL SESSION** **Protein Post-Translational Modifications**	**Ballroom 2**

Session Chairs: Hui Zhang, John Hopkins University, Baltimore, MD and Justyna Fert-Bobber, Cedars Sinai Medical Center, Los Angeles, CA

11:00 AM – 11:30 AM

Functional Proteomics Analysis to Discover Novel Oxygen-sensing Posttranslational Modification Pathway
Yue Chen, Associate Professor of Biochemistry, University of Minnesota, Minnesota, United States

11:30 AM – 12:00 PM

Methods to Illuminate the Regulation and Function of the Phosphoproteome
Judit Villen, University of Washington, Washington, United States

12:00 PM – 12:15 PM	**09.01: Quantitative PTM Maps of Human Pathologic Tau Identify Patient Heterogeneity and Define Critical Steps in Alzheimer's Disease Progression** Mukesh Kumar, Hendrik Wesseling, Waltraud Mair, Christoph Schlaffner, Pieter Beerepoot, Benoit Fatou, Amanda Guise, Long Cheng, Hanno Steen and Judith Steen
12:15 PM – 12:30 PM	**09.02: Proteogenomic Characterization of Pancreatic Ductal Adenocarcinoma** Liwei Cao, Chen Huang, Daniel Cui Zhou, Oliver F. Bathe, Daniel W. Chan, Ralph H. Hruban, Li Ding, Bing Zhang and Hui Zhang

11:00 AM – 12:30 PM — PARALLEL SESSION — Ballroom 1

Top-Down Proteomics and Native MS

Session Chairs: Jeff Agar, Northeastern University, Boston, MA and Huilin Li, Sun Yat-Sen University Guangzhou, China

11:00 AM – 11:30 AM	**Building a Community of Top Down Proteomics to Advance Proteoform Measurement & Biology** Neil Kelleher, Northwestern University, Illinois, United States
11:30 AM – 12:00 PM	**Towards Illuminating the Dark Matter of Mass Spectrometry for a Top-Down View of Native Protein Structures** Joseph Loo, University of California Los Angeles, California, United States
12:00 PM – 12:15 PM	**O10.01: Novel Strategies for Top-down Proteomics of Endogenous Membrane Proteins** Kyle Brown, Trisha Tucholski, Christian Eken and Ying Ge
12:15 PM – 12:30 PM	**O10.02: Automated capillary isoelectric focusing-tandem mass spectrometry for qualitative and quantitative top-down proteomics** Tian Xu and Liangliang Sun

12:30 PM – 1:00 PM — Networking Break — Exhibit Hall

Visit the virtual exhibit hall to connect with our exhibitors.

THURSDAY, MARCH 11

(all times in Central Time)

1:00 PM – 1:45 PM | **Sponsored Seminar** | **Auditorium**

Sponsored By:

Proteograph Product Suite: An Automated, Nanoparticle-Based Solution for Unbiased, Deep, Rapid, and Scalable Plasma Proteomics
Daniel Hornburg, Seer Inc., California, United States

2:00 PM – 3:30 PM | **PARALLEL SESSION**
Machine Learning and Computational Proteomics | **Ballroom 2**

Session Chairs: Gil Omenn, University of Michigan, Ann Arbor, MI and David Fenyo, New York University, New York, NY

2:00 PM – 2:30 PM
Applications of Intelligent Data Acquisition for Multiplexed Proteomics
Devin Schweppe, University of Washington, Washington, United States

2:30 PM – 3:00 PM
Advancing Precision Oncology Through Proteogenomics: Challenges and Opportunities
Bing Zhang, Baylor University, Texas, United States

3:00 PM – 3:15 PM
11.01: Deep learning-derived evaluation metrics for benchmarking computational pipelines for the analysis of phosphoproteomic data
Wen Jiang, Bo Wen, Kai Li, Felipe Leprevost, Jamie Moon, Nathan Edwards, Tao Liu, Alexey Nesvizhskii and Bing Zhang

3:15 PM – 3:30 PM
11.02: hu.MAP 2.0: Integration of Over 15,000 Proteomic Experiments Builds a Global Compendium of Human Multiprotein Assemblies
Kevin Drew, John Wallingford and Edward Marcotte

2:00 PM – 3:30 PM | **PARALLEL SESSION**
Multi-'omics - Methods and Applications | **Ballroom 1**

Session Chairs: Robert Moritz, Institute for Systems Biology, Seattle, WA and Pei Wang, Mount Sinai Hospital, New York, NY

2:00 PM – 2:30 PM	**Proteome Activity Landscapes of Tumor Cell Lines Determine Drug Responses** Bernhard Kuster, Technical University of Munich, Munich, Germany
2:30 PM – 3:00 PM	**Linking the Genome and the Proteome in Cancer** Sam Payne, Brigham Young University, Utah, United States
3:00 PM – 3:15 PM	**O12.01: Dynamic 3D proteomes reveal protein functional alterations at high resolution in situ** Valentina Cappelletti, Hauser Thomas, Ilaria Piazza, Monika Pepelnjak, Liliana Malinovska, Tobias Fuhrer, Yaozong Li, Christian Dörig, Paul Boersema, Ludovic Gillet, Jan Grossbach, Aurelien Dugourd, Julio Saez-Rodriguez, Andreas Beyer, Nicola Zamboni, Amedeo Caflisch, Natalie de Souza, Paola Picotti
3:15 PM – 3:30 PM	**12.02: A New Deconvolution Method for Studying Cell Type Composition in Tumor Microenvironment based on Bulk Proteogenomic Data** Francesca Petralia and Pei Wang

---

3:45 PM – 5:15 PM — **EVENING WORKSHOP** — **Navy Pier**

Targeted Mass Spectrometric Assays for Diabetes and Obesity Research

Presented By: Salvatore Sechi, National Institute of Health / National Institute of Diabetes and Digestive and Kidney Diseases, Maryland, United States
Jennifer Van Eyk, Cedars-Sinai Medical Center, California, United States
Andy Hoofnagle, University of Washington, Washington, United States
Jun Qu, University at Buffalo, New York, United States
Wei-Jun Qian, Pacific Northwest National Laboratory, Washington, United States

---

3:45 PM – 5:15 PM — **EVENING WORKSHOP** — **Auditorium**

Putting Humpty Dumpty Back Together Again: What Does Protein Quant Mean in Bottom-Up Proteomics?

Presented By: Alexey Nesvizhskii, University of Michigan, Michigan, United States
Deanna Plubell, University of Washington, Washington, United States
Lloyd Smith, University of Wisconsin - Madison, Wisconsin, United States
Neil Kelleher, Northwestern University, Illinois, United States
Olga Vitek, Northeastern University, Massachusetts, United States
Nicolas Young, Baylor College of Medicine, Texas, United States
Bobbie-Jo Webb-Robertson, Pacific Northwest National Laboratory, Washington, United States

5:30 PM – 7:30 PM	**Virtual Cocktail Party**	Magnificent Mile Foyer

Grab a snack and your favorite beverage and join your fellow attendees in the Gather.Town foyer to catch up before the conference ends.

9:00 AM – 9:45 AM — **Morning Activities**

Morning Yoga
Navy Pier

Join us in the Navy Pier room for a beginner yoga session to gear you up for a day full of sessions! No experience required.

Coffee Chat
Millennium Park

Not a morning person? Bring your own beverage and join your fellow attendees in the Millennium Park room to catch up over coffee.

10:00 AM – 10:50 AM — **Award Plenary Session** — **Auditorium**

Session Chair: Gilbert Omenn, University of Michigan, Ann Arbor, MI, United States

Gilbert S. Omenn Computational Proteomics Award Presentations

Nuno Bandeira, University of California San Diego, California, United States

Olga Vitek, Northeastern University, Massachusetts, United States

11:00 AM – 12:30 PM — **PARALLEL SESSION**
New Strategies of Interactomics in Structural Biology — **Ballroom 2**

Session Chairs: Kallol Gupta, Yale University, New Haven, CT and Michael Marty, University of Arizona, Tucson, AZ

11:00 AM – 11:30 AM
Pharmacological Insight Through Interactome Studies
Jim Bruce, University of Washington, Washington, United States

11:30 AM – 12:00 PM
Developing Cross-linking Mass Spectrometry for Structural Analysis of Protein Complexes
Lan Huang, University of California Irvine, California, United States

12:00 PM – 12:15 PM
13.01: Cleavable Biotin-Enabled Proximity Labeling Proteomics and Peptidomics
Shiori Sekine, Haorong Li, Ashley Frankenfield and Ling Hao

12:15 PM – 12:30 PM
13.02: Identification of ligand-dependent protein interaction networks of the mu-opioid receptor with temporal and spatial resolution
Ruth Hüttenhain, Braden T Lobingier, Benjamin J Polacco, Qiongyu Li, Emily Blythe, Nevan J Krogan and Mark Von Zastrow

11:00 AM – 12:30 PM	**PARALLEL SESSION** **Stem Cell Proteomics and Chemical Proteomics in Drug Discovery**	**Ballroom 1**

Session Chairs: Stacy Malaker, Yale University, New Haven, CT and Jon D. Williams, Abbvie, Inc., North Chicago, IL

11:00 AM – 11:30 AM
Tissue-Down Cell Surface Proteomics for Benchmarking Stem Cell-Derived Cardiomyocytes
Rebekah Gundry, University of Nebraska Medical Center, Nebraska, United States

11:30 AM – 12:00 PM
Activity Mapping from Single Cells to Whole Organisms in Cancer
Ray Moellering, University of Chicago, Illinois, United States

12:00 PM – 12:15 PM
14.01: High-throughput elucidation method for small molecule - protein interactions
Ziwei Zhang, Xiuyuan Ma and Yu Gao

12:15 PM – 12:30 PM
14.02: Quantitative mass spectrometry-based global proteome and phosphoproteome analyses of thymic epithelial tumors (TET) reveal novel drug targets
Xu Zhang, Ting Huang, Fatos Kirkali, Yue Qi, Tapan Maity, Khoa Dang Nguyen, David Schrump, Olga Vitek, Arun Rajan and Udayan Guha

12:30 PM – 1:00 PM	**Networking Break**	**Exhibit Hall**

Visit the virtual exhibit hall to connect with our exhibitors.

1:00 PM – 2:00 PM	**Closing Award Plenary Session**	**Auditorium**

Session Chair: Ileana Cristea, Princeton University, New Jersey, United States

Catherine E. Costello Lifetime Achievement in Proteomics Award Presentation

Brian Chait, The Rockefeller University, New York, United States

3:00 PM – 5:00 PM	**2021 Board of Directors Meeting**	**By Invitation Only (Zoom)**

US HUPO
from genes to function

2021 VIRTUAL

ASSOCIATION AWARDS

Donald F. Hunt Distinguished Contribution in Proteomics Award

The Donald F. Hunt Distinguished Contribution in Proteomics award recognizes a focused or singular achievement in the field of proteomics. It is fully supported by the Journal of Proteome Research (JPR) and was established to recognize Prof. Hunt's significant contributions to the field of proteomics.

ACS Publications

Peipei Ping, University of California, Los Angeles

Dr. Ping is Professor of Physiology, Medicine, and Biomedical Informatics at the UCLA David Geffen School of Medicine. Dr. Ping is internationally recognized for her expertise in mitochondrial biology and proteome remodeling in cardiovascular diseases, application of data science in molecular phenotyping and diseases, as well as computational analytical platforms in cardiovascular diseases. She currently is the NHLBI T32 Program Director of Integrative Data Science Training in Cardiovascular Medicine at UCLA and she also serves as the Associate Director of Scalable Analytics Institute (ScAI) at Department of Computer Science in UCLA Samueli School of Engineering. From 2014 to 2019, Dr. Ping served as Program Director for NIH BD2K Centers of Excellence at UCLA (HeartBD2K). She has authored over 200 publications. She has been a frequent speaker at national and international conferences; she has delivered more than 300 invited lectures since 1996. Professor Ping has devoted her career to education (84 trainees, all are making important contributions to the society: 17 hold positions in academic institutions globally and 43 are employed by leading technology firms and/or Fortune 500 Companies).

Dr. Ping has received many honors; including a Research Achievement Award from the Canadian Proteomics Network (2009), the MERIT Award from NHLBI/NIH (2010), the Thomas W Smith Lectureship from the American Heart Association (AHA, 2012), the Distinguished Service Award from International Human Proteome Organization (Intl HUPO, 2013), the Robert M Berne Distinguished Lectureship in Cardiovascular Medicine from the American Physiological Society (APS, 2015), the Outstanding Investigator Award from NHLBI/NIH (2017), and the Clinical Translational Award (Intl HUPO, 2018).

Robert J. Cotter New Investigator Award

This award was established to honor the memory of Bob Cotter, a founding member of US HUPO, for his many contributions to scientific research and for his legacy as a mentor to young scientists. Each year, the award will be given to an individual early in his or her career, in recognition of significant achievements in proteomics, broadly defined.

Martin Wuhr, Princeton University

Martin Wühr received his BS in biochemistry from the TU Munich, and a PhD in systems biology from Harvard. For his postdoc, Martin worked jointly with Steven Gygi and Marc Kirschner. Since 2016 Martin has been an Assistant Professor of Molecular Biology at the Lewis-Sigler Institute at Princeton University. The Wühr Lab aims to advance quantitative proteomics via the complementary reporter ion approach. The developed technology enables accurate, sensitive, and precise multiplexed proteomics at the MS2 level with comparatively simple instrumentation. Leveraging this approach, the Wühr Lab studies various aspects of cellular organization, including how differential nucleocytoplasmic partitioning affects developmental progression, which proteins are liquid-liquid phase separated in vivo, and to what extent protein expression is regulated via transcription, translation, or degradation.

Gilbert S. Omenn Computational Proteomics Award

This award recognizes the essential nature of computational methodology and software in proteomics. Specifically, this award acknowledges the specific achievements of scientists that have developed bioinformatics, computational, statistical methods and/or software used by the proteomics community, broadly defined. The award is named in honor of Gil Omenn, a US HUPO Past President, leader of the Human Proteome Project, and influential proteomics researcher.

Nuno Bandeira, University of California, San Diego

Dr. Nuno Bandeira obtained his Ph.D. in Computational Mass Spectrometry in 2007 at the Department of Computer Science, University of California, San Diego. He is currently an Associate Professor with joint appointments with the Dept. of Computer Science and Engineering and with the Skaggs School of Pharmacy and Pharmaceutical Sciences at the University of California, San Diego, where he is also a founding faculty of the Halicioglu Data Science Institute (HDSI). Dr. Bandeira is also the Executive Director of the Center for Computational Mass Spectrometry (CCMS), where he develops algorithms for large scale analysis of proteomics mass spectrometry data. Dr. Bandeira leads the MassIVE repository for proteomics mass spectrometry data and the GNPS repository and knowledge base for metabolomics and natural products mass spectrometry data, altogether serving the community with hundreds of terabytes of mass spectrometry data in over 10,000 public datasets. Building on CCMS's expertise with distributed algorithms for analysis of mass spectrometry data, MassIVE and GNPS have enabled the concept of mass spectrometry `living data', whereby public datasets are continuously reprocessed to transfer new knowledge to existing data, as well as to generate new hypotheses that are curated by active researchers to generate new knowledge (which then loops back to all other datasets). This virtuous cycle of iterative reanalysis, curation and knowledge base extension has resulted in the data-networking of thousands of researchers in 150+ countries, where it was integrated with numerous ongoing research projects and consequently improved public data annotation by over 10-fold since the time of original deposition.

Olga Vitek, Northeastern University

Dr. Olga Vitek holds a PhD in Statistics from Purdue University. After a postdoc in the Aebersold lab in the Institute for Systems Biology in Seattle, she returned to the Department of Statistics at Purdue as Assistant and then Associate Professor and University Faculty Scholar. Currently Olga is Professor in the Khoury College of Computer Sciences at Northeastern University. Her group develops, implements and evaluates methods at the intersection of statistics, machine learning, mass spectrometry and systems biology. Open-source software MSstats and Cardinal developed in her lab are used in academia and industry, and were recently recognized with the Chan Zuckerberg Essential Open Source Software for Science award. Olga's lab hosts May Institute on Computation and Statistics for Mass spectrometry and Proteomics, an educational event that each year attracts over 100 participants. Olga is a Senior Member of the International Society for Computational Biology, and an Elected Member of the Board of Directors of USHUPO and of the Council of HUPO. She serves as Associate Editor of Bioinformatics, and a member of the Editorial advisory board of Molecular and Cellular Proteomics and of Journal of Proteome Research.

US HUPO
from genes to function

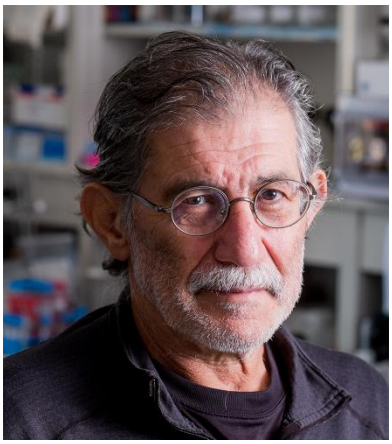
2021 VIRTUAL

ASSOCIATION AWARDS

Catherine E. Costello Lifetime Achievement in Proteomics Award

The Lifetime Achievement in Proteomics Award recognizes a career of discovery that has made a lasting impact on the field of proteomics.

Brian Chait, The Rockefeller University



Brian T. Chait was born in Cape Town, South Africa. He received his B. Sc. (1969) and B. Sc. (Hons) (1970) from the University of Cape Town and D.Phil. (1976) in experimental nuclear physics from Oxford University. For the past 42 years he has worked at The Rockefeller University in New York, where he is Camille and Henry Dreyfus Professor and Head of the Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, and where he led until recently the NIH-funded National Resource for the Mass Spectrometric Analysis of Biological Macromolecules. Professor Chait has received several awards for his group's research in developing instrumentation and methods for characterizing proteins, including the 2002 ACS Field & Franklin Award for Outstanding Achievement in Mass Spectrometry, the 2007 HUPO Distinguished Discovery Award in Proteomics, and the 2015 ASMS Award for a Distinguished Contribution in Mass Spectrometry. He and his group have also been awarded 31 US patents, several of which have been commercialized and are widely used throughout the world.

Interesting is submitting a nomination for a US HUPO 2022 Association Award?

The Call for Nominations will open this fall!

Self-nominations are also encouraged.
Nominations will be held for 3 years.

Parallel Session 01: Single Cell Proteomics

Single-Cell Proteomics: Practical Considerations for Unlocking the Proteome One Cell at a Time

The ability to profile single cells and map the proteome with high spatial resolution across tissue regions provides a new way to understand the tissue microenvironment, substructure, and cellular organization from a global proteome perspective. Recent advances in sample processing, separations, ion mobility and mass spectrometry that have enabled profiling of >1000 proteins from single mammalian cells will be discussed. Efforts to simplify and automate the workflow for improved robustness and broader dissemination will also be described.

Presented By:

Ryan Kelly, Brigham Young University, Utah, United States

Mass Spectrometry Gone Subcellular to Drive Developmental Neurobiology

Understanding development requires knowledge of both the types of molecules that are produced in cells and the functional roles they play. During early patterning of the chordate embryo, including formation of the neural tissues, cells must differentiate to the correct cell type at the correct location and correct time. In this presentation, we will discuss the development of in situ/vivo approaches by subcellular mass spectrometry that enabled us to characterize the proteomic and metabolomic profile of spatiotemporally identified single cells in live embryos developing to tadpoles. Using precision fabricated microcapillaries, the intracellular content was collected from the cells that have reproducible tissue fates in cleavage-stage embryos of Xenopus laevis. Molecular measurements on capillary electrophoresis TOF or Orbitrap mass spectrometer platforms revealed proteomic and metabolomic differences between the cells that correlated with cell type. Such molecular differences were not detectable at the level of the transcriptome in independent studies. Follow-up experiments discovered the formation of cell-by-cell molecular gradients that actively contribute to patterning of the embryo. By downscaling the MS approach to <5% of the cell's volume, 95% of the analyzed embryos successfully developed to sentient tadpoles. Based on anatomy and visual function in a background color preference assay, these tadpoles were indistinguishable from the control siblings, validating in vivo single-cell MS. In situ/vivo subcellular MS expands the molecular toolbox of cell biology to help derive and test hypotheses on the functional roles proteins, peptides, and metabolites play during normal and impaired development.

Presented By:

Peter Nemes, University of Maryland, Maryland, United States

O01.01 | Principled and Reproducible Mass Spectrometry-based Single Cell Proteomics Data Analysis

Laurent Gatto (1), Christophe Vanderaa (1)

(1) de Duve Institute, UCLouvain, Brussels, Belgium

Mass spectrometry-based proteomics is actively embracing quantitative, single cell-level analyses. Indeed, recent advances in sample preparation and mass spectrometry (MS) have enabled the emergence of quantitative MS-based single-cell proteomics (SCP). While exciting and promising, SCP

still has many rough edges. This work describes our computational contribution to SCP. We present a new framework that facilitates the reproducible processing, analysis and exploration of SCP data. We showcase the application of 'scp', a new R/Bioconductor software package, by replicating the analysis of two publicly available SCP data sets: a nanoPOTS label-free quantitation experiment (Zhu et al. 2019) and a TMT-based SCoPE2 data set (Specht et al. 2020). 'scp' is a peer-reviewed, well documented and thoroughly tested software. It provides us with the means to formalise and standardise SCP data processing, analyses and interpretation. The availability of software such as 'scp' as well as pre-packaged and curated data such as those available in the 'scpdata' companion package play an important role in supporting and promoting the emergence of SCP. Well crafted software enable researchers to optimise SCP data analyses, enable reproducible data processing and provide the ground for rigorous development and benchmarking of new state-of-the-art methods.

Contact: Laurent Gatto, laurent.gatto@uclouvain.be

O01.02 | Robust, very low nanoflow LC-MS methods for ultra-sensitive single-cell and phosphoproteomics

Dorte B. Bekker-Jensen (1), Ole B. Hørning (1), Andreas-David Brunner (2), Marvin Thielert (2), Jesper V. Olsen (3), Ole Vorm (1), Matthias Mann (2), Nicolai Bache (1)

(1) Evosep Biosystems, Odense, Denmark

(2) Max Planck Institute for Biochemistry, Martinsried, Germany

(3) Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark

Ultra-sensitive proteomics is a rapidly growing field where cutting-edge technologies are pushed to their limits for applications such as single-cell analyses. One of the main challenges to achieve this, is to maintain sensitivity and robustness throughout the entire workflow, where attention is especially needed during sample preparation and efficient loading of minute sample amounts to the LC-MS system. Here we describe how the Evotip, a disposable trap column, is an integral part for efficient concentration and loading of very low sample amounts. As the electrospray ionization efficiency is concentration dependent, we achieve the highest sensitivity at low flowrates. To take advantage of this, we developed a set of LC methods operating at gradient flows of 25, 50 and 100 nl/min on standard Evosep One hardware. This setup builds on a robust platform allowing for several thousand uninterrupted injections even at very low flow rates.

To verify the expected increase in sensitivity with the nanoflow methods, we performed dilution series of HeLa digests, analyzed the samples with the 30 min method operating at a flow rate of 100 nl/min and benchmarked its performance directly against the established 60 samples per day method, with a gradient flow of 1 ul/min. We found that at sub-ng of HeLa peptides, the nanoflow method increased up to ten-fold compared to the standard method, resulting in more than 50% increase in protein identifications. Interestingly, these proteins were also measured with significantly higher run-to-run reproducibility.

We also challenged our method by analyzing phosphopeptides enriched from just 1ug of peptide starting material. Again, we found a factor of 10x higher average intensity compared to the standard method. In conclusion, the combination of using Evotips and preformed gradients allow operation at extremely low flow rates with undiminished robustness, opening exciting application areas.

Contact: Dorte B. Bekker-Jensen, dbj@evosep.com

Parallel Session 02: Cellular Signaling & Systems Biology

Control of Nuclear Integrity via Lamin Acetylation During Cell Cycle Progression and Viral Infections

Nuclear lamins are ancient type V intermediate filaments with diverse functions that include maintaining nuclear shape, tethering and stabilizing chromatin, regulating gene expression, and contributing to cell cycle progression. The dysregulation of the nuclear lamina has been linked to laminopathy diseases and cancers, as well as to viral infections. A mandatory step for nuclear-replicating viruses is their need to disrupt the nuclear periphery for virus capsid egress and further maturation of infectious particles. Given these critical roles of lamins in health and disease states, an outstanding question has been how lamins are regulated. Here, we establish the first known function for lamin B1 (LMNB1) acetylation in normal proliferating cells, as well as during viral infections. Using virology assays in conjunction with acetylome, spatial proteomics, thermal proximity co-aggregation, and microscopy, we discover that LMNB1 K134 acetylation inhibits nuclear-replicating viruses (human cytomegalovirus and herpes simplex virus 1) by maintaining the integrity of the nuclear periphery. We demonstrate the broader relevance of this acetylation outside the context of infection by uncovering a role for LMNB1 acetylation in regulating cell cycle progression and DNA repair in uninfected cells. We determine that LMNB1 K134 acetylation inhibits the progression of cells through the G1/S cell cycle checkpoint. Using mutational analysis and TurboID mass spectrometry, we characterize the effect of K134 acetylation on the LMNB1 interaction landscape upon DNA damage. This analysis underscores that LMNB1 K134 acetylation regulates the stability of the nuclear periphery. We define the underlying mechanism, demonstrating that LMNB1 K134 acetylation blocks cell cycle progression by impeding 53BP1 recruitment to DNA damage foci and delaying DNA repair through canonical nonhomologous end joining. Altogether, our findings highlight an important regulatory hub for nuclear shape and function, adding to the growing knowledge of mechanisms that drive organelle remodeling and cellular organization in health and disease states.

Presented By:

Ileana Cristea, Princeton University, New Jersey, United States

---

Proteomics-Based Systems Biology of Human Disease

Our mission is to develop novel mass spectrometry-based proteomics and systems biology tools and apply these tools in the fields of Alzheimer's disease and cancer. We seek to obtain the full spectra of multi-omics data from cellular/animal/clinical samples. Integration of such multi-omics data offer a systems or holistic view for unbiased identification of central disease signaling networks, functional modules and master regulators. To validate the derived hypotheses from the big data analysis, we develop animal models and perturb the models with chemical and genetic methods in functional and mechanistic studies. These studies provide novel insights into the pathogenesis for therapeutic intervention, and may discover disease biomarkers for precision medicine.

Presented By:

Junming Peng, St. Jude's Hospital, Tennessee, United States

O02.01 | Phosphorylation-driven Regulation of IFI16 Viral DNA Sensing at the Nuclear Periphery

Dawei Liu (1), Krystal Lum (1), Joshua Justice (1), Ileana Cristea (1)

(1) Molecular Biology, Princeton University, Princeton, New Jersey, United States

In mammalian cells, DNA sensors are critical players in innate immunity, serving to detect the presence of pathogenic DNA and activate innate immune responses. The interferon inducible protein 16 (IFI16) has emerged as a key sensor of foreign DNA in the nucleus. Its function is especially critical in the defense against nuclear-replicating viruses, such as herpes simplex virus 1 (HSV-1). During viral infection, IFI16 quickly localizes to the nuclear periphery, where it binds to incoming viral DNA and oligomerizes. Upon viral DNA binding, IFI16 inhibits virus replication by inducing cytokine expression and inhibiting viral gene expressions. However, little is known about how IFI16 antiviral functions are regulated at the nuclear periphery and the mechanisms through which IFI16 transmits immune signals or whether this relies on nucleo-cytoplasmic shuttling of host factors for communication to the ER-localized STING adaptor protein. To answer these questions, we used a multidisciplinary approach to characterize localization-dependent protein interactions and posttranslational modification of IFI16 at the nuclear periphery. Using biochemical fractionation and quantitative mass spectrometry, we defined the dynamic proteomes of the nuclear periphery compared to the nuclear core in uninfected and infected primary fibroblasts. We then designed a proximity labeling-mass spectrometry (miniTurboID) workflow to capture transient and stable interactions of IFI16 at the nuclear periphery. To gain insights into factors contributing to downstream signaling, we next investigated nucleo-cytoplasmic shuttling early in infection. We infected cells with wild type HSV-1 or a strain that lacks the ability to inhibit IFI16, and performed nuclear-cytoplasmic fractionation and tandem mass tagging (TMT)-MS to predict alterations in translocation events. Cross-referencing the nuclear periphery interactions and proteome datasets, we identified dynamic IFI16 associations with kinases at the nuclear periphery. Through targeted mass spectrometry, mutagenesis, microscopy and molecular virology, we identified key phosphorylation sites on IFI16 that modulate its antiviral functions.

Contact: Dawei Liu, daweiliu@princeton.edu

O02.02 | Protein Communities Reveal New Members When Recognizing the Deeper Coding Potential of Transcripts

Sebastien Leblanc (1), Marie A Brunet (1), Jean-François Jacques (1), Amina M Lekehal (1), Alan A Cohen (2), Michelle S Scott (1), Xavier Roucou (1)

(1) Biochimie et genomique fonctionelle, Universite de Sherbooke, Sherbrooke, Canada

(2) Département de médecine de famille, Universite de Sherbooke, Sherbrooke, Canada

Recent proteogenomic approaches have led to the discovery that regions of the transcriptome previously annotated as non-coding frequently encode proteins (termed alternative proteins, or altProt). These regions correspond to non-coding RNAs or untranslated regions as well as overlapping reading frames in mRNAs. Thus a growing body of literature now shows that human genes have the capacity to encode more than one functional protein. AltProts are missing from common protein libraries used in the analysis of mass spectrometry data which precludes their detection and the discovery of their functional roles. Here we incorporate this increased diversity in a stringent re-analysis of a high throughput human network proteomics dataset (BioPlex 2.0) which includes a peptide centric approach to the validation of peptide spectrum matching. Over 200 altProts were revealed throughout the network and found within clusters associated with a wide variety of cellular functions and pathologies. AltProts integrated the large network seamlessly, occupying a wide range of eigenvector centrality with some bridging otherwise sparsely interconnected regions. This work at once confidently identifies novel human proteins and allows the generation of plausible hypotheses for their functional role in cellular processes. Interesting patterns of interactions involving altProts emerge. Notably, multiple genes were found to encode both an annotated (reference) protein and an altProt with some found close together and others only distantly connected through the network. Additionally, of the altProts encoded by pseudogenes, 28% were found to be direct interactors of reference proteins encoded by their respective parental gene. Our study unveils the contribution of altProts to protein complexes and offers fresh avenues for shedding light onto molecular pathways related to human diseases.

Contact: Sebastien Leblanc, lebs2209@usherbrooke.ca

Parallel Session 03: Omics of the Central Nervous System

Multi-dimensional 'Omics of Neurodegenerative Diseases

Presented By:

Michael Ward, National Institute of Health, Maryland, United States

---

Comprehensive Proteomics and Lipidomics Strategies to Advance Disparities Research in Alzheimer's Disease

Alzheimer's disease will impact an estimated 15 million individuals by 2050 in the United States, and the global incidence is increasingly growing. Certain racial/ethnic subpopulations are disproportionately affected by Alzheimer's disease. For example, African Americans and Hispanics have one and a half to three times higher incidence rates compared to non-Hispanic Whites. African Americans also have higher incidence of hypertension, diabetes, high cholesterol, obesity, and cardiovascular disease in adults. These comorbidities increase risk for Alzheimer's disease and currently a molecular understanding of how this occurs does not exist [1]. We hypothesize that altered lipid metabolism is a major and central contributor to Alzheimer's disease and these noted comorbidities. Recently, we have begun to explore this hypothesis using comprehensive proteomics and lipidomics mass spectrometry-based approaches. Specifically, we have developed robust quantitative assays to measure proteins in post-mortem brain, liver, and heart tissue as well as in plasma and a range of lipids and lipid classes in plasma. Combined, these analyses provide tremendous insight into lipid metabolism and many other biological pathways that are altered in the brain and periphery of Alzheimer's disease. Furthermore, protein and lipid changes that are altered in ways that are unique to a given racial/ethnic background have been assessed. This presentation will discuss the various proteomic and lipidomics approaches established, results from pilots with small cohorts of Alzheimer's disease patients, and the implications of these findings for understanding disease pathogenesis especially as it relates to racial/ethnic disparities in Alzheimer's disease.

Presented By:

Rena Robinson, Vanderbilt University, Tennessee, United States

---

O03.01 | Deconvoluting the Effects of Age and Disease on Protein Aggregation in Mouse Models

Cristen Molzahn (1), Erich Kuechler (2), Irina Zemlyankina (1), Lorenz Nierves (3), Aly Karsan (4), Philipp Lange (3), Thibault Mayor (1)

(1) Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC, Canada

(2) Michael Smith Laboratories, University of British Columbia, Vancouver, BC, Canada

(3) BC Children's Hospital Research Institute, University of British Columbia, Vancouver, BC, Canada

(4) BC Cancer Agency, University of British Columbia, Vancouver, BC, Canada

Normal cellular function relies on maintaining the proteome in an operating state. As an organism ages, cellular response to unfolded proteins declines which results in diminished protein homeostasis (proteostasis). The reduced folding capacity is associated with the onset of aggregation of numerous neurodegenerative disease-related proteins. In C. elegans, the proteostatic decline with age has even been linked to the onset of aggregation of proteins in wild-type animals. However, this has not been extensively characterized in aging mammals. Here we present data on triton-insoluble proteins obtained from wild type mice to identify features associated to proteins that aggregate upon aging. We then performed a meta-analysis to

compare these features among distinct groups of insoluble proteins reported in the literature. To reveal the insoluble portion of the proteome, we analyzed the pellet fraction of mouse brain tissues after high-speed centrifugation using data independent acquisition. We quantified over 4000 proteins in the pellet fraction derived from cortex tissue from mice ages 4- and 23-months revealing age-induced changes in protein solubility. Notably, proteins enriched in the pellet fraction consist of long-lived proteins and proteins associated with oxidative stress. We then compared these proteins with four published proteomic datasets of insoluble proteins from 3 different disease models. Surprisingly, feature analysis of these dataset revealed that the age of the mouse was associated with different types of proteins entering the pellet fraction. In general, insoluble proteins from young models were more disordered and more likely to be found in phase separated granules. With age, these traits are less enriched and proteins with more structure and longer half-lives enter the pellet fraction. Together this data suggests that upon aging the proteome organization is altered leading to a different class of proteins that enters the pellet fraction as a result of loss of proteostasis.
Contact: Cristen Molzahn, cristen.molzahn@msl.ubc.ca

O03.02 | A dynamic mouse peptidome landscape reveals probiotic modulation of the gut-brain axis

Chenxi Jia (1)

(1) National Center for Protein Sciences-Beijing, Beijing, China

Certain probiotics have beneficial effects on the function of the central nervous system through modulation of the gut-brain axis. Here, we describe a dynamic landscape of the peptidome across multiple brain regions, modulated by oral administration of different probiotic species over various time courses. The spatiotemporal and strain-specific changes of the brain peptidome correlated with the composition of the gut microbiome. The hippocampus exhibited the most sensitive response to probiotic treatment. The administration of heat-killed probiotics altered the hippocampus peptidome, but did not significantly change the gut microbiome. We developed a literature-mining algorithm to link the neuropeptides altered by probiotics with potential functional roles. We validated the probiotics-regulated role of corticotropin-releasing hormone (CRH) by monitoring the hypothalamic-pituitary-adrenal axis, the prenatal stress–induced hyperactivity of which was attenuated by probiotics treatment. Our findings provide evidence for modulation of the brain peptidome by probiotics and provides a rich resource for further studies of the gut-brain axis and probiotic therapies.

Contact: Chenxi Jia, cjia99@126.com

Parallel Session 04: Proteomics of Infectious Disease

Host Factor Targeted Drug Discovery for SARS-CoV-2 Through an International Collaboration

The novel coronavirus SARS-CoV-2, the causative agent of COVID-19 respiratory disease is evolving during the current pandemic. New variants show enhanced replication and the potential to evade therapeutic antibodies. In the near future, variants may even evade first generation vaccines. The currently approved direct acting antiviral remdesivir targets the viral RNA-dependent RNA polymerase which is subject to rapid evolution as it is encoded by the viral RNA genome. In order to develop therapeutic approaches which act in a pan-coronavirus manner we and our colleagues at the QBI Coronavirus Research Group (QCRG) have mapped the human proteins (host factors) which multiple Coronaviruses rely on for replication. Through a rapid drug repurposing effort we have identified zotatifin, a clinical eIF4A inhibitor as a host factor targeted therapeutic. Zotatifin which is based on the natural product rocaglamide A works as a molecular glue to trap eIF4A on its target, the (+) RNA viral genome. Other examples of targeting essential host factors, including those for immune evasion will be discussed.

Presented By:

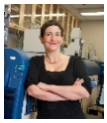
Nevan Krogan, University of California San Francisco, California, United States

---

A Neighboring View of the SARS-CoV-2 Proximal Interactome

Viral replication is dependent on interactions between viral polypeptides and host proteins. Identifying virus-host protein interactions can thus uncover unique opportunities for interfering with the virus life cycle via novel drug compounds or drug repurposing. Importantly, many viral-host protein interactions take place at intracellular membranes and poorly soluble organelles, which are difficult to profile using classical biochemical purification approaches. We used proximity-dependent biotinylation (BioID), to establish a reference map for the human cell at steady state (Go et al., bioRxiv; humancellmap.org). Applying proximity-dependent biotinylation (BioID) with the fast-acting miniTurbo enzyme to selected humancellmap baits and to 27 SARS-CoV-2 proteins in a lung adenocarcinoma cell line (A549), we detected 7810 proximity interactions (7382 of which are new for SARS-CoV-2) with 2242 host proteins (results available at covid19interactome.org). These results complement and dramatically expand upon recent affinity purification-based studies identifying stable host-virus protein complexes, and offer an unparalleled view of membrane-associated processes critical for viral production. The parallel profiling of host organellar markers allowed us to propose modes of action for several viral proteins in the context of host proteome remodelling. In summary, our dataset identifies numerous high confidence proximity partners for SARS-CoV-2 viral proteins, and describes potential mechanisms for their effects on specific host cell functions.

Presented By:



Anne-Claude Gingras, University of Toronto, Canada

O04.01 | Versatile and Multivalent Nanobodies Efficiently Neutralize SARS-CoV-2

Yufei Xiang (1), Sham Nambulli (2), Zhengyun Xiao (1), Heng Liu (3), Zhe Sang (4), W. Paul Duprex (2), Dina Schneidman-Duhovny (5), Cheng Zhang (3), Yi Shi (1)

(1) Cell Biology, University of Pittsburgh, Pittsburgh, PA, United States

(2) Center for Vaccine Research, University of Pittsburgh, Pittsburgh, PA, United States

(3) Pharmacology and Chemical Biology, University of Pittsburgh, Pittsburgh, PA, United States

(4) Computational Biology, University of Pittsburgh/Carnegie Mellon University, Pittsburgh, PA, United States

(5) Computer Science and Engineering, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

Cost-effective, efficacious therapeutics are urgently needed against the COVID-19 pandemic. Here, we used camelid immunization and proteomics to identify a large repertoire of highly potent neutralizing nanobodies (Nbs) to the SARS-CoV-2 spike (S) protein receptor-binding domain (RBD). We discovered Nbs with picomolar to femtomolar affinities that inhibit viral infection at sub-ng/ml concentration and determined a structure of one of the most potent in complex with RBD. Structural proteomics and integrative modeling revealed multiple distinct and non-overlapping epitopes and indicated an array of potential neutralization mechanisms. We constructed multivalent Nb constructs that achieved ultrahigh neutralization potency (IC50s as low as 0.058 ng/ml) and may prevent mutational escape. These thermostable Nbs can be rapidly produced in bulk from microbes and resist lyophilization, and aerosolization. These promising agents are readily translated into efficient, cost-effective, and convenient therapeutics to help end this once-in-a-century health crisis.

Contact: Yufei Xiang, yufeixiang@pitt.edu

O04.02 | Defining Mechanisms Underlying Virus Regulation of Mitochondrial Bioenergetics During Infection

Cora Betsinger (1), Joel Federspiel (1), Connor Jankowski (1), William Hofstadter (1), Ileana Cristea (1)

(1) Molecular Biology Department, Princeton University, Princeton, NJ, USA

Alterations of mitochondrial functions and cellular metabolism are hallmarks of nearly all virus infections. As obligate intracellular parasites, viruses rely on mitochondria for the production of biosynthetic precursors and energy necessary for generating new viral particles. The prevalent pathogen human cytomegalovirus (HCMV) alters both mitochondrial structure and metabolism during its replication. However, how HCMV upregulates mitochondrial bioenergetics remains unknown. Employing a multidisciplinary approach integrating proteome and interactome datasets with super-resolution microscopy and metabolomic assays, we identify a previously uncharacterized viral protein, pUL13, which targets the mitochondria and increases oxidative phosphorylation during infection. Specifically, we use targeted mass spectrometry analysis of the HCMV proteome during infection, coupled with molecular virology techniques, to establish that pUL13 is required for productive HCMV replication. We then quantify temporal cellular proteome changes during infection and demonstrate that pUL13 alters electron transport chain protein abundances. Using a live-cell Seahorse metabolic assay to monitor cellular respiration, we establish pUL13 as necessary and sufficient to increase cellular respiration, not requiring the presence of other viral proteins. To mechanistically define the function of pUL13 in regulating cellular respiration, we characterize the spatio-temporal pUL13 functional interaction network during infection. We discover and validate that pUL13 targets the MICOS complex, a critical regulator of mitochondrial architecture and electron transport chain function. We then visualize the impact of pUL13 on mitochondrial architecture using stimulated emission depletion (STED) super-resolution microscopy analysis of mitochondrial cristae ultrastructure. Our findings address the outstanding question of how HCMV modulates mitochondria to increase bioenergetic output and expands the knowledge of the intricate connection between mitochondrial architecture and electron transport chain function. Importantly, this is the first known instance of a virus protein targeting the MICOS complex to increase bioenergetic output, highlighting a mechanism that other virus pathogens might also possess.

Contact: Cora Betsinger, cnbetsinger7@gmail.com

Parallel Session 05: Biomarkers & Precision Medicine

Beyond Genomics; Proteogenomic Insights Into Cancer Biology

Genomic analyses in cancer have been enormously impactful, leading to the identification of driver mutations and development of targeted therapies. But the functions of the vast majority of somatic mutations and copy number variants in tumors remain unknown, and the causes of resistance to targeted therapies and methods to overcome them are poorly defined. Recent improvements in mass spectrometry-based proteomics now enable the ability to look directly at the consequences of genomic aberrations, providing deep and quantitative analyses of tumor tissues. Integration of proteins and their post-translational modifications identified by proteomics with genomic, epigenomic, and transcriptomic data constitutes the new field of proteogenomics, and it is already leading to new biological and diagnostic knowledge with potential to improve our understanding of malignant transformation and therapeutic outcomes. I will describe recent developments and key findings obtained using proteogenomics to analyze Lung and Breast Cancer (1-2), the two most dominant cancers world-wide, and describe proteogenomic methods (3) being developed to address clinical questions.

Presented By:

Shankha Satpathy, Broad Institute, Massachusetts, United States

---

To Understand One, We Must Study Many - The Need for High-throughout Proteomics and Good Cohort Design

Since its inception, proteomics was poised to revolutionize clinical diagnostics by providing novel biomarkers. Currently, however, proteomics has not met its own expectations particularly because large population sizes are required for meaningful biomarkers that can account for genetic, environmental and lifestyle variance. A main bottleneck has been the suboptimal throughput. Enabling throughput will allow the robust processing and analysis of statistically meaningful samples numbers, i.e., hundreds to thousands of samples within a reasonable timeframe. Our recent, involvement in several NIAID immunophenotyping studies focused on understanding COVID-19 impact on the immune system, the efficacy of vaccination, and ontogeny in early life with thousands of plasma samples, highlighted the importance of throughput in collecting high quality data. The Steen Lab has developed a high throughput proteomics platform that allows for the robust processing and analysis of such large numbers of plasma samples. These efforts leveraged the methodology and experience of the Steen Lab to process and analyze hundreds of urine samples with the aim of identifying non-invasive biomarker candidates. The presentation will describe method development from the Steen Lab towards i) high throughput proteomics, ii) improved depth of the plasma proteome, and iii) meta-analyses converting many small studies and datasets into a large dataset with statistical relevance. Our platform which incorporates these approaches will enable the analysis of large sample numbers which will bring allow us to understand deviations from normal despite natural interpersonal variability, i.e., truly personalized medicine.

Presented By:

Hanno Steen, Harvard Medical School, Massachusetts, United States

---

O05.01 | Comparative Protein Interaction Analysis of Genes Altered in Head and Neck Cancer Instructs the Selection of Therapeutic Targets

Danielle Swaney (1), Dana Steffen (2), Zhiyong Wang (2), Jisoo Park (2), Yusuke Goto (2), Margaret Soucheray (1), Neil Bhola (1), Kyumin Kim (1), Fan Zheng (2), Yan Zeng (1), Rachel O'Keefe (1), John Von Dollen (1), Mehdi Bouhaddou (1), Jason Kreisberg (2), Minkyu Kim (1), Daniel Johnson (1), Natalia Jura (1), Jennifer Grandis (1), J. Silvio Gutkind (2), Trey Ideker (2), Nevan Krogan (1)

(1) UCSF, San Francisco, CA, USA

(2) UC San Diego, San Diego, CA, USA

Tumor DNA sequencing has produced extensive lists of genetic alterations which can be challenging to translate to a molecular or clinical understanding of cancer. Here we outline a framework for elucidating genetic complexity through multidimensional protein-protein interaction maps and apply it to enhance the understanding of head-and-neck squamous cell carcinoma. This network uncovers 771 interactions from cancer and non-cancerous cell states including wild-type and mutant protein isoforms. To compare across different cell states, we have implemented and differential interaction score (DIS) to prioritize cancer-enriched interactions. Utilization of the DIS revealed a novel cancer-enriched association of the FGF receptor tyrosine kinase 3 with Daple, a guanine-nucleotide exchange factor, resulting in activation of Gαi and PAK1/2 to promote cancer cell migration. Furthermore, we have quantitatively compared the interaction landscape for 15 different mutations of PIK3CA, and observe a mutation specific interaction between the HER3 receptor tyrosine kinase and PIK3CA that can inform the response to HER3 inhibition in vivo. We anticipate the application of this framework will be valuable to understand the underlying biology of different cancers as well as other diseases.

Contact: Danielle Swaney, Danielle.swaney@ucsf.edu

O05.02 | Proteogenomic analysis unveils the HLA Class I presented immunopeptidome in melanoma and EGFR mutant lung adenocarcinoma

Yue Andy Qi (1), Tapan Maity (2), Constance Cultraro (2), Vikram Misra (2), Xu Zhang (2), Catherine Ade (3), Shaojian Gao (2), David Milewski (4), Khoa Nguyen (2), Mohammad Ebrahimabadi (5), Ken-ichi Hanada (3), Javed Khan (4), Cenk Sahinalp (5), James Yang (3), Udayan Guha (6)

(1) Thoracic and GI Malignancies Branch, Center for Cancer Research, NCI, NIH; Center for Alzheimer's Related Diseases, National Institute on Aging, NIH, Bethesda, MD, USA

(2) Thoracic and GI Malignancies Branch, Center for Cancer Research, NCI, NIH, Bethesda, MD, USA

(3) Surgery Branch, Center for Cancer Research, NCI, NIH, Bethesda, MD, USA

(4) Genetics Branch, Center for Cancer Research, NCI, NIH, Bethesda, MD, USA

(5) Cancer Data Science Laboratory, Center for Cancer Research, NCI, NIH, Bethesda, MD, USA

(6) Thoracic and GI Malignancies Branch, Center for Cancer Research, NCI, NIH; Bristol Myers Squibb, Bethesda, MD, USA

Immune checkpoint inhibitor and adoptive lymphocyte transfer-based therapies have shown great therapeutic potential for cancers with high tumor mutation burden (TMB). Here, we employed mass spectrometry (MS)-based proteogenomic large-scale profiling to identify potential immunogenic human leukocyte antigen (HLA) Class |-presented peptides in both melanoma, a "hot tumor" with high TMB, and EGFR mutant lung adenocarcinoma, a "cold tumor" with low TMB. We used cell line and patient-specific databases constructed using variants identified from whole-exome sequencing, as well as a de novo search algorithm from the PEAKS search algorithm to interrogate the mass spectrometry data of the Class I immunopeptidome. We identified 12 mutant neoantigens. Several classes of tumor-associated antigen-derived peptides were also identified. We constructed a cancer germline (CG) antigen database with 285 antigens and identified 42 Class I-presented CG antigens. We identified more than 1000 post-translationally modified (PTM) peptides representing 58 different PTMs. Our results suggest that PTMs play a critical role impacting HLA-binding affinity dramatically. Finally, leveraging de novo search and an annotated lncRNA database, we developed a novel non-canonical peptide discovery pipeline to identify 44 lncRNA-derived peptides that are presented by Class I. We validated MS/MS spectra of select peptides using synthetic peptides and performed HLA Class I binding assays to demonstrate binding of select neo-peptides and lncRNA-derived peptides to Class I proteins. In summary, we provide direct evidence of HLA Class I presentation of a large number of mutant and tumor-associated peptides for potential use as vaccine or adoptive cell therapy in melanoma and EGFR mutant lung cancer.

Contact: Yue Andy Qi, andy.qi@nih.gov

Parallel Session 06: Cardiovascular Proteomics

Identifying Translated Alternative Protein Isoforms

Although many genes undergo alternative splicing, the protein-level existence and by extension molecular function of most alternative isoform transcripts are unknown. We developed a pipeline and software that combines RNA sequencing and mass spectrometry to find alternative protein isoforms in the human proteomes. The method analyzes transcript junctions to construct custom tissue-specific protein databases containing likely-translated protein sequences. We re-analyzed >79 million mass spectra in 12 human tissues from public data on ProteomeXchange, and found over 1,500 mon-canonical protein isoforms including candidate undocumented peptides. This method was further applied to understand isoform shifts in human induced pluripotent stem cell differentiation into cardiomyocytes. Functionally, we found a significant enrichment of alternative isoform sequences to intersect with disordered protein regions as well as post-translational modification sites. In ongoing work, we are developing targeted mass spectrometry assays with the aid of machine learning predicted spectra to verify isoform peptides in human cardiac and cell culture models.

Presented By:

Maggie Lam, University of Colorado Denver, Colorado, United States

---

The Extracellular Matrix Proteome in Cardiac Physiology

Following myocardial infarction (MI), the left ventricle (LV) undergoes a series of cardiac wound healing responses that involve stimulation of robust inflammation to clear necrotic myocytes and tissue debris and induction of extracellular matrix (ECM) protein synthesis to generate a scar. Proteomic strategies provide us with a means to index the ECM proteins expressed in the LV, quantify levels, determine molecular and cellular physiology, and explore interactions. This talk will focus on the major cell types that coordinate cardiac wound healing, namely the infiltrating leukocytes and the cardiac fibroblasts. We will discuss efforts in proteomics research that have expanded our understanding of post-MI LV remodeling, concentrating on the strengths and limitations of different proteomic approaches to glean information that is specific to ECM turnover in the MI setting. In summary, this talk will provide an overview of how cardiac ECM proteomics has evolved over the last decade and will provide insight into future directions that will drive our understanding of cardiac ECM turnover in the MI LV.

Presented By:

Merry Lindsey, University of Nebraska Medical Center, Nebraska, United States

---

O06.01 | Induced Pluripotent Stem Cell-derived Engineered Cardiac Tissue Model of Hypertrophic Cardiomyopathy Enabled by Integrated Functional Assessments and Top-down Proteomics

Jake Melby (1), Willem de Lange (2), Jianhua Zhang (3), David Roberts (1), Stanford Mitchell (4), Trisha Tucholski (1), Gina Kim (3), Andreas Kyrvasilis (5), Sean McIlwain (6), Timothy Kamp (3), J. Carter Ralphe (2), Ying Ge (7)

(1) Chemistry; University of Wisconsin-Madison, Madison, WI, United States

(2) Pediatrics; University of Wisconsin-Madison, Madison, WI, United States

(3) Medicine; University of Wisconsin-Madison, Madison, WI, United States

(4) Molecular and Cellular Pharmacology; University of Wisconsin-Madison, Madison, WI, United States

(5) Cellular and Regenerative Biology; University of Wisconsin-Madison, Madison, WI, United States

(6) Biostatistics; University of Wisconsin-Madison, Madison, WI, United States

(7) Chemistry and Cellular and Regenerative Biology; University of Wisconsin-Madison, Madison, WI, United States

Human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CM) are a promising model system for disease modeling, cardiotoxicity screening, and drug discovery. 3-dimensional (3D) engineered cardiac tissue (ECT) constructs made from hiPSC-CMs are particularly appealing due to their closer representation of the structural and functional complexity of the heart compared to 2-dimensional hiPSC-CM monolayers. Recently, we have developed an integrated method that permits sequential assessment of functional properties and top-down proteomics from the same single hiPSC-ECT construct. The ability to perform functional and molecular assessments using the same hiPSC-ECT construct allows for more reliable correlation between observed functional performance and underlying molecular events, and thus is critically needed. Our method allowed for quantification of differences in isometric twitch force and the sarcomeric proteoforms between two groups of hiPSC-ECTs that differed in the duration of time spent in 3D-ECT culture. By using this integrated method we discovered a new and strong correlation between the measured contractile parameters and the phosphorylation levels of alpha-tropomyosin between the two groups of hiPSC-ECTs. We used this method to study a hiPSC-ECT model of a commonly inherited cardiovascular disease, hypertrophic cardiomyopathy (HCM). While the end stage HCM phenotype is well studied, the molecular events leading to disease pathogenesis remain poorly understood. hiPSC-ECTs of control and CRISPR-Cas9 knockouts (KOs) of cardiac myosin binding protein C (cMyBP-C), a protein that is encoded by a gene that approximately 40% of mutations are found in HCM patients, were generated to compare differences in functional assessments and top-down proteomics of sarcomere proteoforms. Our results indicate that the kinetics, twitch force magnitude, and sarcomeric proteoform landscape of the CRISPR-Cas9 KO hiPSC-ECTs were significantly altered compared to the control hiPSC-ECTs. By integrating innovative tissue engineering techniques, functional assessments, and proteomics technologies, this project will establish novel methods to probe the molecular mechanisms underlying HCM.

Contact: Jake Melby, jake.melby@wisc.edu

O06.02 | Distinctive Mechanisms in Hypertrophy and Heart Failure

Aleksandra Binek (1), Justyna Fert-Bober (1), Daniel Soetkamp (1), Simion Kreimer (1), Alejandro Rivas (1), Anna Pyzel (1), Amy Bradshaw (2), Michael Zile (2), Jennifer Van Eyk (1)

(1) Cedars-Sinai Medical Center, Los Angeles, CA, USA

(2) Medical University of South Carolina, Charleston, SC, USA

Introduction: Patients with heart failure and a preserved ejection fraction (HFpEF) present heart function abnormalities that remain poorly understood. Defining proteomic signature of HF that is independent of left ventricular hypertrophy (LVH) should allow for stratification of its subtypes and potential mechanism that contributes to the disease.
Methods: Intraoperative left ventricular (LV) myocardial biopsies were obtained from patients (n=21) recruited to undergo coronary artery bypass grafting (CABG). Patients were categorized to: non-hypertensive control (n=9), LVH (n=5), and HFpEF (n=7). Myocardial tissue was sub-fractionated into: cytoplasmic- (neutral pH), myofilament- (acidic pH), and membrane-enriched extract (SDS-soluble). All fractions were assessed for protein quantity and Lys/Arg modifications using liquid chromatography mass spectrometry (LC-MS).
Results: In HFpEF, 13% of the cardiac LV proteome changed compared to control heart, with a substantial proportion (77%) decreasing in quantity across all three cardiac fractions, while with LVH, 61% of the proteomic LV changes were increased. Although glycolysis and gluconeogenesis increased in both cardiopathies with respect to control, in HFpEF was less pronounced than in LVH. Modified proteome of the HFpEF was dominated by decreases in protein succinylation and to a lesser degree in methylation when compared to control group. This general trend of down-regulation of succinylation can be attributed to depletion in the levels of succinyl-CoA, the cofactor of enzymatic Lys succinylation. Importantly, there was a striking discordant activation/inhibition of cell survival and proliferation pathways between the HFpEF and LVH. Two major upstream regulator clusters linked the proteome changes in cell growth and proliferation to RICTOR and Myc that showed completely opposite trends in LVH and HFpEF groups.
Conclusions: HFpEF has a unique proteome signature compared to LV hypertrophy profile which does not arise from sub-proteome involved in contraction but rather is involved in overall cellular metabolism signaling.

Contact: Aleksandra Binek, bineka@cshs.org

Parallel Session 07: Technology Innovation in Multi-omics

Structural Biology on the Proteome-Wide Scale: An In-Cell Mass Spectrometry-Based Protein Footprinting Method

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

Presented By:

Lisa Jones, University of Maryland, Maryland, United States

Multiplex Chemical Tags for High-Throughput Comparative Multiomic Analyses of Cerebrospinal Fluids in Alzheimer's Disease

Glycosylation is one of the most important and most complex protein post-translational modifications. Alterations in glycomic profiles have been linked to various diseases, including cancer, neurodegenerative disorders, and cardiovascular problems. Thus, new methods are needed for quantitative analysis of glycans to facilitate elucidation of the diverse biological roles of glycans and their roles in human diseases. Advances in mass spectrometry (MS)-based glycoproteomics and glycomics are increasingly enabling qualitative and quantitative approaches for site-specific structural analysis of protein glycosylation. However, quantitative analysis of native glycans and intact glycopeptides remains to be challenging due to high complexity and diversity of glycan structures, difficulty of synthesizing glycan standards, the relatively low response in MS detection, and the wide dynamic range of glycans in clinically relevant samples. In this presentation, I will describe our recent efforts in developing both MS1 and MS2-based relative quantification strategies for proteomic, glycoproteomic and glycomic analyses in biological samples. Specifically, we are developing multiplexed isobaric and isotopic tagging strategies to discover, identify and evaluate candidate biomarkers of Alzheimer's disease (AD) in cerebrospinal fluids (CSFs) obtained from asymptomatic cognitively-healthy middle-aged adults, older cognitively-normal adults, and patients with mild cognitive impairment (MCI) and AD. A large-scale comparative glycoproteomic analysis via the 12-plex DiLeu (N,N-dimethyl leucine) tagging strategy revealed distinct glycosylation patterns and dynamic changes of certain glycoforms in CSF samples collected from the control, MCI, and AD patients. More than 1500 intact N-glycopeptides were quantified, with 19 N-glycopeptides exhibiting dysregulation in the progression of AD, revealing disease stage-specific N-glycopeptides. Furthermore, we also designed and synthesized a set of multiplex isobaric multiplex reagents for carbonyl containing compounds (SUGAR) tags for high throughput glycomic analyses. Collectively, these cost-effective and novel mass defect-based labeling reagents enable robust, sensitive, and accurate proteomic, glycopeptidomic, and glycomic analysis with enhanced quantitative performance and structural elucidation capabilities.

Presented By:

Lingjun Li, University of Wisconsin Madison, Wisconsin, United States

---

O07.01 | A Multi-Omic Investigation into the Role of the APOE Genotype in Alzheimer's Disease

Erin Baker (1), Melanie Odenkirk (1), Kristin Burnum-Johnson (2), Anna Lipton (2), Karen Butler (1), Jaclyn Kalmar (1), Brendan MacLean (3), Michael MacCoss (3), Edward Fox (4), Thomas Montine (4), David Muddiman (1), Blaine Roberts (5)

(1) Department of Chemistry, North Carolina State University, Raleigh, NC, USA

(2) Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

(3) Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA, USA

(4) Department of Pathology, Stanford University, Stanford, CA, USA

(5) Department of Biochemistry, Emory University School of Medicine, Atlanta, GA, USA

Alzheimer's disease (AD) is a progressively common neurodegenerative disease that affects 47 million people worldwide. Despite extensive efforts, AD neuropathology still remains confounded as underlying factors such as age and the APOE allele type alter AD progression mechanisms and challenge the development of effective treatments and intervention strategies. To gain a better understanding of the neuropathology of AD, we evaluated molecular mechanisms reflective of APOE allele types in two brain regions. Thus, post-mortem cerebellum (CBM) and frontal cortex (FCX) brain tissue collected from 62 patients [29 healthy control (HC) and 33 AD] having ε2/ε3, ε3/ε3, ε3/ε4, or ε4/ε4 allele groups. Multi-omic evaluations of the proteins, lipids and metals for each brain tissue sample were conducted to holistically characterize the genotype differences. Bottom-up proteomic and lipidomic analyses were performed by coupling liquid chromatography, ion mobility spectrometry and mass spectrometry (LC-IMS-MS) measurements to provide confident peptide and lipid identifications, and 15 metals were assessed with ICP-MS. Evaluation of the multi-omic studies showed molecular changes in the AD and HC comparisons for the different allele types, reflecting distinct AD progression mechanisms. Namely, proteins that mapped to the KEGG Alzheimer's pathway were dysregulated to the greatest degree for the ε3/ε4 allele group in the CBM and the ε3/ε3 group in the FCX. Conversely, the ε2/ε3 AD versus HC comparison was the least perturbed in both brain regions. Lipidomic analyses revealed heightened dysregulation in the FCX for the ε3/ε3 and ε3/ε4 allele types, while the ε2/ε3 AD group showed little dysregulation relative to the ε2/ε3 HC except for lysolipids within the CBM being significantly downregulated. Metallomics further elucidated Zn66, Cu63 and Fe56 all being upregulated in the FCX tissue for AD patients. Currently, glycan assessments are also being performed to gain an even further understanding of molecular changes occurring in the AD brains.

Contact: Erin Baker, ebaker@ncsu.edu

O07.02 | Uncovering Isoform Diversity of the Human Proteome by Long-Read Proteogenomics

Rachel Miller (1), Simi Kaur (2), Ben Jordan (3), Rob Millikin (1), Kyndalanne Pike (1), Isabella Whitworth (1), Michael Shortreed (1), Christina Chatzipantsiou (4), Raymond LeClair (5), Anne Deslattes Mays (6), Lloyd Smith (1), Gloria Sheynkman (3)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

(2) Department of Biophysics, University of Wisconsin-Madison, Madison, WI, USA

(3) Department of Molecular Physiology & Biological Physics, Center for Public Health Genomics, UVA Cancer Center, University of Virginia School of Medicine, Charlottesville, VA, USA

(4) Lifebit Biotech Ltd., London, UK

(5) Springbok, LLC, Boston, MA, USA

(6) Science and Technology Consulting, LLC, Farmington, CT, USA

The human proteome originates from ~20,000 protein-coding genes and its complexity results from genetic variation, alternative splicing and post-translational modifications. Mass spectrometry (MS)-based characterization of the human proteome typically employs a bottom-up approach, in which identification of peptides and inference of proteins is critically dependent on the accuracy of the database used. An ideal database would mirror the proteins expressed in the sample. Used in most analyses, generic reference protein databases are designed to broadly represent a proteome

regardless of biological variability. Unfortunately, such reference databases fail to capture the complexity of individual samples, deviating significantly from the protein isoforms present in the sample by both lacking isoform diversity for some genes, and having excessive diversity for others. Fortunately, new sequencing technologies capable of capturing full-length transcripts with high fidelity have opened the door for the generation of sample-specific databases encompassing the isoform diversity of complex proteomes. Here we leverage the latest generation of PacBio long-read sequencing to provide a full-length transcriptome for translation into a protein isoform database capable of capturing sample-specific biological diversity. PacBio sequencing was obtained for the Jurkat cell line, producing 4 million reads which clustered to 175,916 distinct transcripts. Widespread differences were observed between the PacBio-derived database and the UniProt and GENCODE reference databases, with over 75% of the PacBio transcripts annotated as novel. We have developed a long-read proteogenomic analysis pipeline based on the Nextflow/CloudOS framework to identify transcripts, generate a sample-specific protein database, perform database searching, compare analysis results against reference protein databases, and leverage transcript abundances to inform protein inference. The sample-specific isoform database constructed via the pipeline contains 112,479 protein entries for approximately 13,000 genes. This work will provide a foundation for the proteomics community to integrate PacBio long-read sequencing into their work for the delineation of a proteome's isoform diversity.

Contact: Rachel Miller, rmmiller22@wisc.edu

Parallel Session 08: Epigenetics and Chromatin Biology

Quantitative Proteomics for Understanding Genome Mutations that Affect Human Health

With the large scale genomics data that has been generated over the years, many mutations on chromatin related proteins have been discovered. Some of these are mutations to enzymes that modify chromatin proteins such as histones, while other mutations are found on chromatin remodeling complexes and even the histones themselves. In fact, hundreds of histone mutations have been recently described to be high frequency mutations in a variety of cancers, and even linked to neurological disorders. Here we explore how some of these mutations affect chromatin function, and cause dysregulation of gene expression patterns needed for proper cellular signaling and development. Using a combination of proteomics, genomics and cell biology approaches, we share our latest data towards understanding how histone mutations lead to human disease.

Presented By:

Ben Garcia, University of Pennsylvania, Pennsylvania, United States

---

Functional Characterization of Lysine Lactylation and Hydroxybutyrylation

Short-chain fatty acids and their corresponding acyl-CoAs sit at the crossroads of multiple metabolic pathways and play important roles in diverse cellular processes. Two noteworthy examples are the newly identified protein posttranlation modifications (PTMs), lysine lactylation (Kla) and β-hydroxybutyrylation (Kbhb), which are derived from lactate and β-hydroxybutyrate, respectively. We demonstrated that histone Kla directly stimulates transcription, and established a novel function for lactate to regulate gene expression in macrophages. Here we also report that the acyltransferase p300 can catalyze the enzymatic addition of Kbhb, while HDAC1 and HDAC2 enzymatically remove Kbhb, both in vitro and in vivo. We demonstrate that p300-dependent histone Kbhb can directly mediate gene transcription. In addition, a comprehensive analysis of Kbhb substrates in mammalian cells has identified 3248 Kbhb sites on 1397 substrate proteins. These results not only illustrate the landscape of new lysine acylation pathways, but also open new avenues for studying diverse functions of cellular metabolites associated with these PTM pathways.

Presented By:

He Huang, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, China

---

O08.01 | Elucidating the Mechanism of Germline Histone H3.3 Mutations on Neurodevelopment

Khadija D. Wilson (1), Elizabeth Porter (1), Peter Klein (2), Elizabeth Bhoj (3), Benjamin A. Garcia (1)

(1) Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, US

(2) Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, US

(3) Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, US

DNA in eukaryotic cells is organized around histone octamers comprising of two copies of histones H2A, H2B, H4, and H3. Histone H3.3 (H3.3), a histone variant, often found at actively transcribed loci, plays a role in cellular inheritance and identity. Ablation of H3.3 expression leads to loss of active gene states and dysfunction of heterochromatin and telomeric structures. H3F3A/B, the two genes known to encode H3.3, are ubiquitously expressed with highest expression in the brain. Our collaborators at the Children's Hospital of Philadelphia discovered a cohort of patients with germline missense mutations on both H3F3 genes. These patients suffer from a severe neurodevelopmental disorder characterized by intellectual disability,

craniofacial and congenital defects. Surprisingly, no malignancies were detected. The underlying cause of the disease is unknown, however H3.3 has been implicated in many basic central nervous system functions. We hypothesized the mutant H3.3 (mH3.3) changes the regulatory capacity of mH3.3-containing nucleosomes thereby altering gene expression and ultimately the cellular proteome. To this end, we are employing a combination of techniques and methodology including bottom-up proteomics and neurodevelopmental assays in Xenopus laevis. We developed stable cell lines expressing either H3.3 or mH3.3 and we assessed histone post translational modifications changes. We observed significant alterations to histone acetylation on H3 and H4 peptides. Performing quantitative proteomics in the HEK293T cells led us to determine that these mutations affect several cellular processes including folic acid metabolism, and post-synaptic density, all processes known to be dysregulated in other neurological syndromes. Additionally, we introduced mH3.3 in X. laevis embryos to assess development and craniofacial abnormalities. We observed reduced craniofacial cartilage, abnormal head shape, and impaired motility in the mutant tadpoles. Ultimately, our studies aim to identify potential protein targets involved in the neurodevelopmental impairments and to shed light into the molecular mechanisms of H3.3 in neurodevelopment.

Contact: Khadija Wilson, wilsonkh@pennmedicine.upenn.edu

O08.02 | Improving dimethyl labeling for its application to histone lysine methylation quantitative analysis

Junfeng Huang (1), Ngai Ting Chan (2), Zexin Zhu (1), Shengjie Zhang (2), Wei Xu (2), Lingjun Li (1)

(1) School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA

(2) McArdle Laboratory for Cancer Research, University of Wisconsin-Madison, Madison, WI, USA

Histone lysine methylation is an important post-translational modification (PTM) occurring on the lysine residue, and plays crucial roles in many biological processes. Abnormal histone lysine methylation has been associated with many fetal diseases. Each of lysine residues can be mono-, di- or trimethylated and the number of methyl groups added on the same modification site can have different functions. Moreover, histone lysine methylations can perform active or repressive transcription according to different modification sites. Therefore, a comprehensive characterization of site occupancy of histone lysine methylation is helpful to fully elucidate its function. As histones are lysine-rich and super-hydrophilic proteins, trypsin digestion of histones results in very small and hydrophilic peptides which often suffers from substantial losses during sample preparation and are difficult to be detected by conventional reversed phase LC-MS. Moreover, methylated lysine will be blocked for trypsin digestion, which is different from free lysine and will affect the site occupancy calculation of lysine methylation. Previously, lysine propionylation label has been developed to overcome this drawback. However, propionylation label can increase the peptide hydrophobicity and decrease peptide charge state compared with native lysine methylation peptides and affect peptide LC retention and MS signal intensity, which will make the relative quantitation of lysine methylation inaccurate. Herein, we introduced stable isotope dimethyl labeling to block free lysine residues that will generate same tryptic peptide sequence and distinguish it from native lysine methylation. After tryptic digestion, all peptide N-termini were labeled with propionylation to increase their hydrophobicity and eliminate the sample loss of small and hydrophilic peptides. Moreover, we have optimized reaction condition to avoid a general defect of dimethyl labeling following standard protocol. We used this improved method to quantify changes in histone H3K27 methylation induced by small molecule inhibitors of lysine demethylases and siRNA knockdown of epigenetic modifiers ASH2L and WDR5.

Contact: Junfeng Huang, jhuang373@wisc.edu

Parallel Session 09: Protein Post-Translational Modifications

Functional Proteomics Analysis to Discover Novel Oxygen-sensing Posttranslational Modification Pathways

Oxygen availability governs the energy homeostasis and development of living organisms and is widely implicated in diseases such as cancer and metabolic disorders. Cellular adaptation to the changes of the oxygen concentration are regulated by diverse signaling and transcriptional mechanisms. Posttranslational modifications (PTMs) including proline hydroxylation (Hyp), phosphorylation and ubiquitination are important signaling mechanisms that mediates such responses. Despite of its physiological significance, system-wide characterization of oxygen-response pathways has been largely focused on transcriptome profiling and gene expression analysis. We will discuss our recent efforts to apply high resolution mass spectrometry and quantitative proteomics analysis to discover and biochemically characterize novel oxygen-dependent regulatory pathways and PTM signaling events. Through PTM-specific analysis, we established a functional annotation database of oxygen-dependent proline hydroxylation proteome and revealed hypoxia-response phosphorylation and ubiquitination signaling in cancer and neuronal cells. Our study showed that proline hydroxylation participates in protein structural maintenance and is significantly enriched on functional domains in highly-connected interaction networks. Combining with quantitative interactome and protein dynamics analysis, we identified new ubiquitination events regulating hypoxia-response enzyme activity and DNA damage repair pathways that promote genome instability under hypoxic microenvironment. Overall, these studies expand the landscape of oxygen-dependent regulation of protein structure and functions and provide mechanistic insights into the hypoxia-mediated regulation of cellular networks in disease.

Presented By:

Yue Chen, Associate Professor of Biochemistry, University of Minnesota, Minnesota, United States

---

Methods to Illuminate the Regulation and Function of the Phosphoproteome

Protein phosphorylation functions as a highly versatile signaling unit integral to almost every cellular process. Protein phosphorylation and cellular signaling have been studied for decades, with mass spectrometry developments enabling signaling studies at a proteome-wide scale. These proteomic studies have been crucial for cataloguing thousands of protein phosphorylation events and their regulation. However, as phosphoproteomics experiments have increased in scale, new challenges have emerged, including scaling-up sample processing, dealing with data sparsity and inaccurate phosphosite localizations, challenges in elaborating network models, and the limited knowledge on which phosphosites may be functionally relevant. I will present some of our recent technological developments in phosphoproteomics that aim to bridge some of these gaps with the final goal of achieving a systematic and accurate characterization of the phosphoproteome.

Presented By:

Judit Villen, University of Washington, Washington, United States

---

O09.01 | Quantitative PTM Maps of Human Pathologic Tau Identify Patient Heterogeneity and Define Critical Steps in Alzheimer's Disease Progression

Mukesh Kumar (1), Hendrik Wesseling (1), Waltraud Mair (1), Christoph Schlaffner (1), Pieter Beerepoot (1), Benoit Fatou (2), Amanda Guise (1), Long Cheng (1), Hanno Steen (2), Judith Steen (1)

(1) Neurobiology, Harvard University, Boston, MA, USA

(2) Pathology, Harvard University, Boston, MA, USA

To elucidate the role of Tau protein isoforms and its post translational modification (PTM) stoichiometry in Alzheimer's disease (AD), we generated a high resolution quantitative proteomic map of 95 PTMs on multiple isoforms of Tau isolated from the post-mortem human tissue from 49 AD and 42 control (CTRL) subjects. While Tau PTM maps reveal heterogeneity across subjects, a subset of PTMs display high occupancy and patient frequency for AD suggesting importance of these PTMs in disease. Hierarchical clustering analyses indicate that PTMs occur in an ordered manner leading to Tau aggregation. Further, we performed an in-depth characterization of Tau proteoforms with prion-like behavior and defined a minimum set of PTMs associated with seeding of Tau aggregation. We first used biochemical fractionation methods to isolate the detergent (sarkosyl) insoluble pathological Tau aggregates from AD and CTRL brain tissue lysate. Insoluble fractions were further processed and analyzed by LC-MS/MS. We employed FLEXITau, an in-house developed targeted, high-throughput, quantitative mass spectrometry (MS) method which provides absolute quantification and unbiased stoichiometric information of Tau modifications on all detectable peptides. The same samples were also analyzed by shotgun proteomics using data dependent acquisition method for discovery analysis of PTM and for label-free quantification. In summary, we identified critical molecular features and mechanisms that leads to Tau aggregation. Our data on pathological Tau shows the enrichment of 0N and 4R isoforms, underrepresentation of the protein C-terminal, enrichment of the microtubule binding domain (MBD), hyperphosphorylation in the proline rich region (PRR) leading to an increase in negative charge and ubiquitination and acetylation of multiple lysine residues in MBD leading to a decrease in positive charge and these modifications facilitates Tau aggregation. Our data provides a guide for diagnostic criteria and therapeutic strategies for each stage of Alzheimer's disease and we propose a processive model for Tau fibrilization.

Contact: Mukesh Kumar, mukesh.kumar@childrens.harvard.edu

O09.02 | Proteogenomic Characterization of Pancreatic Ductal Adenocarcinoma

Liwei Cao (1), Chen Huang (2), Daniel Cui Zhou (3), Oliver F. Bathe (4), Daniel W. Chan (1), Ralph H. Hruban (1), Li Ding (3), Bing Zhang (2), Hui Zhang (1)

(1) Department of Pathology, Johns Hopkins University, Baltimore, MD, USA

(2) Lester and Sue Smith Breast Center, Baylor College of Medicine, Houston, TX, USA

(3) Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA

(4) Departments of Surgery and Oncology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Pancreatic ductal adenocarcinoma (PDAC) is deadliest of all solid malignancies with a 5-year survival rate less than 10%. Previous genomic characterizations have delineated the genetic alterations associated with PDAC, with subsequent transcriptomic analyses resulting in the development of transcriptomic-based subtypes of PDAC. Despite these substantial efforts, the molecular mechanism of this cancer type is still not fully uncovered for precision medicine. To this end, The Clinical Proteomic Tumor Analysis Consortium conducted the first comprehensive characterization of 140 pancreatic cancers, 67 normal adjacent tissues, and 9 normal pancreatic ductal tissues using whole genome sequencing, whole exome sequencing, methylation, RNA-seq, miRNA-seq, proteomics, phosphoproteomics, and glycoproteomics. Proteomics, phosphoproteomics, and glycoproteomics measurements were deployed to characterize proteins and protein modifications, resulting in identification of 11,662 proteins, 51,469 phosphosites, and 34,024 glycopeptides, respectively. Genomic, transcriptomic, and epigenetic characterization were conducted on the same tissues for integrated proteogenomic analysis. Tumors with sufficient neoplastic purity were identified by histology-based, DNA methylation-based and RNA-based molecular deconvolution to address the inherent low neoplastic cellularity of pancreatic tumor samples. We verified KRAS is the major driver gene in PDAC. In addition, we found that genetic events, including somatic mutations and copy number alterations could impact gene and protein expression, as well as protein modifications including phosphorylation and glycosylation. Integrating global proteomic and phosphoproteomic measurements, we identified over-expressed kinase substrates and their corresponding kinases, uncovering potential targets for therapy. The results of our molecular and cellular subtyping revealed immune-hot subtype tumors that would be ideal for immunotherapy, as well as the underlying mechanisms associated with immune-cold subtypes, including endothelial cell remodeling, glycolysis, and cell junction dysregulation. This comprehensive proteogenomic characterization of PDAC provides a valuable resource for the uncovering the molecular mechanism of this cancer type and thus paves the way for discovery of novel early detection and therapeutic targets.

Contact: Liwei Cao, lcao11@jhmi.edu

Parallel Session 10: Top-Down Proteomics and Native MS

Building a Community of Top Down Proteomics to Advance Proteoform Measurement & Biology

While top down mass spectrometry has become synonymous with the direct analysis of intact proteins and their complexes, the term more generally denotes an approach to measurement that recognizes the value of retaining as much information as possible about a system prior to analysis. By avoiding proteolytic digestion, proteoform-specific identifications can be made -directly. A series of vignettes will focus on both denatured and native modes of Top Down Proteomics. I will also describe a few recent advances top down MS, like the most recent breakthroughs in individual ion mass spectrometry (i2MS) in orbitraps (joint work with the group of Mike Senko at Thermo Fisher Scientific). I will also describe a new sample introduction approach called “SampleStream” that enables simultaneous buffer exchange and >20x concentration of samples in either denaturing or native modes. Providing information from intact complex mass (MS1) to subunits and their backbone fragment ions, native top-down MS even enables identification and characterization of unknown protein complexes. By more faithfully preserving post-translational modifications and non-covalent interactions throughout the measurement process, top down mass spectrometry is positioned to make basic and translational proteomics more efficient and valuable, particularly in the detection and assignment of function to proteoforms and their PTMs underlying human wellness and disease.

Presented By:

Neil Kelleher, Northwestern University, Illinois, United States

---

Towards Illuminating the Dark Matter of Mass Spectrometry for a Top-Down View of Native Protein Structures

Native mass spectrometry (MS) of proteins and protein assemblies reveals size and binding stoichiometry. But elucidating their structures to understand their function is more challenging. Using electrospray ionization (ESI), relative charging by native ESI-MS appears to give some information on protein folding. We show that native MS and native top-down MS (TDMS), i.e., fragmentation of the gas-phase protein, can be effective for deriving structural information for soluble and membrane protein complexes, and much of this information can be correlated to the solution-phase structure. Native top-down MS generates information on the surface topology, ligand binding sites, and post-translational modifications (PTMs) of protein complexes. We use native top-down MS to investigate the molecular action of compounds that prevent amyloid fibril formation in neurodegenerative diseases such as Alzheimer’s and Parkinson’s disease. Product ions generated by TDMS contain either the N- or C-terminus (terminal fragments) or without either termini (internal fragments). Conventionally, internal fragments have been ignored in analyses due to the inability to reliably assign them. Previous reports by others and more recently by our group (Zenaidee et al., JASMS 2020) suggested that including internal fragment assignments in top-down protein MS could increase sequence coverage significantly, which will allow for efficient analysis of protein structure and will be useful for localizing sites of PTMs and protein-ligand sites of interactions. We hope that illuminating the “dark matter” of previously unassigned peaks will transform how MS can elucidate protein structure and proteoforms.

Presented By:

Joseph Loo, University of California Los Angeles, California, United States

---

O10.01 | Novel Strategies for Top-down Proteomics of Endogenous Membrane Proteins

Kyle Brown (1), Trisha Tucholski (1), Christian Eken (1), Ying Ge (2)

(1) Department of Chemistry University of Wisconsin-Madison, WI, 53719, Madison, WI, USA

(2) Department of Chemistry, Department of Cell and Regenerative Biology, Human Proteomics Program, University of Wisconsin-Madison, WI, USA

Membrane proteins play important cellular roles and represent critical drug targets. Post-translational modifications (PTMs) and changes in protein isoforms can greatly influence the function of membrane proteins at a molecular level. Top-down mass spectrometry (MS) has emerged as a powerful technology to achieve proteoform-resolved information and represents a promising tool to characterize endogenous membrane proteins. However, the low expression level and general hydrophobicity of endogenous membrane proteins coupled with the overall complexity of the membrane subproteome makes their extraction, purification, and characterization a challenge for top-down practitioners. Here, we developed complementary top-down proteomic approaches to characterize endogenous integral membrane proteins. First, we utilized our recently developed photocleavable surfactant, Azo, to solubilize membrane proteins after soluble proteins were depleted, enabling a streamlined workflow for proteoform characterization. Utilizing online liquid chromatography (LC) coupled to MS, we identified proteins from mitochondria, nucleus, plasma membrane, cytoskeleton, endoplasmic reticulum, cytoplasm and extracellular region including every subunit of the ATP synthase complex from cardiac tissue. Additionally, we observed various PTMs, such as acetylation, methylation, phosphorylation and localized a palmitoylation on the transmembrane domain (C36) of phospholamban (a crucial cardiac regulatory protein). This approach afforded high-throughput analysis, but the overall representation of membrane proteins among the identified species was lower than desired. Therefore, we developed a complementary method utilizing a cloud point extraction to enrich membrane proteins in a single step followed by size-exclusion chromatography and LC-MS/MS for membrane proteoform characterization. Significantly, we found that 70% of the proteins identified were membrane protein and 54% were integral membrane proteins with as many as 19 transmembrane domains (voltage-dependent anion-selective channel protein). However, the cloud point/SEC/LC-MS/MS approach was lower throughput compared to the Azo-enabled membrane proteomics approach. Overall these methods represent exciting opportunities to further develop integral membrane top-down proteomics for basic and clinical research.
Contact: Kyle Brown, kbrown33@wisc.edu

O10.02 | Automated capillary isoelectric focusing-tandem mass spectrometry for qualitative and quantitative top-down proteomics

Tian Xu (1), Liangliang Sun (1)

(1) Michigan State University, East Lansing, MI, United States

Top-down proteomics (TDP) aims to delineate proteomes in a proteoform-specific manner, which is vital for accurately understanding protein function in cellular processes. It requires a high-capacity separation of proteoforms before mass spectrometry (MS) and tandem mass spectrometry (MS/MS). Capillary isoelectric focusing (cIEF)-MS has been recognized as a useful tool for TDP in the 1990s because cIEF is capable of high-resolution separation of proteoforms. Previous cIEF-MS studies concentrated on measuring protein's mass without MS/MS, impeding the confident proteoform identification in complex samples as well as the accurate localization of post-translational modifications (PTMs) on proteoforms. Herein, for the first time, we present automated cIEF-MS/MS-based TDP for large-scale delineation of proteoforms in complex proteomes. Single-shot cIEF-MS/MS identified 711 proteoforms from an Escherichia coli (E. coli) proteome consuming only nanograms of proteins. Coupling two-dimensional size exclusion chromatography (SEC)-cIEF to ESI-MS/MS enabled the identification of nearly 2000 proteoforms from the E. coli proteome. Label-free quantitative TDP of zebrafish male and female brains using the SEC-cIEF-MS/MS quantified thousands of proteoforms and revealed sex-dependent proteoform profiles in brains. Particularly, we discovered several proteolytic proteoforms of pro-opiomelanocortin and prodynorphin with significantly higher abundance in male zebrafish brains as potential endogenous hormone proteoforms. Multi-level quantitative proteomics (TDP and bottom-up proteomics) of the brains revealed that the majority of proteoforms having statistically significant differences in abundance between genders showed no abundance difference at the protein group level. This work represents the first multi-level quantitative proteomic study of sexual dimorphism of the brain.
Contact: Liangliang Sun, lsun@chemistry.msu.edu

Parallel Session 11: Machine Learning and Computational Proteomics

Applications of Intelligent Data Acquisition for Multiplexed Proteomics

High-throughput quantitative proteomics requires efficient sample handling, fast mass spectrometric acquisition, and precise quantitation. Modern methods for multiplexed proteomics enable the utilization of isobaric chemical labeling to compare up to 16 samples simultaneously. However, acquisition of multiplexed proteomic quantitative data generally suffers from either slow acquisition speeds (SPS-MS3) or limited dynamic range (HRMS2). To bridge this divide we developed intelligent data acquisition strategies to acquire high accuracy quantifying scans based on real-time spectral matching. Real-time database searching of spectra during data acquisition enabled a 100% increase in data acquisition speed and improved quantitative accuracy. Since its initial development we have applied the improved efficiency of real-time search to interrogate cysteine reactivity, mTOR inhibition, and murine redox regulation. Thus, the integration of intelligent data acquisition with quantitative proteomics offers new avenues to explore and understand the proteome across conditions, tissues of origin or species of interest.

Presented By:

Devin Schweppe, University of Washington, Washington, United States

---

Advancing Precision Oncology Through Proteogenomics: Challenges and Opportunities

Powered by next generation sequencing-based genomics and mass spectrometry-based proteomics, proteogenomics is a rapidly evolving field in cancer research. Despite great potential, there are many informatics challenges in translating cancer proteogenomics data into novel therapeutic insights. In this talk, I will summarize efforts in my group in addressing challenges in data storage, quality control, data analysis, data interpretation, and data and tool dissemination. I will also show an example that demonstrate the use of proteogenomics data to identify a small panel of proteins biomarkers that may help match head and neck cancer patients to different targeted and immunotherapy drugs.

Presented By:

Bing Zhang, Baylor University, Texas, United States

O11.01 | Deep learning-derived evaluation metrics for benchmarking computational pipelines for the analysis of phosphoproteomic data

Wen Jiang (1), Bo Wen (2), Kai Li (2), Felipe Leprevost (3), Jamie Moon (4), Nathan Edwards (5), Tao Liu (4), Alexey Nesvizhskii (3), Bing Zhang (2)

(1) Department of Quantitative and Computational Biosciences, Baylor College of Medicine, Houston, Texas, United States

(2) Lester and Sue Smith Breast Center, Baylor College of Medicine, Houston, Texas, United States

(3) Department of Pathology, University of Michigan, Ann Arbor, Michigan, United States

(4) Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, United States

(5) Department of Biochemistry and Molecular & Cellular Biology, Georgetown University, Washington, DC, United States

Tandem mass spectrometry (MS/MS)-based phosphoproteomics is a powerful platform for global phosphorylation analysis. Applying different computational pipelines to the same dataset, however, may produce different phosphopeptide identifications and phosphosite localizations. Efforts have been made to compare computational pipelines using data from synthetic peptides, but performance comparison in actual application data has

been largely limited to numbers of phosphopeptide identifications due to the lack of ground-truth or appropriate evaluation metrics. We considered three deep learning-derived features as potential evaluation metrics. MusiteDeep predicts phosphosite probability for a given sequence as previously reported (AUROC=0.896); DeltaRT is defined as the absolute retention time (RT) difference between RTs observed and predicted by AutoRT; and spectral similarity is defined as the Spearman's correlation coefficient between spectra observed and predicted by pDeep2. Using a synthetic peptide dataset, we showed that DeltaRT and spectral similarity provided excellent discrimination between correct and incorrect peptide-spectrum-matches (PSMs) when incorrect PSMs were associated with wrong peptide sequences (AUROC=0.95 and 0.88, respectively). DeltaRT also provided excellent discrimination between correct and incorrect PSMs (AUROC=0.88) even when the incorrect PSMs were caused by only wrong phosphosite localization, while spectral similarity did relatively poorly in this scenario (AUROC=0.77). Based on these results, we used phosphosite probability and DeltaRT as evaluation metrics to benchmark four informatics pipelines, including the popular MaxQuant and three independent CPTAC pipelines, CPTAC1, CPTAC2, CPTAC3, on a tandem mass tag (TMT) dataset from analysis of human uterine tumors. CPTAC1 outperformed the other three pipelines in both sensitivity and specificity. In a TMT dataset from murine cell lines and a label-free dataset from human cell culture, CPTAC1 also outperformed MaxQuant. In summary, two deep learning-derived features, DeltaRT and phosphorylation probability, enabled systematic evaluation of computational pipelines for the analysis of phosphoproteomic data, and the CPTAC1 pipeline showed the best performance in our evaluation.

Contact: Wen Jiang, Wen.Jiang@bcm.edu

O11.02 | hu.MAP 2.0: Integration of Over 15,000 Proteomic Experiments Builds a Global Compendium of Human Multiprotein Assemblies

Kevin Drew (1), John Wallingford (2), Edward Marcotte (2)

(1) Biological Sciences, University of Illinois at Chicago, Chicago, IL, USA

(2) Molecular Biosciences, University of Texas at Austin, Austin, TX, USA

A general principle of biology is the self-assembly of proteins into functional complexes. Characterizing their composition is, therefore, required for our understanding of cellular functions. Unfortunately, we lack a comprehensive set of protein complexes for human cells. To address this gap, we developed a machine learning framework to identify protein complexes in over 15,000 mass spectrometry experiments which resulted in the identification of nearly 7,000 physical assemblies. We show our resource, hu.MAP 2.0, is more accurate and comprehensive than previous resources and gives rise to many new hypotheses, including for 274 completely uncharacterized proteins. Further, we identify 259 promiscuous proteins that participate in multiple complexes pointing to possible moonlighting roles. We have made hu.MAP 2.0 easily searchable in a web interface (http://humap2.proteincomplexes.org/), which will be a valuable resource for researchers across a broad range of interests including systems biology, structural biology, and molecular explanations of disease.

Contact: Kevin Drew, ksdrew@uic.edu

Parallel Session 12: Multi-'omics - Methods and Applications

Proteome Activity Landscapes of Tumor Cell Lines Determine Drug Responses

Integrated analysis of genomes, transcriptomes, proteomes and drug responses of cancer cell lines (CCLs) is an emerging approach to uncover molecular mechanisms of drug action. We extend this paradigm to measuring proteome activity landscapes by acquiring and integrating quantitative data for 10,000 proteins and 55,000 phosphorylation sites (p-sites) from 125 CCLs. These data are used to contextualize proteins and p-sites and predict drug sensitivity. For example, we find that Progesterone Receptor (PGR) phosphorylation is associated with sensitivity to drugs modulating estrogen signaling such as Raloxifene. We also demonstrate that Adenylate kinase isoenzyme 1 (AK1) inactivates antimetabolites like Cytarabine. Consequently, high AK1 levels correlate with poor survival of Cytarabine-treated acute myeloid leukemia patients, qualifying AK1 as a patient stratification marker and possibly as a drug target. We provide an interactive web application termed ATLANTiC (http://atlantic.proteomics.wzw.tum.de), which enables the community to explore the thousands of novel functional associations generated by this work.

Presented By:

Bernhard Kuster, Technical University of Munich, Munich, Germany

---

Linking the Genome and the Proteome in Cancer

Cancer is often called a disease of the genome. Although DNA mutations are a common root cause, the cancer treatment focuses on the dysfunctional cellular state including aberrant protein abundance or phosphorylation signaling. Thus, improvements for cancer care require an integrated multi-omics perspective. This talk discusses data sources and analysis methods for multi-omics cancer analysis. We focus specifically on the functional consequences of DNA mutation and the context-dependent nature of driver mutations.

Presented By:

Sam Payne, Brigham Young University, Utah, United States

O12.01 | Dynamic 3D proteomes reveal protein functional alterations at high resolution in situ

Valentina Cappelletti (1), Hauser Thomas (1), Ilaria Piazza (1), Monika Pepelnjak (1), Liliana Malinovska (1), Tobias Fuhrer (1), Yaozong Li (2), Christian Dörig (1), Paul Boersema (1), Ludovic Gillet (1), Jan Grossbach (3), Aurelien Dugourd (4), Julio Saez-Rodriguez (4), Andreas Beyer (5), Nicola Zamboni (1), Amedeo Caflisch (2), Natalie de Souza (1), Paola Picotti (1)

(1) ETH Zürich, Institute of Molecular Systems Biology, Zürich, CH

(2) Department of Biochemistry, University of Zurich, Zürich, CH

(3) CECAD Research center, University of Cologne, Cologne, Germany

(4) Institute for Computational Biomedicine, University of Heidelberg, Heidelberg, Germany

(5) Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany

Biological processes are regulated by molecular events, such as intermolecular interactions, chemical modification and conformational changes, which do not affect protein levels and therefore escape detection in classical proteomic screens. Reasoning that these events affect protein structure, we tested whether a global readout of protein structure could detect various types of functional alterations simultaneously and in situ. We tested this idea using limited proteolysis coupled to mass spectrometry (LiP-MS), which monitors structural changes in thousands of proteins within a complex, native-like environment. In bacteria adapting to different nutrient sources and in yeast responding to acute stress, the structural readout, visualized as structural barcodes, captured enzyme activity changes, allosteric regulation, phosphorylation, protein aggregation and protein complex formation, with the resolution of individual regulated functional sites such as binding and active sites. Comparison with prior knowledge, including flux, phosphoproteomics and metabolomics data, showed that LiP-MS detects many known functional alterations within well-studied pathways. It suggested novel metabolite-protein interactions and enabled identification of a fructose-1,6-bisphosphate-based regulatory mechanism of glucose uptake in E. coli. The structural readout dramatically increases the coverage of classical protein expression profiling, generates mechanistic hypotheses, better links holistic and reductionist approaches, and paves the way for a new in situ structural systems biology.

Contact: Paola Picotti, picotti@imsb.biol.ethz.ch

O12.02 | A New Deconvolution Method for Studying Cell Type Composition in Tumor Microenvironment based on Bulk Proteogenomic Data

Francesca Petralia (1), Pei Wang (1)

(1) Genetics and Genomic Sciences, Icahn School of Medicine at Mt Sinai, New York, NY, USA

Characterizing the tumor microenvironment is extremely important to capture the heterogeneity across different patients and suggest targeted treatments. The fraction of different cell types in the tumor microenvironment can be estimated based on transcriptomic profiling of bulk tumor via deconvolution algorithms. A limitation of current deconvolution methods such as Cibersort is that they rely on a signature matrix derived based on RNA data which might not be suitable when performing the deconvolution based on other data types such as proteomic profiling. xCell – another very popular algorithm for tumor deconvolution – is a more flexible algorithm which takes as input only a list of key cell specific markers; however, it lacks to model directly cell type proportions. In order to overcome these limitations, we propose Rep-Bulk (Repulsive Bulk) – a new Bayesian method for the deconvolution of bulk tumor data based on transcriptomic and/or proteomic profiling. Contrary to xCell and Cibersort, our framework estimates different cell type fractions and the mean of gene expression in different cell types from the data, simultaneously. Prior knowledge is flexibly incorporated in the model through a repulsive prior. We demonstrate Rep-Bulk can outperform xCell and Cibersort on various synthetic data examples and real data examples based on RNA and proteomic data from public databases.

Contact: Francesca Petralia, francesca.petralia@mssm.edu

Parallel Session 13: New Strategies of Interactomics in Structural Biology

Pharmacological Insight Through Interactome Studies

Chemical cross-linking and mass spectrometry (XL-MS) technologies have been developed that enable new insight on the interactome, potentially bridging the fields of structural and systems biology. Our XL-MS experiments are investigating how interactomes in cells are altered during drug treatment, as well as what and how proteins bind with emerging therapies designed to treat mitochondrial dysfunction. This presentation will highlight our recent advances in cross-linker molecular design, informatics capabilities and our applications to drug treated cells, tissues and organelles to provide new pharmacological insights uniquely possible through interactome studies.

Presented By:

Jim Bruce, University of Washington, Washington, United States

---

Developing Cross-linking Mass Spectrometry for Structural Analysis of Protein Complexes

Presented By:

Lan Huang, University of California Irvine, California, United States

---

O13.01 | Cleavable Biotin-Enabled Proximity Labeling Proteomics and Peptidomics

Shiori Sekine (1), Haorong Li (2), Ashley Frankenfield (2), Ling Hao (2)

(1) School of Medicine, University of Pittsburgh., Pittsburgh, PA, USA

(2) Department of Chemistry, George Washington University, Washington, DC, USA

Proximity labeling (PL)-based proteomic techniques offer unique ways to capture both stable and transient protein networks and molecular microenvironment surround a target protein in living cells. However, a number of technical challenges still exist, such as interferences from highly abundant streptavidin signals and endogenously biotinylated proteins, how to select the proper controls to minimize false discoveries, and experimental variations among biological/technical replicates. In our recent study, we developed and systematically optimized an endogenous lysosomal PL probe to capture the proteins in the lysosomal microenvironment in human neurons (Analytical Chemistry, Nov. 2020). Although bioactive peptides also represent major components of the molecular microenvironment, it is technically challenging to capture bioactive peptides with PL-MS method. Therefore in this study, we developed a new cleavable biotin-enabled PL-MS technique that solved major technical hurdles in the field and enabled both proteins and bioactive peptides to be captured in living cells.

We first established the labeling conditions and LC-MS/MS method in vitro using NHS-cleavable biotin reagents to label BSA protein standard and WT human cell lysate. We then used cleavable biotin as the substrate for mitochondrial TurboID cell line to capture proteins and endogenous peptides in the mitochondrial matrix in vivo. With our cleavable biotin-enabled PL-MS method, biotinylated proteins can be cleaved off streptavidin-coated beads after enrichment without the need for on-beads digestion. Overwhelming streptavidin signals and interference endogenously biotinylated proteins can therefore be removed from the sample, achieving greatly improved sensitivity and specificity. Our method also revealed specific biotinylation sites with better specificity and reduced cost compared to previously published methods using biotin antibodies. Protein and peptide fractions can be enriched separately from both cell lysate and secreted exosomes, achieving a comprehensive profile of molecular microenvironment with PL-MS.
Contact: Ling Hao, linghao@gwu.edu

O13.02 | Identification of ligand-dependent protein interaction networks of the mu-opioid receptor with temporal and spatial resolution

Ruth Hüttenhain (1), Braden T Lobingier (2), Benjamin J Polacco (1), Qiongyu Li (1), Emily Blythe (3), Nevan J Krogan (1), Mark Von Zastrow (4)

(1) Cellular Molecular Pharmacology, University of California San Francisco, San Francisco, CA, USA

(2) Chemical Physiology and Biochemistry, Oregon Health & Science University, Portland, OR, USA

(3) Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA

(4) University of California San Francisco, University of California San Francisco, San Francisco, CA, USA

G protein-coupled receptors (GPCRs) represent the largest family of signaling receptors and drug targets. Following ligand-induced activation of GPCRs signal transduction is mediated by protein interaction networks operating on short timescales and across multiple cellular locations. While temporal dynamics of protein interactions have been previously characterized, a major challenge remains largely unmet: how to interrogate the protein interaction networks engaged by GPCRs while capturing both their spatial and temporal context.

Previously, we have developed an approach combining APEX-based proximity labeling with quantitative proteomics and a system of spatial references, which delivers, with sub-minute temporal resolution, protein interaction networks and subcellular location of the receptor. We not only validated capture of proteins known to interact with the receptors, including those with transient or low affinity interactions, but demonstrated that our approach can be used to discover new network components regulating receptor function.

We recently extended the approach to understand how chemically distinct agonists targeting the mu-type opioid receptor (MOR) produce different receptor based effects. Specifically, we developed a novel data analysis strategy, which allowed data-driven prediction of the subcellular location of MOR and determination of the protein interaction networks engaged by the receptor after stimulation with full, partial, and G protein biased agonists. The data revealed distinct intracellular trafficking comparing the agonists, while we observed rapid endocytosis for the full, unbiased agonist, the G protein biased agonist did not provoke any endocytosis. Interestingly, we discovered EYA4 and KCTD12 as interactors of MOR, which are common for all three agonists and have not been previously linked with MOR. We validated that the recruitment of both EYA4 and KCTD12 to the activated receptor is dependent on G protein signaling, indicating that these two novel interactors might play a role in regulating G protein signaling downstream of MOR activation.

Contact: Ruth Huttenhain, ruth.huttenhain@ucsf.edu

Parallel Session 14: Stem Cell Proteomics and Chemical Proteomics in Drug Discovery

Tissue-Down Cell Surface Proteomics for Benchmarking Stem Cell-Derived Cardiomyocytes

Accurate phenotyping of stem cell-derived cardiomyocytes is necessary to promote their use in drug screening, disease modeling, and regenerative medicine. However, benchmarking stem cell derivatives to their adult counterparts is hampered by the lack of cell type resolved views of the human heart. Here, we describe an advanced analytical platform, CellSurfer, that enables sensitive cell surface profiling of adult human cardiac cells. We applied CellSurfer to develop the first quantitative map of cell surface glycoproteins in cardiomyocytes isolated from different anatomical regions of the human heart. Emerging from these data are new markers for benchmarking stem cell derived cardiomyocytes to their relevant human counterparts and new subtype-specific adult cell phenotypes.

Presented By:

Rebekah Gundry, University of Nebraska Medical Center, Nebraska, United States

---

Activity Mapping from Single Cells to Whole Organisms in Cancer

Biological systems are inherently and profoundly heterogeneous, both at the molecular level (e.g. encoded proteins existing in distinct posttranslational modification states and macromolecular complexes) and the cellular level (e.g. intra- and intercellular localization of biomolecules). Despite growing awareness and appreciation for this level of molecular complexity in living systems, most studies still rely on reductionist strategies to interrogate the proteome, owing to the significant technical challenges associated with studying protein structure, function and organization in native environments. We believe that in order to understand molecular information flow under basal or diseased conditions we must be able to probe biomolecular function and organization in native environments across scales of space and time. Therefore, innovation in the development probes and technology platforms is needed. In this talk I will describe the development of new chemical probes and complementary proteomic platforms that enable quantitative detection of protein functional states on length scales ranging from single cells to live animals. I will discuss the integration of these platforms for the discovery, high-throughput profiling and diagnostic detection of enzyme activities associated with tumor progression and metastasis. Beyond cancer, this talk will emphasize the integration of chemical proteomic platforms as a discovery engine to identify novel targets for diagnostic and therapeutic development in human disease.

Presented By:

Ray Moellering, University of Chicago, Illinois, United States

---

O14.01 | High-throughput elucidation method for small molecule - protein interactions

Ziwei Zhang (1), Xiuyuan Ma (1), Yu Gao (1)

(1) College of Pharmacy, University of Chicago, Chicago, IL, US

Identification of target proteins of any small molecule is a crucial step in new drug development. The protein binding profile of a small molecule not only defines the specificity and effectiveness of a drug but also strongly correlates with its toxicity and side effects. Conventional methods that test the binding affinity between a small molecule and individual purified protein could only detect on-target effects, without the ability to identify off-target

effects. Here we present a study that tries to elucidate small molecule binding proteins on a proteome scale. We use an energetics-based protein separation method coupled with tandem MS/MS to rapidly and unbiasedly identify the interacting proteins for a given compound.

This approach is based on the hypothesis that a protein bound by a given drug would see its thermal stability changed upon binding. If protein could bind to small molecule spontaneously, we can infer an entropy decrease and a more stable energy state of the new complex. By step-wise adding denature solution, cell lysate is gradually precipitated out and collected for MS analysis. Each protein in cell lysate will have an intensity distribution in different precipitates. The distribution reflects the stability of protein and will be altered upon binding to a small molecule. This enables us to simultaneously assay thousands of proteins stability changes in cell lysate. The protein stability changes are calculated and ranked with a python script to select potential protein targets. These potential targets could be further validated by known protein-protein interactions to find highly confident unknown targets.

Our results show that this method is easy to handle and could be validated by known small molecule protein interactions, which makes it a powerful and high-throughput tool for large-scale target identification.

Contact: Ziwei Zhang, zzhan60@uic.edu

O14.02 | Quantitative mass spectrometry-based global proteome and phosphoproteome analyses of thymic epithelial tumors (TET) reveal novel drug targets

Xu Zhang (1), Ting Huang (2), Fatos Kirkali (1), Yue Qi (1), Tapan Maity (1), Khoa Dang Nguyen (1), David Schrump (3), Olga Vitek (2), Arun Rajan (1), Udayan Guha (4)

(1) Center for Cancer Research, NCI/NIH, Bethesda, MD, USA

(2) Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA

(3) Thoracic Surgery Branch, Center for Cancer Research, NCI/NIH, Bethesda, MD, USA

(4) Center for Cancer Research, NCI/NIH; Bristol-Myers Squibb, Lawrence Township, NJ, USA

Thymic epithelial tumors (TETs, which consist of thymoma and thymic carcinoma) are rare epithelial tumors of the thymus having amongst the lowest mutational burden of all cancers. There are limited treatment options for patients with advanced disease. The complexity and rarity of the disease hampers the development of better therapeutics. Here, we report large-scale identification and quantitation of protein and phosphorylation abundance by proteomic approach and reveal the differences between thymomas and thymic carcinomas. Three-state SILAC quantitative mass spectrometry was employed to quantify the global proteome and phosphoproteome of the two thymoma cell lines (IU-TAB1 and T1682), and three thymic carcinoma cell lines (MP57, T1889 and Ty82). TMT quantitative proteomics was used to quantify the global proteome and phosphoproteome of 54 TETs from 26 patients. Basic RPLC separation was performed on the first dimension; 96 fractions were collected and concatenated into 12 fractions, followed by TiO2 enrichment before LC-MS/MS analysis. We identified 4756 proteins/5690 phosphosites from TET cell lines, and 8320 proteins/17716 phosphosites from TET patients. Hierarchical clustering of both SILAC and TMT ratios demonstrated that thymomas and thymic tumors clustered separately. Clusters based on proteins and phosphosites were slightly different. For patient samples, different locations of tumors from the same patient were grouped together. Several protein groups, such as MCM protein complex, solute carrier family, and collagens had different abundance between thymoma and thymic carcinoma. The metabolic enzyme, GSTP1 had increased abundance in a subset of TET cell lines and patients. Three GSTP1 inhibitors were tested in cell survival analysis on TET cell lines. Furthermore, Cibersort analysis identified significant differences in immune cellular protein signatures of the tumor microenvironment between thymoma and thymic patients. T cell receptor signaling pathway was the top pathway enriched among the significantly decreased proteins of the immune cells in thymic patients.

Contact: Xu Zhang, xu.zhang@nih.gov

These awards are proudly supported by:

TA01.01

Top-Down Nanoproteomics Enables Comprehensive Analysis of Low-Abundance Cardiac Troponin I Proteoforms from Human Serum

David S. Roberts (1), Timothy N. Tiambeng (1), Kyle A. Brown (1), Yanlong Zhu (2), Bifan Chen (3), Zhijie Wu (4), Stanford D. Mitchell (5), Tania M. Guardado-Alvarez (1), Song Jin (1), Ying Ge (2)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

(2) Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, USA

(3) Genentech Inc., South San Francisco, CA, USA

(4) Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA

(5) Molecular and Cellular Pharmacology Program, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI, USA

Cardiovascular disease (CVD) is the leading cause of mortality worldwide. Cardiac troponin I (cTnI) is widely utilized as a 'gold-standard' biomarker for CVDs, especially acute myocardial infarction (AMI). cTnI is released into the bloodstream following cardiac injury where it circulates with low abundance (typically < 50 ng/mL) and in myriad protein forms, "proteoforms" (e.g., phosphorylated, acetylated, etc.), making detection and analysis extremely challenging. Recent studies have shown that cTnI is heavily modified and its proteoforms arising from various post-translational modifications (PTMs) can provide new insights to the molecular mechanisms underlying cardiovascular diseases. The PTM profiles of cTnI can function as molecular fingerprints of cellular signaling pathway activity with the potential to serve as the next-generation biomarkers. Therefore, a comprehensive proteoform-resolved cTnI analysis that can detect cTnI in blood with detailed PTM information is urgently needed. However, due to the extraordinary dynamic range (~1012) of the blood proteome in the presence of high-abundance blood proteins, accurate detection and analysis of circulating cTnI remains an unsolved challenge. Here we developed a novel nanoproteomics strategy to specifically capture and enrich cTnI using functionalized nanoparticles (NPs) and subsequently employ top-down mass spectrometry (MS) to identify and quantify various cTnI proteoforms in human serum. We demonstrate that carefully designed peptide-functionalized NPs can directly capture and enrich cTnI from human serum with high specificity and reproducibility, while simultaneously depleting highly abundant blood proteins, such as human serum albumin (HSA). These NPs not only outperform conventional monoclonal antibody platforms for serum cTnI enrichment, but also faithfully and holistically preserve all endogenous cTnI proteoforms. This antibody-free approach can be potentially leveraged in future clinical cTnI diagnostic assays. By further applying to a large human cohort, patient blood samples can be analyzed to comprehensively detect all cTnI proteoforms and establish the relationships between cTnI proteoforms and underlying disease etiology.

Contact: David Roberts, dsroberts@wisc.edu

TA01.02

Optimization of Pervanadate BOOST Channel in TMT Multiplexing Delivers Deeper Quantitative Characterization of the Tyrosine Phosphoproteome

Xien Yu Chua (1), Ricky Edmondson (2), Arthur Salomon (3)

(1) Department of Molecular Pharmacology, Physiology, and Biotechnology, Brown University, Providence, RI, United States

(2) Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, AR, United States

(3) Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI, United States

The activation of T cell signaling pathways, regulated predominantly by tyrosine phosphorylation, is fundamental to the functions of the adaptive immune system. A comprehensive examination of the tyrosine phosphoproteome using high-throughput proteomics will expand our understanding in the signaling network in T cells. To overcome challenges in the detection and quantitation of low-abundance phosphotyrosine(pTyr)-containing peptides in mass spectrometry, we recently developed the Broad-spectrum Optimization Of Selective Triggering (BOOST) method (PMID: 32071147) to deliver deeper quantitative characterization of the tyrosine phosphoproteome. We leveraged the multiplexing capability of tandem mass tags (TMT) and the use of pervanadate (PV) boost channel (cells treated with the broad-spectrum tyrosine phosphatase inhibitor PV) to selectively increase the relative abundance of pTyr-containing peptides. PV boost channel facilitates selective fragmentation of pTyr-containing peptides to deliver accurate TMT reporter ion quantitation of the sample channels. As a result, pTyr peptide identification and quantitation were vastly improved. Since then, we have continued to expand the capabilities and robustness of this methodology, including the optimization of sample preparation, TMT channel arrangement and data acquisition parameters. We report that this method can be applied to understand fundamental biological questions about pTyr signaling pathways, revealing the basis of thymic selection threshold. We also validated this approach beyond cell lines using scarce mice primary T cells. Overall, we provide an update to this strategy, which can potentially be applied in analyses of other post-translational modifications where treatments that broadly elevate the levels of those modifications across the proteome are available.

Contact: Xien Yu Chua, xien_yu_chua@brown.edu

TA01.03

caAtlas: An Immunopeptidome Atlas of Human Cancer

Xinpei Yi (1), Yuxing Liao (1), Bo Wen (1), Kai Li (1), Yongchao Dou (1), Sara Savage (1), Bing Zhang (1)

(1) Baylor College of Medicine, Houston, Texas, USA

At the core of immunotherapy is the T cell recognition of tumor antigens. Thus, comprehensive characterization of tumor antigens is essential for the design of effective and safe cancer immunotherapy strategies. Mass spectrometry (MS)-based immunopeptidomics allows high-throughput, direct identification of major histocompatibility complex (MHC)-bound peptides in vivo. Here, we collected 42 published human immunopeptidomic datasets, which include 327 cancer samples covering nine cancer types and 216 non-cancerous samples. A systematic, quality-controlled analysis of all 51.6 million MS/MS spectra using an open search algorithm identified 364,283 peptides, including 11,647 modified peptides associated with 80 types of modifications. Unmodified peptides showed high levels of predicted binding affinities by NetMHCpan; however, the predicted binding affinities for the modified peptides were much lower, suggesting a deficiency of existing computational tools in making predictions for modified peptides. Analysis of the 416 phosphorylated and 1,239 acetylated HLA class I peptides revealed distinct sequence motifs, and the pattern for the phosphopeptides was consistent with previous reports. Further interrogation of all identified peptides revealed evidence for tumor specific-presentation of hundreds of known and novel cancer-testis antigens as well as tumor associated antigens. We also found non-tumor-specific presentation of 44 previously annotated cancer-testis antigens. Applying NeoQuery, a newly developed computational pipeline, to our dataset identified immunopeptidomic evidence for more than 1,000 somatic mutation-derived neoantigens, and a subset of them were further validated through a traditional MS/MS peptide identification pipeline. We make all these data together with annotated MS/MS spectra supporting each individual antigen identification in a publicly accessible, easily browsable web portal named cancer antigen atlas (caAtlas, http://www.zhang-lab.org/caatlas/). caAtlas provides a fundamental resource for the selection and prioritization of MHC-bound peptides for immunogenicity testing and cancer immunotherapy development.

Contact: Xinpei Yi, Xinpei.Yi@bcm.edu

TA01.04

A high dimensional map of the budding yeast phosphorylation-dependent signaling network

Mario Leutert (1), Noelle Fukuda (1), Judit Villen (1)

(1) Genome Sciences, University of Washington, Seattle, WA, USA

The budding yeast Saccharomyces cerevisiae is among the best-studied experimental organisms, however a quantitative understanding of the phosphorylation dependent control of its signaling network is missing. Most phosphoproteomic studies in biological model systems are low dimensional (<5 conditions) due to limitations in reproducibility and throughput of sample preparation and mass spectrometric measurement workflows. Much higher dimensionality is required to obtain insight into architecture and usage of the signaling network. We have addressed these challenges by introducing rapid-robotic phosphoproteomics (R2-P2), an end-to-end automated method that uses magnetic particles to process cell lysates to deliver mass spectrometry-ready phosphopeptides in a 96-well format. We have shown that R2-P2 in combination with data-independent acquisition (DIA) mass spectrometry achieves reproducible and deep quantification of the phosphoproteome. We now employed the R2-P2-DIA-MS workflow to gather the most comprehensive quantitative phosphoproteomic perturbation map in yeast to date. To achieve high dimensionality, we systematically exposed yeast cells to a wide range of ~100 distinct environmental and chemical perturbations that touch many aspects of yeast

biology. The enormous perturbation-phosphorylation map facilitates identification of functional phosphorylation events, sub-network usage, pathway crosstalk and kinase-substrate predictions. This resource will eventually allow integration of the descriptive signaling mechanism with the predicted cellular functions obtained by the multitude of other large-scale assays performed in yeast.

Contact: Mario Leutert, mleutert@uw.edu

Advances in Technology

P01.01 Analysis of Human Cell Digests Using Fast Microflow SWATH® Acquisition and Cloud Computing
Alexandra Antonoplis, Christie Hunter

P01.02 TurnoveR: A Skyline External Tool for the Analysis of Protein Turnover from Metabolic Labeling Studies
Nathan Basisty, Cameron Wehrfritz, Nicholas Shulman, Alexandra Marsh, Michael MacCoss, Brendan MacLean, Birgit Schilling

P01.03 Robust Plasma Protein Profiling Workflow for Clinical Research Using a UHPLC and a Modified Orbitrap Mass Spectrometer
Debadeep Bhattacharyya, David Sarracino, Kerry Hassell

P01.04 Optimization of a long-gradient ultrahigh pressure liquid chromatography (UPLC) separation for HDX-MS
Mulin Fang, Zhe Wang, Kellye A. Cupp-Sutton, Thomas Welborn, Kenneth Smith, Si Wu

P01.05 Proteograph's multi nanoparticle proteins coronas enable deep plasma proteomics studies at scale with unmatched sensitivity in combination with trapped ion mobility
Shadi Ferdosi, Tristan Brown, Patrick Everley, Mike Figa, Matthew McLean, Eltaher Elgierari, Xiaoyan Zhao, Veder Garcia, Tianyu Wang, Kateryna Riedesel, Jessica Chu, Martin Goldberg, Asim Siddiqui, Juan Cruz Cuevas, John Blume, Daniel Hornburg, Omid Farokhzad

P01.06 EasyPep Sample Preparation Technology for Rapid and Efficient Mass Spectrometry-based Proteomics
Amarjeet Flora, Sergei Snovida, Penny Jensen, Bhavin Patel, Ryan Bomgarden

P01.07 A Proteomic sample preparation for mass spectrometry using an automated workstation
Qin Fu, Johnson Casey, Danica-Mae Manalo, Jennifer Van Eyk

P01.08 Computational simulation improves PASEF ion distribution and diaPASEF-based proteomics
Huanhuan Gao, Youqi Liu, Siqi Zheng, Xianming Liu, Xiaoxian Du, Nagarjuna Nagaraj, Weigang Ge, Tiejun Li, Yi Zhu, Tiannan Guo

P01.10 A high-throughput automated sample preparation workflow for LC-MS-based protein analysis: SP3+iST- On the Tecan Freedom EVO liquid handling platform
Katrin Hartinger, Claudia Martelli, Katarzyna Buczak, Alexander Schmidt, Manuel Bauer, Nils A. Kulak

P01.11 Spray-capillary Based Capillary Electrophoresis Mass Spectrometry for Metabolite Analysis in Single Cells
Lushuang Huang, Mulin Fang, Kellye Cupp-Sutton, Zhe Wang, Kenneth Smith

P01.12 Photocleavable Surfactant-Enabled Extracellular Matrix Proteomics
Samantha Knott, Kyle Brown, Harini Josyer, David Inman, Suzanne Ponik, Andreas Friedl, Ying Ge

P01.13 Comprehensive analysis of RBM20 modifications by middle-down and bottom-up high-resolution mass spectrometry
Eli Larson, Yutong Jin, Yanghai Zhang, Wei Guo, Ying Ge

P01.14 Automation of Cell and Tissue Sample Preparation Workflow for High Throughput Proteomics
Danica-Mae Manalo, Vineet Vaibhav, Andrea Matlock, Vidya Venkatraman, Jennifer Van Eyk

P01.01

Analysis of Human Cell Digests Using Fast Microflow SWATH® Acquisition and Cloud Computing

Alexandra Antonoplis (1), Christie Hunter (1)

(1) SCIEX, Technical Marketing, Redwood City, CA, USA

New data processing solutions are needed to match the speed and scale at which SWATH Acquisition results are collected. In this study, the OneOmics™ Suite, a cloud-based full proteomics solution, was used to quantitate and compare protein expression differences across six human cell digests. The cell digest proteins were also quantitated using two distinct microflow LC gradients, a 20 min gradient and a 45 min gradient, to evaluate impacts of shortening the LC method.

Methods: Whole cell lysates (HEK293, Hela, MCF-7, A549 and MG132-treated Hela and MG132-treated HEK293) were digested with trypsin and loaded at 3-5 µg total protein per injection. A NanoLC™ 425 system was used for microflow chromatography at two gradient lengths. SWATH Acquisition data was collected using the OptiFlow® Ion Source on the TripleTOF® 6600+ System. All SWATH Acquisition Data was processed in the SCIEX Cloud Platform using the OneOmics Suite.

Results: Shortening the microflow LC gradient from 45 to 20 min resulted in similar numbers of proteins quantified per digest at 1% FDR and 20% CV and no significant reduction in quantitative accuracy. More peptides per protein at 1% FDR and 20% CV were quantified with the 45 min gradient as expected. Using the Browser and Biological Review tools in the OneOmics Suite, expression profiles of individual proteins could be compared across the cell lines, along with enrichment of ontologies and biological pathways.

The OneOmics Suite enabled fast processing and results interrogation of SWATH Acquisition data collected for six distinct cell digests at two microflow gradient lengths. The numbers of proteins quantified were very similar between the 20 and 45 min gradients, with no significant degradation of quantitative accuracy.
Contact: Alexandra Antonoplis, alexandra.antonoplis@sciex.com

P01.02

TurnoveR: A Skyline External Tool for the Analysis of Protein Turnover from Metabolic Labeling Studies

Nathan Basisty (1), Cameron Wehrfritz (1), Nicholas Shulman (2), Alexandra Marsh (2), Michael MacCoss (2), Brendan MacLean (2), Birgit Schilling (1)

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

(2) Department of Genome Sciences, University of Washington, Seattle, WA, United States

Loss of protein homeostasis is a hallmark of aging and age-related diseases, including neurodegeneration, sarcopenia, and diabetes. While disruption of protein turnover machinery, such as autophagy and the ubiquitin-proteasome system, are often associated with aging and associated pathologies, the turnover rates of proteins do not necessarily reflect a reduction of these processes. Therefore, methods to measure the turnover rates of proteins directly, rather than surrogate measurements of translation and degradation machinery, are critically needed to accurately examine the stability of the proteome during aging and disease processes. Additionally, while the measurement of protein turnover is relevant in many biological settings, including aging, conducting a protein turnover study remains very computationally complex and difficult for most scientists. The development of versatile computational tools on widely accessible, open-source platforms is needed to make this approach more user-friendly.

Here, we have developed a new computational tool – TurnoveR – for the accurate calculation of protein turnover rates from mass spectrometry analysis of metabolic labeling experiments in Skyline, a free and open-source proteomics software platform. Using data generated from metabolic labeling of mice with heavy leucine, we demonstrate how this tool enables the calculation of protein turnover rates entirely within a Skyline workspace using raw data acquired on multiple mass spectrometric platforms: TripleTOF 6600 and Orbitrap Velos. We re-analyze previously published data in calorie restricted and ad libitum-fed mice to show this approach re-capitulates turnover rates and differential changes in turnover rates between treatment groups calculated in previous studies. We hope that the addition of this external tool to the widely used Skyline proteomics software will facilitate wider utilization of metabolic labeling and protein turnover analysis in highly relevant biological models, including aging, neurodegeneration, and skeletal muscle atrophy.

Acknowledgements: Nathan Shock Pilot Award (UW, Schilling), R01 AR071762 (NIH, Adams), and K99 AG065484 (NIH, Basisty).

Contact: Nathan Basisty, nbasisty@buckinstitute.org

P01.03

Robust Plasma Protein Profiling Workflow for Clinical Research Using a UHPLC and a Modified Orbitrap Mass Spectrometer

Debadeep Bhattacharyya (1), David Sarracino (1), Kerry Hassell (1)

(1) ThermoFisher Scientific, Boston, MA, USA

Profiling the human plasma proteome using liquid chromatography (LC) coupled to mass spectrometry (MS) has been a sought-after application for routine monitoring of health. While there are 70 FDA cleared tests for analysis of proteins, implementation of an LC-MS based assays has been limited owing to the demand for high analytical dynamic range. Recent technology platforms comprising Ultra High-Performance Liquid Chromatography (UHPLC) coupled with Orbitrap mass spectrometer (HRAM) has revolutionized the depth at which this proteome can be identified and quantified.
In this study, we present LC-MS/MS data of the plasma samples. A spectral library was created by offline fractionation of a pool of digested plasma samples into 24 fractions. The combined libraries contain complete information of more than three hundred protein groups and more than two thousand peptides. More than one hundred proteins evaluated in the experiments demonstrated a combined CV of 10% or less. The plasma samples differ in platelet, white blood cell and red blood cell numbers, and associated proteins from these contaminants are well characterized and the data obtained in this study can also enable biotherapeutic characterization processes, and other high dynamic range applications that do not experience sample limitation issues.

Contact: Debadeep Bhattacharyya, debadeep.bhattacharyya@thermofisher.com

P01.04

Optimization of a long-gradient ultrahigh pressure liquid chromatography (UPLC) separation for HDX-MS

Mulin Fang (1), Zhe Wang (1), Kellye A. Cupp-Sutton (1), Thomas Welborn (1), Kenneth Smith (2), Si Wu (1)

(1) Department of Chemistry and Biochemistry, University of Oklahoma, Norman, OK, USA

(2) Department of Arthritis and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

Hydrogen deuterium exchange coupled with mass spectrometry (HDX-MS) is a powerful protein foot-printing technique for the characterization of protein dynamics and protein interactions. Currently, due to the advances in liquid chromatography (LC)-MS technology and the availability of improved data analysis software, HDX-MS is expanding to examine protein interactions in complex matrices such as multi-protein complexes. However, throughput is still the key technical bottleneck for the application of HDX-MS in the analysis of complex biological samples such as cell lysates. The difficulty associated with complex sample analysis using HDX-MS stems from the co-elution of peptides resulting in increased peak complexity after deuterium labeling. Here, we evaluated and optimized ultrahigh-pressure liquid chromatography (UPLC) separation conditions with respect to mobile phase modifiers, flow rates and pressures, and gradient lengths at -10 °C for the HDX-MS analysis of E. coli cell lysate digest. We demonstrated that, under the optimized conditions, a long-gradient separation (e.g., 90 minutes) at -10 °C was capable of characterizing 1419 deuterated peptides from 320 proteins, which was a great improvement when compared with a short-gradient separation (e.g., 15 minutes). Additionally, our results showed that the deuterated peptides eluted throughout the long gradient separation maintain a high level of fractional deuterium incorporation at -10 °C. High deuterium retention during subzero-temperature long-gradient UPLC separation makes this a promising approach for the high-throughput HDX-MS analysis of complex biological samples. Overall, our study demonstrated that the optimized, sub-zero, long-gradient UPLC separation enabled the characterization of thousands of peptides from hundreds of proteins in a single HDX-MS analysis. As a result, this technique holds great promise for characterizing protein interactions in samples with high complexity and in more native-like environments such as cell lysates.

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

P01.05

Proteograph's multi nanoparticle proteins coronas enable deep plasma proteomics studies at scale with unmatched sensitivity in combination with trapped ion mobility

Shadi Ferdosi (1), Tristan Brown (1), Patrick Everley (1), Mike Figa (1), Matthew McLean (1), Eltaher Elgierari (1), Xiaoyan Zhao (1), Veder Garcia (1), Tianyu Wang (1), Kateryna Riedesel (1), Jessica Chu (1), Martin Goldberg (1), Asim Siddiqui (1), Juan Cruz Cuevas (1), John Blume (1), Daniel Hornburg (1), Omid Farokhzad (1)

(1) Seer Inc., Redwood City, CA, USA

Blood plasma is a rich, readily available source of proteins that is commonly used in clinical profiling studies. However, proteome research is inherently constrained by the large dynamic concentration range and complexity of proteins within plasma. The ability to overcome these constraints while interrogating deeply and broadly into the plasma proteome has only been partially addressed by laborious, unscalable, and low throughput workflows. To fully enable high-throughput plasma proteomics, we recently have developed a quantitative profiling platform called Proteograph that consists of a panel of five nanoparticles (NPs) with distinct physicochemical properties. This panel of NPs is used in parallel to provide optimized identification of plasma proteins in terms of depth and breadth with precise and reproducible quantification. Here we explore the synergy of the Proteograph assay with the TIMS-TOF pro mass spectrometer (MS). Using this MS, we investigated short and long LC gradients ranging from 7 to 90 min using both data-dependent- and data-independent-acquisition (i.e., DDA and DIA) for the five distinct NP-plasma protein corona evaluating depth, dynamic range, coverage, throughput, and precision of Proteograph™ proteome profiling. We also evaluated linearity of response employing a multi-level proteome spike-in experiment. The high efficiency of ion-beam sampling enabled by the trapped ion mobility and parallel accumulation-serial fragmentation of the TIMS-TOF increased sensitivity by up to 10x compared to alternative MS instrumentation. For DDA, we observed a 20% gain in protein identification as compared to conventional MS acquisition strategies while using only 25% of the injection mass (160 ng). The precision of protein and peptide quantification was improved by 2x (CV of 21% vs 11%). In DIA, we observed 40% higher identification with comparable precision. In summary, Proteograph in combination with the TIMS-TOF pro provides a high-performance combination for rapid deep, precise, and accurate proteome profiling for biomedical research and biomarker discovery.

Contact: Shadi Ferdosi, sferdosi@seer.bio

P01.06

EasyPep Sample Preparation Technology for Rapid and Efficient Mass Spectrometry-based Proteomics

Amarjeet Flora (1), Sergei Snovida (1), Penny Jensen (1), Bhavin Patel (1), Ryan Bomgarden (1)

(1) Thermo Fisher Scientific, Rockford, IL, USA

Introduction Advances in mass spectrometry (MS) instrumentation have enabled routine analysis of complex protein samples. However, protein sample preparation for mass spectrometry still largely lacks cohesive standardization, which generally leads to inconsistent and irreproducible analyses. Recently, we expanded our EasyPepTM sample preparation technology by introducing two new formats – a large-scale (Maxi) sample preparation kit to streamline the analysis of protein post-translational modifications and a 96-well plate format kit to enable high throughput and automation. Here, we describe sample type-specific examples to highlight the unique features of each format for different sample types and

applications. Methods Several cell and tissue types, including FFPE and plasma samples, were processed using our standardized EasyPep sample preparation procedures according to the format-specific procedures for both label-free and Tandem Mass TagTM (TMTTM) reagent labeling experiments. Peptide yields were assessed using PierceTM Quantitative Colorimetric Peptide Assay prior to LC-MS analysis using Thermo Scientific™ Q-Exactive™ Plus hybrid quadrupole-Orbitrap™ mass spectrometer.

Results Our optimized chemistry enabled efficient and reproducible processing of cultured mammalian cells and tissues significantly reduced both hands-on and total sample processing time. We were able to routinely obtain 10-20% higher protein and peptide identifications with lower missed cleavages than other common workflows. In addition, we demonstrate that EasyPep sample prep chemistry and cleanup is fully compatible with TMT or TMTpro reagent multiplexing for processing and quantitative analysis of large numbers of samples within a few days. Three operation modes including centrifugation, vacuum, and positive pressure were evaluated and optimized for 96-well EasyPep sample preparation plate. Both vacuum and centrifugation modes were evaluated for the larger Maxi column format. We have also adapted EasyPep 96 sample preparation protocol to a fully automated workflow using a HamiltonTM STARletTM automated liquid handling system. Conclusion: Optimized workflow for quick and efficient sample preparation for discovery and quantitative proteomics.

Contact: Amarjeet Flora, amarjeet.flora@thermofisher.com

P01.07

A Proteomic sample preparation for mass spectrometry using an automated workstation

Qin Fu (1), Johnson Casey (2), Danica-Mae Manalo (1), Jennifer Van Eyk (2)

(1) Heart Institute, LA, CA, USA

(2) Heart Institute/Precision Bio. Labs, LA, CA, USA

Sample preparation serves as the foundation for any proteomic analysis. Prior to MS analysis, the proteins in a biological sample are typically denatured, reduced, alkylated, digested into tryptic peptides, and desalted. Each of these steps requires optimization, and the entire process is traditionally performed manually, allowing for the introduction of analytical errors. We have developed an automated workstation workflow to perform all necessary plasma sample preparation procedures including desalting step. Here, we report the development of a more efficient and better controlled workflow with the following advantages: 1) The number of liquid transfer steps is reduced from nine to six by combining reagents; 2) Pipetting time is reduced by selective tip pipetting using a 96-position pipetting head with multiple channels; 3) Potential throughput is increased by the availability of up to 45 deck positions; 4) Complete enclosure of the system provides improved temperature and environmental control and reduces the potential for contamination of samples or reagents; and 5) The addition of stable isotope labeled peptides, as well as β-galactosidase protein, to each sample makes monitoring and quality control possible throughout the entire process. These hardware and process improvements provide good reproducibility and improve intra-assay and inter-assay precision (CV of less than 20%) for LC-MS based protein and peptide quantification. The entire workflow for digesting 96 samples in a 96-well plate can be completed in approximately 5 hours. The automated workflow presented here provides for consistent enzymatic digestion with improved reproducibility and throughput compared to manual methods.

Contact: Qin Fu, qin.fu@cshs.org

P01.08

Computational simulation improves PASEF ion distribution and diaPASEF-based proteomics

Huanhuan Gao (1), Youqi Liu (2), Siqi Zheng (3), Xianming Liu (4), Xiaoxian Du (4), Nagarjuna Nagaraj (5), Weigang Ge (2), Tiejun Li (6), Yi Zhu (1), Tiannan Guo (1)

(1) 1. Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. 2. Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China 3. Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China, Hangzhou, Zhejiang, China

(2) 4. Westlake Omics (Hangzhou) Biotechnology Co.,Ltd. No.1, Yunmeng Road, Cloud Town, Xihu District, Hangzhou 310024, Zhejiang Province, China., Hangzhou, Zhejiang, China

(3) 5. School of mathematics science, Peking University, Summer Palace Road 5th, 100871, Beijing, China., Beijing, Beijing, China

(4) 6. Bruker Daltonics, Shanghai, China., Shanghai, Shanghai, China

(5) 7. Bruker Daltonics GmbH, Breme, Germany., Breme, Breme, Germany

(6) 5. School of mathematics science, Peking University, Beijing, Beijing, China

Introduction: Data independent acquisition (DIA) coupled with PASEF, i.e. diaPASEF, is emerging as a robust methodology for high-throughput quantitative proteomic analysis (1). The default diaPASEF method adopts a series of gas phase isolation window of fixed width along the m/z space and ion mobility space. Interaction between m/z and ion mobility of peptide precursors has not been investigated. Methods: We generated DDA data using PASEF in a timsTOFPro instrument from various sample types including Hela cell line and clinical tissues using LC with gradient lengths ranging from 30min, 45min, 90 min, to 120 min. We analyzed the interaction between m/z and ion mobility of peptide precursors, and built computational models to optimize the PASEF to improve the diaPASEF. Results: We analyzed the interaction between the m/z and ion mobility of the peptide precursors and divided all the MS1 ions into different isolation windows with equally distributed complexity in each isolation window using multiple computational modeling methods as will presented in the presentation. Overall, this strategy achieved an increase of 10%-30% of peptide and protein identifications comparing to the conventional diaPASEF method. The applicability of this improved method was further validated in clinical specimens. Conclusions: We optimized the PASEF windows using computational modelling which improved the proteomic depth of diaPASEF.

(1) Meier, F., et al. Nat Methods 17, 1229–1236 (2020).

Contact: Huanhuan Gao, gaohuanhuan@westlake.edu.cn

P01.10

A high-throughput automated sample preparation workflow for LC-MS-based protein analysis: SP3+iST- On the Tecan Freedom EVO liquid handling platform

Katrin Hartinger (1), Claudia Martelli (2), Katarzyna Buczak (3), Alexander Schmidt (3), Manuel Bauer (2), Nils A. Kulak (1)

(1) PreOmics GmbH, Martinsried/Planegg, Bavaria, Germany

(2) Tecan, Männendorf, Zurich, Switzerland

(3) Biozentrum, University of Basel, Basel, Basel, Switzerland

Liquid chromatography-mass spectrometry (LC-MS) has emerged as a powerful analytical tool for protein identification and quantification. The arise of simplified, commercially available sample preparation solutions with continuously evolving LC-MS instrumentation facilitates fast, reproducible and highly sensitive LC-MS analysis while making the technique also accessible for non-expert users. However, LC-MS sample preparation is largely performed manually and thus, often remains the critical step in terms of time and robustness, particularly for large sample cohorts. Aiming for a high-throughput automation solution for standardized sample preparation, we developed an automated iST-based workflow using the liquid handling platform Freedom EVO® 100 as well as the positive pressure module Resolvex® A200 from Tecan. Starting from lysed and denatured samples, this automated workflow provides ready-to-measure peptides in 96-well format for hundreds of samples within a single day while reducing overall hands-on time to only 20 minutes. Besides automation, also the iST procedure itself has become more versatile by coupling it to an upstream SP3-based protein clean-up and concentration step. This allows the removal of contaminants interfering with digestion and LC-MS analysis as well as the concentration of samples making this novel workflow applicable for various sample types and compatible with all kinds of buffer conditions. The complete automated workflow covers the SP3 sample handling, iST-based protein alkylation and digestion on the Freedom EVO as well as the iST peptide clean-up on the Resolvex A200 and can be implemented by simple script-based programming, thus providing utmost user-friendliness and accessibility for everyone. Here, we demonstrate high-throughput automation of sample processing resulting in excellent sample quality and high reproducibility for cell lysates, scaling from 1-50 µg. By combining the proven quality of the PreOmics iST technology with the power of liquid handling automation platforms from Tecan, we simplify and standardize efficient and highly reproducible LC-MS sample preparation.

Contact: Katrin Hartinger, hartinger@preomics.com

P01.11

Spray-capillary Based Capillary Electrophoresis Mass Spectrometry for Metabolite Analysis in Single Cells

Lushuang Huang (1), Mulin Fang (1), Kellye Cupp-Sutton (1), Zhe Wang (1), Kenneth Smith (2)

(1) Chemistry and biochemistry, University of Oklahoma, Norman, Oklahoma, United States

(2) Department of Arthritis and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, United States

Capillary electrophoresis mass spectrometry (CE-MS) has been extensively used for single-cell analysis. Conventionally, cell content was extracted through microsampling method (e.g., pump-based approach) and off-line coupled with CE separation by several sample preparation steps, which often requires dedicated transferring devices and may suffer from potential sample loss during intermediate steps. Here, we demonstrated the first application of online coupling microsampling method with CE-MS platform for single-cell analysis (e.g., onion cell) that features on high reproducibility and simplified workflow. Our group has developed an electrospray-assisted microsampling device, “spray-capillary”, which utilizes electrospray process to generated highly consistent tiny vacuum force for sample loading (e.g., ~ 15 pL/s injection flow rate). To adopt the device for single-cell analysis, the sample inlet end was laser pulled to decrease the bore size to about 15 µm. We first utilized the device for semi-quantitative microsampling coupled with direct MS detection. Sample injection time was varied to test quantitative sample extraction performance. Onion metabolites such as amino acid, sulfur containing compound and oligosaccharides were putatively assigned by accurate mass measurement and confirmed by targeted MS/MS. In total, we identified about 80 metabolites by both MS/MS and cross-verification with previous reports. Next, we performed spray-capillary based single-cell CE-MS analysis, which features at online coupling microsampling process with CE separation. The spray-capillary device with a laser pulled sample inlet end was directly used as both microsampling tool and CE separation column. High voltage was applied right after microsampling process by manually moved sample inlet end to background electrolyte reservoir. Reproducibility was demonstrated by performing triplicate trials on different single onion cells and displaying separation merits such as relative migration time, peak area

and peak height. By adding CE before MS detection, more than 160 mass features were putative assigned and verified by cross-checking with previous reports.

Contact: Lushuang Huang, lshuangh001@gmail.com

P01.12

Photocleavable Surfactant-Enabled Extracellular Matrix Proteomics

Samantha Knott (1), Kyle Brown (1), Harini Josyer (2), David Inman (1), Suzanne Ponik (2), Andreas Friedl (3), Ying Ge (4)

(1) Department of Chemistry, University of Wisconsin Madison, Madison, WI, USA

(2) Department of Cell and Regenerative Biology, University of Wisconsin Madison, Madison, WI, USA

(3) Department of Pathology and Laboratory Medicine, University of Wisconsin Madison, Madison, WI, USA

(4) Department of Chemistry, Department of Cell and Regenerative Biology, Human Proteomics Program, University of Wisconsin Madison, Madison, WI, USA

The extracellular matrix (ECM) provides an architectural meshwork that surrounds and supports cells. The dysregulation of heavily post-translationally modified ECM proteins directly contributes to various diseases. Mass spectrometry (MS)-based proteomics is an ideal tool to identify ECM proteins and characterize their post-translational modifications, but ECM proteomics remains challenging owing to the extremely low solubility of the ECM. Herein, enabled by effective solubilization of ECM proteins using our recently developed photocleavable surfactant, Azo, we have developed a streamlined ECM proteomic strategy that allows fast tissue decellularization, efficient extraction and enrichment of ECM proteins, and rapid digestion prior to reversed-phase liquid chromatography (RPLC)-MS analysis. A total of 173 and 225 unique ECM proteins from mouse mammary tumors have been identified using 1D and 2D RPLC-MS/MS, respectively. Moreover, 87 (from 1DLC-MS/MS) and 229 (from 2DLC-MS/MS) post-translational modifications of ECM proteins, including glycosylation, phosphorylation, and hydroxylation, were identified and localized. This Azo-enabled ECM proteomics strategy will streamline the analysis of ECM proteins and promote the study of ECM biology.

Contact: Samantha Knott, sknott2@wisc.edu

P01.13

Comprehensive analysis of RBM20 modifications by middle-down and bottom-up high-resolution mass spectrometry

Eli Larson (1), Yutong Jin (1), Yanghai Zhang (2), Wei Guo (2), Ying Ge (3)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

(2) Department of Dairy and Animal Science, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

(3) Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, Wisconsin, United States of America

RNA binding motif 20 (RBM20) is a splicing factor which is highly abundant in heart and muscle tissue. Similar to other splicing factors, RBM20 contains a serine-arginine (SR) domain which is critical for interactions with other proteins in the spliceosome assembly and protein-RNA interactions. The modification of SR domains is known to play a key role in splicing control, however, the SR domain modification status of RBM20 and its effect on the function of RBM20 is wholly unexplored. In this study, we present a comprehensive approach to the characterization of RBM20 modifications. This has been achieved through the use of both middle-down and bottom-up mass spectrometry (MS). First, RBM20 was analyzed by tryptic digestion followed by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) using the new Bruker timsTOF Pro. The timsTOF Pro utilized PASEF (parallel accumulation-serial fragmentation) to achieve 100% instrument duty cycle. This allowed mapping of primary sequence variants, identification of potential post-translational modifications (PTMs), and label-free quantitation of changes in protein isoform expression with high speed and sensitivity. Second, a middle-down MS approach using limited proteolytic digestion was applied to RBM20 by offline LC and direct infusion with ultrahigh-resolution Fourier transform ion cyclotron resonance (FTICR) MS/MS on a 12 tesla Bruker solariX XR. The use of FTICR MS/MS provides high mass accuracy detection of polypeptides and a suite of fragmentation techniques including electron capture dissociation (ECD), which preserves phosphorylation to allow relative quantitation of PTMs and determination of modification site occupancy. This multifaceted approach provides a platform for assessing splicing factor modification states and paves the way for further studies of proteomic changes in disease models.

Contact: Eli Larson, ejlarson2@wisc.edu

P01.14

Automation of Cell and Tissue Sample Preparation Workflow for High Throughput Proteomics

Danica-Mae Manalo (1), Vineet Vaibhav (1), Andrea Matlock (1), Vidya Venkatraman (1), Jennifer Van Eyk (1)

(1) Advanced Clinical Biosystems Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

Introduction

Sample preparation, involving enzymatic digestion of protein to peptides, is crucial for mass spectrometry-based proteomics experiments. Variation in the sample processing steps can influence the overall outcome of the experiment and affect the detection of proteins in the mass spectrometry acquisition stage, thus impact the downstream analysis and biological interpretation of the study. There is a need to develop an efficient workflow for sample processing of complex samples. This can be accomplished by implementing automation for completely hands-free sample preparation. Automation provides a minimum amount of variation compared to manual sample preparation methods, and is especially beneficial for working with large study cohorts. Here we outline 60-70 iPSC-derived motor neuron obtained from hundreds of ALS patients and healthy donors as a model system for cell/tissue automation.

Methods

Batches of 60-70 iPSC-derived motor neurons from the Answer ALS consortium were processed using the Biomek i7 Automated Liquid Handling Workstation (Beckman-Coulter). Samples were acquired by DDA-MS for sample-specific library generation and analyzed using DIA-MS methods on Triple TOF 6600 (Sciex). Batch technical control (BTC) pellets were used to monitor experimental performance within each batch. To measure precision, analyses of %CV were conducted on protein and peptide level. OpenSWATH [1] was used to quantitate the number of peptides and proteins identified across each sample per batch.

Results

Results show consistency in protein and peptide quantitation across batches (~3000 non-reducndnat proteins and ~20,000 peptides). BTC analyses reveal that 80% of proteins had a %CV < 25%.

Conclusion

Integration of automation in our sample preparation workflow for complex cell/tissue samples represents good data quality based on consistency of the quantitative data and tightness of %CVs. This quality data positively influences downstream analyses, which facilitates in the discovery of meaningful biological information in ALS research.

1. Röst, H.L., et al. 2014. Nature Biotechnology 32, 2019-223.

Contact: Danica-Mae Manalo, danica-mae.manalo@cshs.org

P01.15

Global Phosphoproteome Analysis Using High-field Asymmetric Waveform Ion Mobility Spectrometry on a Hybrid Orbitrap Mass Spectrometer

Laura Muehlbauer (1), Alexander Hebert (2), Michael Westphall (2), Evgenia Shishkova (2), Joshua Coon (1)

(1) Chemistry, University of Wisconsin-Madison, Madison, WI, USA

(2) Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI, USA

Mass spectrometry is the premier tool for identifying and quantifying protein phosphorylation on a global scale. Analysis of phosphopeptides requires enrichment, and even after the samples remain highly complex and exhibit broad dynamic range of abundance. Achieving maximal depth of coverage for phosphoproteomics therefore typically necessitates offline liquid chromatography prefractionation, a time-consuming and laborious approach. Here, we incorporated a recently commercialized aerodynamic high-field asymmetric waveform ion mobility spectrometry (FAIMS) device into the phosphoproteomic workflow. We characterize the effects of phosphorylation on the FAIMS separation, describe optimized compensation voltage settings for unlabeled phosphopeptides, and demonstrate the advantages of FAIMS-enabled gas-phase fractionation. Standard FAIMS single-shot analyses identified around 15-20% additional phosphorylation sites than control experiments without FAIMS. In comparison to liquid chromatography prefractionation, FAIMS experiments yielded similar or superior results when analyzing up to four discrete gas-phase fractions. Although using FAIMS led to a modest reduction in the precision of quantitative measurements when using label-free approaches, data collected with FAIMS yielded a 26% increase in total reproducible measurements. Overall, we conclude that the new FAIMS technology is a valuable addition to any phosphoproteomic workflow, with greater benefits emerging from longer analyses and higher amounts of material.

Contact: Laura Muehlbauer, lmuehlbauer@wisc.edu

P01.16

Preserving Trypsin Stability for Accelerated Digestion in the ProTrap XG

Jessica Nickerson (1), Alan Doucette (1)

(1) Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, Canada

Introduction | Efficient and reproducible enzymatic digestion is integral in achieving optimal MS data in a bottom-up proteome workflow. Recently, cartridge and bead-based digestion strategies have promoted generation of abundant peptide lists following accelerated digestion. These strategies capitalize on elevated temperature, high enzyme/sample ratios or digestion on solid supports. The desired outcome of a tryptic digestion is to maximize reproducibility and efficiency. This is best achieved through a complete peptide cleavage, which no strategy appears to produce. We will herein determine the conditions that sustain maximal enzyme activity across the digestion period, providing complete and reproducible protein digestion in minimal time. Preliminary experiments indicate the essential role of calcium in sustaining trypsin activity at elevated temperature, with a three-fold enhancement of activity over the duration of the digest. LC-MS/MS assays will be used to determine optimal digestion conditions that maximize proteomic data and sample throughput.

Methods | Standard protein mixtures will be precipitated, re-solubilized and digested within the ProTrap XG (Proteoform Scientific, Canada) with digestion conditions ranging in temperature, calcium, trypsin acetylation and digestion time. Relative quantitation will be achieved by differential dimethylation, followed by LC-MS/MS analysis. Comparative abundance of fully-cleaved peptides will be used to determine a novel digestion strategy with improved reproducibility and throughput.

Preliminary Results | BAEE activity assays conducted in a range of conditions demonstrated that the optimal temperature for trypsin activity is 47 ˚C, showing a 3-fold increase compared to the conventional temperature, 37 ˚C. This enhanced activity is unfortunately lost within 30-60 min digestion. However, adding 1-10 mM Ca^{2+} sustains >90% of the initial activity across 4 h at pH 8.0.

Conclusions | Quantitative assessment of fully cleaved peptides generated by a range of digestion conditions and times will contribute to a rapid digestion protocol within the ProTrap XG.

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

P01.17

BoxCarmax: a High-Selectivity Data-Independent Acquisition Mass Spectrometry Method for the Analysis of Protein Turnover and Complex Samples

Barbora Salovska (1), Wenxue Li (1), Yi Di (1), Yansheng Liu (2)

(1) Yale Cancer Biology Institute, Yale University, West Haven, CT, US

(2) Department of Pharmacology, Yale University School of Medicine, New Haven, CT, US

The data-independent acquisition (DIA) performed in the latest high-resolution, high-speed mass spectrometers offers a powerful analytical tool for biological investigations. The DIA mass spectrometry (DIA-MS) combined with the isotopic labeling approach holds a particular promise for increasing the multiplexity of DIA-MS analysis, which could assist the relative protein quantification and the proteome-wide turnover profiling. However, the wide isolation windows employed in conventional DIA methods lead to a limited efficiency in identifying and quantifying isotope-labelled peptide pairs. Here, we optimized a high-selectivity DIA-MS named BoxCarmax that supports the analysis of complex samples, such as those generated from Stable isotope labeling by amino acids in cell culture (SILAC) and pulse SILAC (pSILAC) experiments. BoxCarmax enables multiplexed acquisition at both MS1- and MS2- levels, through the integration of BoxCar and MSX features, as well as a gas-phase separation strategy. We found BoxCarmax modestly increased the identification rate for label-free and labeled samples but significantly improved the quantitative accuracy in SILAC and pSILAC samples. We further applied BoxCarmax in studying the protein degradation regulation during serum starvation stress in cultured cells, revealing valuable biological insights. Our study offered an alternative and accurate approach for the MS analysis of protein turnover and complex samples.

Contact: Barbora Salovska, barbora.salovska@yale.edu

P01.18

PASEF-PRM-LIVE and its Application in Functional Protein Profiling

He Zhu (1), Scott Ficarro (1), William Alexander (1), Laura Fleming (1), Wai Cheung Chan (1), Guillaume Adelmant (1), Shourjo Ghose (2), Matthew Willetts (2), Jens Decker (3), Sven Brehmer (3), Gary Kruppa (4), Jarrod Marto (1)

(1) Dana-Farber Cancer Institute, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

(2) Bruker Daltonics Inc, Billerica, MA, USA

(3) Bruker Daltonik GmbH, Bremen, Germany

(4) Bruker S.R.O., District Brno-City, Czech Republic

Introduction
Integrated TIMS with PASEF affords an opportunity for highly selective and rapid detection of peptides in complex matrices. In a traditional PRM analysis, retention time (RT) drift is often problematic. Use of wider detection windows can offset RT drift, albeit at the cost of fewer peptides in each targeted analysis. Here we introduce PASEF-PRM-LIVE, a framework implementing on-the-fly adjustment of detection window to maximize target coverage. This platform is ideally suited for functional protein profiling.

Methods
We have developed a Python API allowing real-time control and execution of PASEF-PRM scans which we call PASEF-PRM-LIVE. Our acquisition framework monitors the elution of user-specified landmark peptides and then uses a regression model to correct the targeting window of upcoming peptide targets. We use activity-based probes to enrich kinases or other enzymes, followed by trypsin digest to generate peptides for use as targets in PASEF-PRM-LIVE.

Preliminary Results
As an initial benchmark, we queried ~80 peptide targets in human biofluids based on 45-min DDA run to test the performance of the time-warp algorithm to dynamically adjust precursor scheduling parameters at 30min, 45min, and 60min gradient, respectively. We found that the RT of targets between PRM and DDA had a good correlation (Pearson r > 0.99) at each LC gradient, indicating that PASEF-PRM-LIVE framework accurately tracked RT drift and dynamically adjusted acquisition parameters. We next used PRM LIVE to target and quantify 1875 peptides from HeLaS3 cells in a 60-minute gradient. Across five replicates, 1753 peptides exhibited good reproducibility (CV ≤ 20%), which demonstrates the ability to profile large numbers of targets in a complex mixture. We further apply this technique in an ABPP assay to quantify the targets of kinase inhibitors.

Conclusions
Our framework for PASEF-PRM-LIVE dramatically increases the throughput of peptide PRM quantification, and is successfully used in activity-based protein profiling.

Contact: He Zhu, he_zhu@dfci.harvard.edu

Aging and Neurodegenerative Diseases

P02.01

A Tissue Specific Atlas of Autophagosome Cargo Proteins.

YoungJoo Yang (1), Vijay Raja (2), Ai Yamamoto (1), Noah Dephoure (2)

(1) Department of Neurology, Columbia University, New York, NY, USA

(2) Department of Biochemistry, Weill Cornell Medical College, New York, NY, USA

Autophagy is a highly conserved, lysosome mediated process for degrading cellular components that is widely implicated in human disease. We set out to explore its tissue specific roles by quantitatively profiling the protein content of purified autophagic vacuoles (AVs) in mouse liver and brain. To isolate highly purified autophagic vacuoles (AVs), we used mice engineered to express a GFP fusion of the LC3 autophagosome marker coupled with immunopurification of GFP-LC3. Using 10plex TMT labeling and a SPS-MS3 method on an Orbitrap Fusion mass spectrometer, we performed triplicate analysis of brain and liver AVs from fed and starved mice. We observed a range of tissue specific and starvation induced changes in the AV

proteome. Many of them reflected tissue specific protein expression patterns, e.g., p450 proteins in liver and synaptic proteins in brain. However, proteins common to both tissues were also found to traffic to AVs differently: peroxisomal, endoplasmic reticulum (ER) and ribosomal proteins were more abundant in liver AVs, whereas mitochondrial proteins were more abundant in the brain. Upon 24 hr starvation, in the liver there was an increase in ER and peroxisomal proteins and a decrease of mitochondrial proteins. Further analysis revealed that while most mitochondrial proteins were decreased, proteins of the outer mitochondrial membrane (OMM) increased. Subsequent experiments showed that starvation triggered a shift in liver autophagy from eliminating intact mitochondria to selectively targeting the OMM. In brain, we observed no changes in neuron-specific AV cargoes but saw increases in proteins unique to non-neuronal cells, such as astrocytes, oligodendrocytes, and brain endothelial cells. These studies revealed core insights into the physiological response of autophagy in two organs in which this pathway has been deemed essential and provide a blueprint for more focused studies to examine the role of autophagy in human disease.

Contact: Noah Dephoure, nod2007@med.cornell.edu

P02.04

Proteomic Investigation of the I1061T Point Mutation Mouse Model of Niemann-Pick Type C

Wenping Li (1), Melissa Pergande (1), Stephanie Cologna (1)

(1) Chemistry, Chicago, IL, USA

Niemann-Pick Type C (NPC) is an autosomal recessive, neurodegenerative disease. Mutations of either NPC genes (NPC1 or NPC2) lead to the accumulation of unesterified cholesterol and sphingolipids in the late endo/lysosomal system. NPC is fatal and hard to diagnose at an early stage with no FDA approved therapy; although Miglustat, a glycosylceramide synthase inhibitor, has shown to ameliorate neurological manifestations. To date, most studies comparing diseased and control animals have been done on NPC1-null mice, which recapitulates the human disease, however, does not accurately mimic the most common genetic mutation in patients. Previously, we investigated the differential proteins in symptomatic (11-week), NPC1-null animals compared to age matched healthy controls. Major findings of that dataset were the identification of altered LIMP-2/SCARB-2, a lysosomal membrane protein, fatty acid transport via fatty acid binding proteins and calcium signaling defects. Currently, our efforts are two-fold: 1) to investigate whether the same differential proteome is consistent between the cerebellum and cerebral cortex of both the NPC1-null mouse model and the NPC1 I1061T point mutation ($NPC1^{I1061T}$), 2) to evaluate earlier time points of the disease to find other potential therapeutic targets as the 11-week mice are near terminal. To do this, tissue lysates will be digested using S-Trap method followed by isobaric labeling of peptides (TMT-10). Data will be acquired on a Q-Exactive nLC-MS system and analyzed using the SEQUEST algorithm via the Thermo Proteome Discoverer software. Proteins involved in different pathways will be analyzed using Ingenuity Pathway Analysis. In this study, we expect to observe both common and unique altered proteins and biological pathways due to the I1061T mutation in NPC1. This will help identify potential biomarkers which can be used for monitoring of potential therapeutic treatments and disease progression.

Contact: Wenping Li, jklwp007@uic.edu

P02.05

Dynamic changes of exosome proteins and post-translational modifications during cellular senescence and aging.

Sandip Kumar Patel (1), Jacob Rose (1), Nathan Basisty (1), Lindsay Pino (2), Roland Bruderer (3), Lukas Reiter (3), Judith Campisi (1), Birgit Schilling (1)

(1) Buck Institute for Research on Aging, Novato, CA, USA

(2) Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA, USA

(3) Biognosys AG, Schlieren, Switzerland

Aging is a complex biological process associated with progressive loss of physiological function leading to several diseases, including neurodegeneration. Through paracrine signaling of the senescence-associated secretory phenotype (SASP), senescent cells drive aging phenotypes. Exosomes play a vital role in cell-to-cell signaling; thus, we hypothesized that senescence-associated exosome markers might qualify as biomarkers for aging and neurodegenerative diseases. SASP exosomes and their cargo, including proteins, miRNAs, and post-translational modifications (PTMs) have so far been unexplored.

We optimized exosome extraction to obtain highly pure exosomes from various biological samples, including cell secretomes and human plasma, by sequential size-exclusion chromatography and ultrafiltration (SEC/UF). We generated a human plasma exosome-specific deep spectral library comprising ~2,300 exosome proteins by data-dependent acquisitions (DDA) measuring 25 offline generated high-pH reversed-phase fractions. Subsequently, in a pilot aging study, we used comprehensive data-independent acquisitions (DIA) to classify plasma exosomes from young and old individuals. We quantified 1,310 exosome proteins, with two unique peptides, of which 144 were significantly changed in young and old plasma groups ($Q<0.05$ and >1.5-fold change). Several pathways, including negative regulation of wound healing and antioxidant, were upregulated, while defense response to bacteria, intermediate filament, cytoskeleton, and complement activation pathways was downregulated in old individuals.

Further, we investigated exosome protein PTMs in aged plasma. Re-analyzing the DIA data, we have identified ~100 phosphopeptides with high confidence in old plasma exosome samples. We also investigated tissue culture soluble SASP and performed western blot analyses of the senescent IMR90 SASP, which showed changes in phosphoprotein patterns. These intriguing results provide insights into PTM-protein profiles of the SASP and dynamic regulation during cellular senescence/aging. This highly translational work may lead to 'Biomarkers of Aging' for prognostic and diagnostic applications.

Acknowledgments: U01AG060906 (Schilling), K99AG065484 (Basisty), P01AG017242 and R01AG051729 (Campisi), Glenn Fellowship (Patel).

Contact: Sandip Kumar Patel, spatel@buckinstitute.org

P02.06

Brain Myelin Proteomics of Neurodegenerative Disorder Niemann-Pick type C1 Mouse Model

Chandimal Pathmasiri (1), Stephanie Cologna (1)

(1) Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois, USA

Niemann-Pick disease, type C1 (NPC1) is a fatal, lysosomal storage disorder, which results in progressive neurodegeneration. Complications affect most severely on central nervous system. Dysmyelination in NPC1 is an understudied phenotype and has mostly been described using standard immunostaining markers. While proteins like myelin basic protein (MBP) and myelin-associated glycoprotein (MAG) have been shown to be altered in NPC1, we sought to understand the myelin proteome defects in a more comprehensive manner in relation to known lipid alterations. For example,

MAG binds ganglioside lipids on axons to ensure axon-myelin interaction and gangliosides are altered in NPC1. Our overarching hypothesis is that alterations of brain myelin proteins in NPC1 could lead to disruption of axon-myelin interactions and myelin compactness.

All studies used the Balb/c $Npc1^{nih}$ (Npc1-null) mouse model. Myelin was isolated using sucrose gradient centrifugation from Npc1-null and control cerebellum, cortex, and midbrain tissues. The purity of isolated myelin is assessed using Western blot and protein lysates are digested using S-trap. Peptide separation is done using nano flow on a C18 reverse phase column and detection via Thermo Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer. Protein identification was obtained using Proteome Discoverer equipped with the SEQUEST search engine and a false-discovery rate set to 1%. Protein identifications required at least 2 unique peptides. In our first study at 7-week of age, using cerebellar myelin, we were able to identify more than 900 proteins including major myelin specific proteins using above workflow. We observed alterations of several lipid binding proteins including type 1 phosphatidylinositol 4,5-bisphosphate 4-phosphatase, lysosomal membrane proteins, cathepsins and chloride transporter proteins. Proteins with Log_2 fold change higher than 1.5 and P-value less than 0.05 considered as significant. Our current work is focused on understanding myelin defects prior to and after myelination is completed using 2,5 and 9-week old mice animals.

Contact: Chandimal Pathmasiri, kpathm2@uic.edu

P02.07

Differential proteomics in a mouse model of Niemann-Pick, Type C following 2-hydroxypropyl-beta-cyclodextrin treatment

Melissa Pergande (1), Antony Cougnoux (2), Wenping Li (1), Fidel Serna-Perez (1), Nigina Khamidova (1), Forbes Porter (2), Stephanie Cologna (1)

(1) University of Illinois at Chicago, Chicago, IL, USA

(2) National Institutes of Health, Bethesda, MD, USA

Niemann-Pick, Type C (NPC) disease is a fatal, neurodegenerative lysosomal storage disorder caused by mutations in the NPC1 or NPC2 gene. As a result of the genetic defect, unesterified cholesterol trafficking is impaired and accumulates within endo/lysosomal compartments. In vitro and animal studies have revealed that 2-hydroxypropyl-beta-cyclodextrin (HPβCD) reduces cholesterol storage, cerebellar degeneration and extends animal lifespan. Current clinical investigation is underway in patients with NPC. In the present study, we sought to investigate the differential proteome following chronic treatment with HPβCD. Label free quantitative proteomics was carried out using standard flow chromatography and JetStream electrospray ionization coupled to a Q-TOF mass spectrometer. Enriched pathways that are affected by HPβCD in the NPC mouse model include downregulation of the CREB signaling, synaptogenesis signaling, synaptic long-term depression and neuroinflammation, whereas an up-regulation of toll-like receptor signaling was observed with treatment. The results of this study show differences induced by HPβCD treatment in wild type animals including the endocannabinoid neuronal synapse pathway. Our current investigations are geared towards understanding how HPβCD affects synapse function with focus on endocannabinoid signaling between presynaptic and postsynaptic neurons in NPC. Results from the proteomic analysis in the NPC mouse model, neuronal cell culture model, as well as validation and functional studies will be presented.

Contact: Melissa Pergande, mperga2@uic.edu

P02.08

Evaluating ventricular CSF as a common reference material in lumbar CSF studies

Deanna Plubell (1), Michelle Emrick (2), Gennifer Merrihew (1), C. Dirk Keene (2), Andrew Hoofnagle (2), Michael MacCoss (1)

(1) Genome Sciences, University of Washington, Seattle, WA, USA

(2) Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

Studies on the lumbar cerebrospinal fluid (CSF) proteome have increased in number and in size over the recent years. To calibrate measurements across studies, we advocate for use of a common reference material to enable comparison of results between laboratories and clinical cohorts. This requires a sufficient sample volume to consistently prepare and measure with each "batch" and appropriate composition of reference proteins. As a potential reference CSF source, we investigated the suitability of post-mortem ventricular CSF collected from research brain donors with <8 hour postmortem interval (n=11). Lumbar CSF (n=11) was collected by lumbar puncture for routine testing and biobanking. A portion of each sample was aliquoted to form a common pool for each CSF sampling location, and a common pool across locations. Digested peptides were analyzed by DIA for both library generation and individual sample quantitation according to Pino et al., MCP, 2020. Peptides were detected with EncyclopeDIA using a PROSIT predicted spectral library. Ventricular CSF has 1.4x higher protein concentration, and 2.8x as many peptides detected in the ventricular GPF-DIA library (31,451) compared to lumbar (11,349). These peptides map to 2.3x as many proteins in ventricular (3,107) compared to lumbar (1,314). For individual samples, 22,986 peptides were quantified in ventricular compared to 9,003 peptides in lumbar. Of the peptides detected in lumbar, 70.8% are also detected in ventricular, while 83.4% of proteins detected in lumbar are also detected in ventricular. CSF from both sites have a similar dynamic range of summed protein intensity; spanning a range of 10^6. Of proteins of interest in neurodegenerative disease research, we detected 5 in both lumbar and ventricular (APOE, APP, SNCA, CHI3L1, NPTXR) and an additional 8 in ventricular CSF alone (Tau, SNCB, SNCG, NFL, NFM, NFH, SYT1, SNAP25).

Contact: Deanna Plubell, plubell@uw.edu

P02.09

Comprehensive proteomic profiling of induced pluripotent stem cells (iPSCs)-derived neurons using integrated FAIMS data-independent and data-dependent acquisitions.

Yue Andy Qi (1), Luke Reilly (1), Lirong Peng (1), Erika Lara Flores (1), Daniel Ramos (1), Julia Stadler (1), Caroline Pantazis (1), Mark Cookson (2), Michael Ward (3)

(1) Center for Alzheimer's Related Diseases, National Institute on Aging, NIH, Bethesda, MD 20892, USA

(2) Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD 20892, USA

(3) National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD 20892, USA

Induced pluripotent stem cells (iPSCs)-derived neuron has become an essential model to study neurodegenerative diseases. We have developed a rapid, simple 1-step differentiation protocol using doxycycline-induced expression of neuronal transcriptional regulator neurogenin-2. Although neuronal morphology of iPSC-derived neurons has been validated, their protein expression profiles have not been well characterized. We profiled

the whole-cell proteome of WTC11 iPSCs and their isogenic derived neurons (n = 4) using an Orbitrap Eclipse mass spectrometer (MS) coupled with a FAIMS interface. Protein extraction and digestion were performed on the KingFisher robotics. The resulting tryptic peptides were separated on a 2-hour LC gradient. MS/MS analysis was conducted by data-independent acquisition (DIA) and data-dependent acquisition (DDA) coupled with FAIMS using various compensation voltages (CVs) ranged from -20V to -85V. For the FAIMS-DIA experiment, 66,707 peptides from 8,621 proteins were identified. FAIMS with -35V provided the highest number of identifications; notably, nearly 50% of the peptides identified at -35V were not detected at other CVs. For the FAIMS-DDA experiment, we identified 49,848 peptides from 8,469 proteins. Combing DIA and DDA datasets, we identified 108,769 distinct peptides belonging to 10,120 proteins. Less than 10% of total peptides (~8,000 peptides) were identified by both approaches. Differential expression analysis of DIA results showed 2,565 proteins were significantly upregulated and 2,331 proteins were downregulated in the iPSC-derived neurons compared to iPSCs, indicating dramatic protein profile alternation. We validated several neuronal markers including SLC17A7, SLC17A6, MAP2, DCLK1 and SYN1 that were significantly activated in iPSC neurons. Pathway analysis showed proteins involved in synapse, cell junction and ribosome signaling were upregulated in neurons, whereas cell cycle, cell division, and DNA repair cascades were inhibited in neurons. Therefore, we provide a rich resource for iPSC neuron proteome; our results suggest that the integrated FAIMS-DIA and FAIMS-DDA approach might largely increase the proteome coverage.

Contact: Yue Andy Qi, andy.qi@nih.gov

P02.10

Quantitative Proteomic Analysis of Huntington's Disease Medium Spiny Neurons Reveals Keys Molecular Determinants of Neuropathogenesis

Kizito-Tshitoko TSHILENGE (1), Carlos Galica Aguirre (1), Nathan Basisty (1), Sicheng Song (2), Joanna Bons (1), Alejandro Lopez-Ramirez (1), Ashley Loureiro (1), Cameron Wehrfritz (1), Anja Holtz (1), Sean Mooney (2), Birgit Schilling (1), Lisa Ellerby (1)

(1) The Buck Institute for Research on Aging, Novato, California, US

(2) University of Washington, Seattle, Washington, US

Huntington's disease (HD) is a rare and incurable neurodegenerative disease associated with a sequential loss of neurons in the striatum, globus pallidus and cortex. HD is caused by unstable expansion of CAG repeats in exon 1 of the huntingtin (HTT) gene, which produces a polyglutamine (polyQ) tract that alters the localization and conformation of the HTT protein. Although the polyQ-expanded HTT is expressed throughout different tissues and organs, the medium spiny neurons (MSN) in the striatum are one of cell types predominantly affected in HD. How abnormal CAG repeats within mHTT leads to HD neuropathogenesis has not been fully understood. To study HD pathogenesis, we used human HD patient-derived neurons to differentiate MSNs and investigate the functional consequences of polyQ-HTT defect in terms of protein expressions and cellular vulnerability. We performed quantitative proteomic analysis of HD-MSNs when compared to isogenic corrected-MSNs using two independent approaches: data-dependent acquisitions with FAIMS (FAIMS-DDA) for label free quantification (LFQ) on the Orbitrap Lumos. This approach identified 6,323 proteins with at least two peptides (FDR ≤ 0.01). Of these, 497 proteins were significantly altered (q-value <0.05 and fold change 1.5) in the HD72-MSN compared to isogenic controls. Subsequently, we quantitatively validated protein candidates by comprehensive data-independent acquisitions (DIA) on a TripleTOF 6600 quantifying 3,106 proteins. We identified signaling pathways not previously known to be dysregulated in HD proteomes including major histocompatibility complex class (MHC) proteins, cellular senescence and regulation of cellular response to heat. The differential protein expression and pathway analysis revealed dysregulation of proteins involved in the maintenance and regulation of the extracellular matrix that are significantly associated with human HD signature. Our omics analysis using HD-MSN identified relevant pathways that are dysregulated in MSN and confirm current and new therapeutic targets for HD.

Contact: Kizito-Tshitoko TSHILENGE, KTshilenge@buckinstitute.org

Biomarkers and Targeted MS Assays

P03.01

Accelerating the Speed of Ion Library Generation Using Fast Microflow DDA and Cloud Processing

Alexandra Antonoplis (1), Nick Morrice (2), Christie Hunter (1)

(1) SCIEX, Technical Marketing, Redwood City, CA, USA

(2) SCIEX, Application Support, Alderley Park, Cheshire, UK

As DIA for quantitative proteomics continues to be used more broadly across studies, the ability to rapidly generate large ion libraries on selected samples is key. Microflow LC enhances both the robustness and speed of this process. Here the use of very fast gradients for DDA acquisition combined with data processing in a cloud-based processing solution was explored to accelerate ion library generation.

Methods: 40 high pH fractions of both plasma and colon cancer FFPE samples were run by DDA using a 10 min microflow gradient on a TripleTOF® 6600+ System, requiring 10 hours of MS acquisition. Each dataset was uploaded to the SCIEX Cloud Platform and searched with ProteinPilot™ App

in OneOmics™ Suite. A thorough search with biological modifications on was performed, meaning that over 800 modifications were considered in the search space using feature probabilities.

Results: Searching DDA data in the SCIEX Cloud Platform can provide improvements in search speed of 3-7x depending on the power of the desktop computer as well as allowing parallel processing. Here, cloud searching was used to accelerate the creation of ion libraries for SWATH Acquisition data processing. From the plasma fractions, an ion library consisting of 1178 proteins with 23981 peptides at < 1% global FDR rate was created with a processing time of 1.25 hours. From the FFPE fractions, an ion library of 7282 proteins with 94411 peptides was created with a processing time of 4.7 hours. The observed median peak widths of the datasets were ~4 sec wide at half height, highlighting the quality of the chromatography that facilitated such high quality DDA datasets.

Given the speed and scale of proteomics experiments today, fast LC-MS and cloud-based database searching, ion libraries to power SWATH acquisition studies can be generated easily in a less than a day.

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

P03.02

Deciphering human immune system from a droplet of blood using proteomics

Benoit Fatou (1), Kinga Smolen (2), Alexia Belperron (3), Zainab Wurie (1), Steven Kleinstein (4), Ruth Montgomery (3), Ofer Levy (5), Linda Bockenstedt (3), Hanno Steen (1)

(1) Department of Pathology, Precision Vaccines Program Boston Children's Hospital / Harvard Medical School, Boston, MA, USA

(2) Precision Vaccines Program, Division of Infectious Diseases Boston Children's Hospital / Harvard Medical School, Boston, MA, USA

(3) Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

(4) Department of Pathology, Yale School of Medicine, Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA

(5) Precision Vaccines Program, Division of Infectious Diseases Boston Children's Hospital / Harvard Medical School,, Broad Institute of Massachusetts Institute of Technology & Harvard, Boston, MA, USA

Blood specimens are widely used to diagnose disease or monitor health status because blood represents an ideal source for biomarkers: it is a systemic body fluid, easily and readily accessible using minimally invasive collection procedures. However, mapping the blood proteome is challenging as the 10 most abundant proteins account for 99% of the protein content and thereby mask detection of the less abundant proteins. Further, the wide variations in the population require the processing and analysis of a large number for meaningful clinical proteomics study. Using an in-house developed method for high throughput microtiter plate-based serum/plasma processing, we are able to rapidly process submicroliter volumes of 96 individual samples per batch resulting to detect and quantify ~400 classical plasma proteins, covering a protein concentration range of up to 7 orders of magnitude. This initial version of our plasma proteomics platform was applied to samples collected in the context of Lyme disease and West Nile Virus infection to identify biomarkers for disease stratification and diagnosis, respectively. However, there is a need to increase the throughput of the entire platform, i.e. to remove any bottlenecks. We have developed a targeted and multiplexed LC/MS method dedicated for human plasma/serum samples, monitoring the top classical plasma proteins in a high throughput process (<15 min per sample). This method is based on a data-driven approach by analyzing the results from various studies focused on a wide range of immune challenges including bacterial and viral diseases as well as inflammation- and vaccination-related projects, which allowed us to develop a fast, sensitive, accurate and precise LC/MS method, which is currently being applied to various large scale studies within the Human Immunology Project Consortium (HIPC) and the IMMuno Phenotyping Assessment in a COVID-19 cohort (IMPACC) study with up to analyze 10,000 plasma samples.

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

P03.03

Label-free Mass Spectrometry Identification of Vascular Integrity Markers in Major Trauma

Holly Hinson (1), Jon Jacobs (2), Shannon McWeeney (3), Tujin Shi (2), Kendall Martin (2), Karin Rodland (4)

(1) Neurology, Oregon Health and Science University, Portland, OR, US

(2) Biological Science Division, Pacific Northwest National Laboratory, Richland, WA, US

(3) Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland, OR, US

(4) Developmental and Cancer Biology, Oregon Health and Science University, Portland, OR, US

Antibody mediated strategies of protein biomarker detection are common, but may limit discovery. We hypothesized that label-free proteomics is feasible for detecting protein biomarkers in plasma of patients sustaining major trauma.

A subset of subjects with major trauma were analyzed from prospective observational trial. Patients were assigned to one of 4 groups based on their presenting Abbreviated Injury Severity Scores (AIS): 1) Polytrauma, 2) Isolated Head Injury 3) Isolated Body Injury, and 4) Mild injury. Baseline blood samples were collected within 6 hours of the antecedent trauma. Sensitive, antibody-free selective reaction monitoring (SRM) mass spectrometry, with spiked-in isotopically labelled synthetic peptides, was used for targeted protein quantification of a panel of prospective targets: Ang2, VCAM1, ICAM1, PAI1, and MMP9. Multiple peptides for each target where identified and optimized for precursor-to-fragment ion transitions using an SRM platform based upon a triple quadrupole mass spectrometer. An overall tiered sensitivity analytical approach was used for peptide detection and quantification based upon plasma immunoaffinity depleted and PRISM fractionation.

Forty-four patients were included in the analysis, of which 82% were men with a mean age of 50 (+/-19). Half had isolated head injury (n=22), with the remaining patients experiencing polytrauma (n=14), isolated body (n=2), or minor injury (n=6). Peptides from three proteins (VCAM1, ICAM1, and MMP9) were detected and quantified in non-depleted processed plasma. Peptides from two proteins (Ang2 and PAI1) were detected and quantification in depleted plasma. VCAM1 (P=0.02) and MMP9 (P=0.03) were significantly upregulated in in the major trauma groups (1-3) v. mild injury (4), while the others were not. There were no differences in protein expression between patients with TBI (1-2) v. those without (3-4).

We detected non-specific up-regulation of proteins reflecting blood brain barrier breakdown in severely injured patients, indicating label-free MS techniques are feasible and may be informative.

Contact: Holly Hinson, hinson@ohsu.edu

P03.04

Reproducible targeted peptide profiling using multiplexed MRM assays on a highly sensitive triple quadrupole mass spectrometer

Christie Hunter (1)

(1) SCIEX, Redwood City, CA, USA

MRM analysis for targeted protein quantification studies is driven by the renowned sensitivity and selectivity of QQQ-based MS systems. But quantitative robustness as well as sensitivity remains key, to ensure both large and small biological changes are accurately measured across large sample cohorts. Biological matrices also require the LC-MS/MS system to have high sensitivity as well as wide linear dynamic range. Methods: A study was performed in simple and complex matrices to test the performance of the SCIEX Triple Quad™ 7500 LC-MS/MS System – QTRAP® Ready for large scale targeted peptide quantification. Methods with up to 4000 MRM transitions were run using Scheduled MRM Algorithm using microflow chromatography (5 and 30 min gradients) and the quantitative reproducibility was evaluated.

Results: The average retention time %RSD observed for the 10-15 replicates analyzed using each method was 0.13% and 0.2% for both the 5 and 30 min gradients. This allowed small detection windows to be used such that the dwell times remained above 5 msec even for the 4000 MRM methods.

The effect of higher numbers of MRM transitions on the reproducibility of 50 peptide MRMs (3 protein digests in buffer) across 15 replicates on the SCIEX 7500 System was assessed using microflow chromatography (30 min gradient). High reproducibility was observed for methods with up to 4000 MRMs per method, with ~85-90% of peptides showing %CV<5%. A similar study was then performed with the same 3 protein digests spiked in complex matrix with 10 replicates using the 5 min gradient. Again, high reproducibility of up to 3000 MRMs was achieved, with 90% of peptides having %CV <4.5%.

Sensitivity and quantitative reproducibility of the SCIEX Triple Quad 7500 System with highly reproducible microflow chromatography provides a powerful tool for large scale targeted peptide quantification studies.

Contact: Christie Hunter, CHRISTIE.HUNTER@SCIEX.COM

P03.05

SureQuant targeted MS workflow for quantitative analysis of oncogenic signaling pathways

Penny Jensen (1), Bhavin Patel (1), Aaron Gajadhar (1), Sebastien Gallien (1), Andreas Huhmer (1), Daniel Lopez-Ferrer (1), Kay Opperman (1), John Rogers (1)

(1) Thermo Fisher Scientific, Rockford, IL, USA

Introduction: The RAS/MAPK and AKT/mTOR pathways represent key mechanisms for cells to regulate cell survival, proliferation, and motility. The cross-talk between two pathways plays a central role in tumor progression and anti-cancer drug resistance. The quantitation of pathway protein expression and modifications are critical for characterization of disease, monitoring cancer progression and determining treatment response. A major limitation in the quantitation of pathway proteins is the lack of rigorously validated methods/reagents and a reliance on semi-quantitative results from Western blotting. We have utilized a novel SureQuant internal standard (IS)-triggered method applying a pool of reference internal standards to quantitate abundance of pathway proteins in a single MS run.

Methods: Multiple cancer cell lines (HeLa, A549, HCT116) were treated and processed using either EasyPep™ MS sample prep kits. 100 fmol of a combined mixture of 96 peptides for 46 pathway targets was spiked into 1 µg of digests. Samples were subjected to LC-MS analysis using Thermo

Scientific Dionex nanoLC™ systems coupled to an Orbitrap Eclipse Tribrid Mass Spectrometer for SureQuant and Thermo Scientific™ Q Exactive™ HF Hybrid Quadrupole-Orbitrap Mass Spectrometer for PRM analysis. Data analysis was performed with Proteome Discoverer and Skyline software.

Results: In this study, we have developed a complete workflow solution by combining EasyPep MS sample prep kits and SureQuant targeted MS assays to quantitate biologically relevant signaling pathway proteins. The internal standards were spiked into digests from multiple treated cancer cell lines. Two targeted MS methods (PRM and SureQuant) were compared to assess the relative performance differences for quantitation of the desired endogenous peptides. More than 90% of the peptides were quantitated using the SureQuant analysis from multiple cancer cell lines.

Conclusions: SureQuant targeted MS analysis allows reproducible, routine and simultaneous quantitation of functionally relevant targeted peptides from signaling pathway proteins.

Contact: Penny Jensen, penny.jensen@thermofisher.com

P03.06

Evaluation of prm-PASEF for multiplexed targeted proteomics

Antoine LESUR (1), Jens DECKER (2), Sven BREHMER (2), Elisabeth LETELLIER (3), François BERNARDIN (1), Pierre-Olivier SCHMIT (4), Gunnar DITTMAR (1)

(1) Quantitative Biology Unit, Luxembourg Institute of Health, Strassen, Luxembourg

(2) Bruker Daltonik GmbH, Bremen, Germany

(3) Department of Life Sciences and Medicine, University of Luxembourg, Luxembourg, Luxembourg

(4) Bruker Daltonics S.A, Wissembourg, France

Introduction

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

Methods

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

Results

The Aqua peptides calibration curves showed a signal response fitted by a linear regression over a concentration factor of 2900 (from 17.2 to 50,000 amole injected column), and an averaged RSD of 3% for the heavy/light peptide signal ratios. For label free quantification, the 2 fmole 186 AQUA peptides measured over 30 LC-MS runs showed a median RSD of 10.5% and a median of 25 data points across the chromatographic peaks. Finally,

the G12 mutations of Ras protein family and the three isoforms (NRas, KRas and HRas) were identified and quantified in 10 colon and lung cancer cells lines using a 10 min chromatography separation.

Conclusions

This new targeted acquisition method that s well suited for clinical applications that require to measure high density of targets with fast chromatography separations.

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

P03.07

Peptide Quantification in Plasma Using the Agilent 6495 Triple Quadrupole LC/MS Coupled with the Agilent 1290 Infinity II LC System

Linfeng Wu (1)

(1) Agilent Technologies, Inc, Santa Clara, CA, USA

In translational research, increased throughput as well as increased multiplexing mass spectrometry using MRM-based LC/MS methods is popular for targeted, bottom-up quantification of protein biomarkers. Researchers often target signature peptides from a limited number of proteins in large cohorts and monitor their expression levels during a specified time period, often leading to hundreds or thousands of biological samples. In such large-volume studies, high-throughput, robustness, and reproducibility are essential when deploying LC/MS methods. For this reason, standard-flow triple quadrupole LC/MS (LC/TQ) is superior to other low-flow LC/MS platforms.

In addition to reproducibility and robustness, analytical sensitivity is another important consideration for peptide quantification. Therefore, we evaluated the quantitative analytical precision, accuracy and sensitivity of 99 signature peptides from 75 protein biomarkers in plasma using a standard flow-based LC/TQ system including the Agilent 1290 Infinity II LC and the 6495 triple quadrupole LC/MS systems coupled with AJS ionization source. To assess overall quantitative performance for protein biomarker peptides, a commercially available mixture of stable isotope-labeled standard (SIS) peptides was spiked into human plasma digest at nine different concentrations. Standard curves of 99 SIS peptides were generated to determine their lower limit of quantification (LLOQ) in plasma matrix. The results show excellent analytical performance under standard LC flow conditions with this LC/TQ system for targeted peptide quantification in heavy matrix.

Contact: Linfeng Wu, linfeng.wu@agilent.com

Cancer Proteomics

P04.01 Selective Pulse Chase-SILAC Labeling of Three-Dimensional Multicellular Spheroids for Global Proteome Analysis
Nicole Beller, Jessica Lukowski, Amanda Hummon

P04.02 Quantitative proteomics reveals T-cells specific response from pancreatic cancer stem cells when compared to PANC1 in 2D and 3D cultures
Raghothama Chaerkady, Farah Alayli, Hsiang-En Hsu, Shao Huan Samuel Weng, Junmin Wang, Wen Yu, Lisa Cazares, Lilian vanvlerken-ysla, Sonja Hess

P04.03 PSMA, STIP1 and PPIA Detection in Urine-Derived Extracellular Vesicles in Multiple Cancers
Lyssa Dimapanat, Alex Rai

P04.04 Protein Signatures to Distinguish Aggressive from Indolent Prostate Cancer
Fernando Garcia-Marques, Shiqin Liu, Sarah M. Totten, Cheylene Tanimoto, Abel Bermudez, En-Chi Hsu, Amy Hembree, Tanya Stoyanova, James D. Brooks, Sharon J. Pitteri

P04.05 Proteomic profiling of IDH-mutant gliomas enables prediction of chromosomal copy number variations
Marius Felix, Dennis Friedel, Uwe Warnken, Damian Stichel, Christel Herold-Mende, Laura Heikaus, Andreas von Deimling, David E. Reuss

P04.06 The Prostate Cancer in vivo Secretome
Amanda Khoo, Zhuyu Qiu, Joseph J. Otto, Vladimir Ignatchenko, Andrew Macklin, Katharina Fritsch, Meinusha Govindarajan, Danny Vesprini, Julius O. Nyalwidhe, Stanley Liu, O. John Semmes, Paul C. Boutros, Thomas Kislinger

P04.07 Proteomic features of multiparametric magnetic resonance imaging visibility in prostate cancer
Amanda Khoo, Taylor Y. Sadun, Lydia Y. Liu, Kathleen Houlahan, Vladimir Ignatchenko, Aydin Pooli, Steven S. Raman, Anthony E. Sisk Jr., Paul C. Boutros, Robert E. Reiter, Thomas Kislinger

P04.08 Label-Free Proteomics Analysis of Prostate Cancer Tissue Samples on a New Orbitrap Mass Spectrometer with High-Field Asymmetric Waveform Ion Mobility Spectrometry
Daniel Lopez-Ferrer, Abel Bermudez, Fernando Garcia-Marques, Sharon Pitteri

P04.09 Oncogenic signaling drives histone lysine lactylation via the Warburg effect and MOF1 lactyltransferase activity
Di Zhang, Lu Yang, Ruilong Liu, Jinjun Gao, Kevin Huang, Ziyuan Li, He Huang, Yingming Zhao

P04.10 Proteomic Profiling of Microsatellite Instability in Colorectal Cancer
Fernando Tobias, Emily Sekera, Heather Hampel, Amanda Hummon

P04.11 PDX models reflect the proteome landscape of pediatric acute lymphoblastic leukemia but divert in select pathways
Anuli Uzozie, C. James Lim, Christopher A. Maxwell, Gregor S. D. Reid, Philipp F. Lange

P04.12 The Proteome of the Microbiome in Peripheral Fluids: Application to Breast Cancer Detection
Ngoc Vuong, Rayan Alhammad, Weidong Zhou, Neha Shetty, Jennifer May, Alessandra Luchini, Claudius Mueller, Virginia Espina, Lance Liotta, Richard Hoefer

P04.13 Combined Surface Proteome and Phosphoproteomics Reveals Novel Therapeutic Targets for Hepatocellular Carcinoma

Shao Huan Samuel Weng, Hsiang-En Hsu, Junmin Wang, Raghothama Chaerkady, Lisa Cazares, Elaine Hurt, Sonja Hess

P04.01

Selective Pulse Chase-SILAC Labeling of Three-Dimensional Multicellular Spheroids for Global Proteome Analysis

Nicole Beller (1), Jessica Lukowski (2), Amanda Hummon (1)

(1) Ohio State University, Columbus, Ohio, United States

(2) University of Notre Dame, Notre Dame, Indiana, United States

Accurate model systems are crucial to cancer research. Three-dimensional cell culture models have gained attention in the scientific community for their abilities to recapitulate the tumor microenvironment more accurately than cells grown in monolayers. Spheroids have distinct chemical and pathophysiological gradients which result in three distinct cellular populations: the necrotic core, a quiescent layer, and a proliferating layer. By utilizing stable isotope labeling in cell culture (SILAC), these layers can be defined, and their proteomic changes tracked in response to external agents. Utilizing HCT116 cells, spheroids were grown beginning with either heavy (Arginine, Lysine $^{13}C^{15}N$), medium (Arginine, Lysine $^{13}C^{14}N$), or light (Arginine, Lysine $^{12}C^{14}N$) media to form multicellular cores. Every four days of growth, labeled media were changed to distinctly label the three layers of the spheroid. After twelve days, the spheroids were harvested and underwent serial trypsinization followed by lysis. Subsequent sample preparation was conducted using S-Trap filters and the resulting peptides underwent off-line fractionation. Samples were analyzed via reversed-phased liquid chromatography coupled to a Q-Exactive HF. To confirm that the different labels do not alter the proteome of each of the layers, the spheroids were labeled with different concatenations of media and, after serial trypsinization, the individual populations were analyzed and compared to their differently labeled counterparts in similar spheroid regions. Our data confirms that a positive correlation exists between spheroid regions, regardless of the specific label present. These results indicate that the labels do not affect the proteome of the individual layers of the spheroid. Additionally, to better understand how the metabolic labeling affects growth, the spheroids were monitored throughout their 14-day growing period, and their diameters were measured. As we move forward in testing chemotherapeutics, the pulse-chase SILAC spheroid model will allow for a more accurate proteome depiction with greater insight into spatial distribution.

Contact: Nicole Beller, beller.25@osu.edu

P04.02

Quantitative proteomics reveals T-cells specific response from pancreatic cancer stem cells when compared to PANC1 in 2D and 3D cultures

Raghothama Chaerkady (1), Farah Alayli (2), Hsiang-En Hsu (1), Shao Huan Samuel Weng (1), Junmin Wang (1), Wen Yu (3), Lisa Cazares (1), Lilian vanvlerken-ysla (2), Sonja Hess (1)

(1) Dynamic Omics, ADPE, AstraZeneca, Gaithersburg, MD, USA

(2) Early Oncology, AstraZeneca, Gaithersburg, MD, USA

(3) Data Science & AI, AstraZeneca, Gaithersburg, MD, USA

Cancer stem cells (CSCs) are a small subset of tumor cells with unique properties responsible for chemotherapy resistance, escape from immune surveillance, relapse and metastasis. Understanding the distinct features of CSCs versus bulk tumor cells and their differences after coculture with activated T-cells can help in identifying potential therapeutic targets. This study is focused on differential protein expression profiling of PANC1 cell line enriched for cancer stem cells or grown in 2D and 3D culture conditions in the presence or absence of T-cells using tandem mass tag (TMT) labeling-based quantitative proteomics. We identified ~7,200 proteins from these 3 different culture conditions and nearly 10% of the proteins were differentially expressed in CSCs when compared to 2D and 3D cultured cells. These distinct differentially expressed proteins revealed different known and novel hallmarks of CSCs. For example, abnormal lipid metabolism including increased cholesterol synthesis enzymes, altered lipid metabolism and immune signaling proteins were observed, although the exact mechanism of action of these pathways in cancer still remains unclear. Aryl hydrocarbon receptor and Wnt signaling were among the important canonical pathways enriched in Ingenuity Pathway Analysis. Several key players of stemness in the Wnt signaling pathway including PPARgamma, SOX and T-cell factor/lymphoid enhancer factor (TCF/LEF) transcription factors were up-regulated in CSCs compared to 2D and 3D cultures. Other interesting affected pathways include oxidative phosphorylation and nuclear factor erythroid 2–related factor 2 (NRF2) mediated oxidative stress pathways. The top 5 proteins upregulated in CSCs with a potential role in metastasis and stemness include tubulin polymerization promoting protein family member 3 (TPPP3), cellular retinoic acid-binding protein 1 (CRABP1), basal cell adhesion molecule (BCAM), transforming growth factor beta 1 (TGFB1) and gelsolin (GSN). Taken together, this study provided multiple potential targets important to the biology of pancreatic CSCs in the presence of activated T-cells.

Contact: Raghothama Chaerkady, raghothama.chaerkady@astrazeneca.com

P04.03

PSMA, STIP1 and PPIA Detection in Urine-Derived Extracellular Vesicles in Multiple Cancers

Lyssa Dimapanat (1), Alex Rai (1)

(1) Department of Pathology and Cell Biology, New York, NY, 10032

Biofluids are a favorable sample type for the identification of novel cancer biomarkers. In particular, urine is convenient, non-invasively obtained, and available in large sample volumes. It can recapitulate biological information typically obtained from tissue biopsies. We are interested in exploring protein biomarkers in urine extracellular vesicles (EV). Recent studies have revealed the critical role of EV-mediated intercellular communication and in regulating cancer processes, including tumorigenesis and metastasis. We isolated urine EVs using an optimized workflow based on serial centrifugation. The presence of EVs was confirmed by detection of canonical EV markers (ALIX, flotillin, and ACTN4) using western blotting. Nanoparticle tracking analysis and transmission electron microscopy revealed particle distribution and EV size. Subsequently, we surveyed 117 cancer samples and 23 control samples for the presence of three candidate protein markers identified in our earlier studies: PSMA, STIP and PPIA. Cancer samples were categorized into five different groups: genitourinary (GU), gastrointestinal (GI), brain, lymphomas and leukemias. PSMA was detected in 62% (73/117) of samples across the various cancer types, and only 2 of 23 controls. Highest PSMA signals were obtained from GU cancers (prostate and bladder), with higher levels in metastatic disease. STIP1 was detectable in 29% (35/117), and PPIA in 5% (6/117) of the cancer samples. STIP1 was detected at highest levels in GI cancers (colorectal and liver), while PPIA was restricted to few samples from patients with higher tumor burden. Our candidate biomarkers are detectable in urine EVs and can fulfill different unmet clinical needs for patient management. We propose that PSMA has potential application for monitoring therapeutic response in GU cancers, STIP1 may be useful for GI cancer subtyping, and PPIA's restricted detection may define a subset of aggressive cancers.

Contact: Lyssa Dimapanat, lpd2125@cumc.columbia.edu

P04.04

Protein Signatures to Distinguish Aggressive from Indolent Prostate Cancer

Fernando Garcia-Marques (1), Shiqin Liu (1), Sarah M. Totten (1), Cheylene Tanimoto (1), Abel Bermudez (1), En-Chi Hsu (1), Amy Hembree (1), Tanya Stoyanova (1), James D. Brooks (2), Sharon J. Pitteri (1)

(1) Radiology, Canary Center at Stanford for Cancer Early Detection, Palo Alto, California, US

(2) Urology, Stanford University School of Medicine, Stanford, California, US

In 2020 there will be approximately 174,675 new prostate cancer cases in the US making it the second most common male cancer. Blood-based measurements of prostate-specific antigen (PSA) levels are used for prostate cancer screening, but most newly detected prostate cancer cases are localized, and low-grade. Distinguishing men with indolent versus aggressive disease remain a major clinical challenge with many men undergoing unnecessary treatment. Here we present proteomic analysis of surgically resected prostate cancer tissue (Gleason Score 3 to 5), and patient-matched normal tissue samples from 22 prostate cancer patients including 8 men with recurrent PSA levels post-surgery. We identified 24,037 unique peptides (FDR<1%) belonging to 3,313 proteins whose absolute abundance ranges seven orders of magnitude. After label-free quantification on proteins with at least two peptides identified in 20% of the samples, we detected 107 proteins ($P < 0.01$) with different levels in tissues from recurrent versus non-recurrent cancers. The proteome changes reflect important biological pathways in tissue that are integral to tumor initiation and progression. Interestingly, a systems biology analysis identified a network of complexes where ~24% of the proteins in the network were associated with the YY1 transcription factor (adjusted $P < 0.001$), potentially explaining, at molecular level, the differences found between recurrent and non-recurrent prostate cancer. On the basic of our LC-MS/MS results, we used immunohistochemistry on a tissue microarray containing 124 patient tissue samples to validate three proteins: POSTN, CALR, and CSTD, as potential indicators of prostate cancer progression. Kaplan-Meier survival curves show significant or near significant ($P = 0.027$, 0.041, and 0.015) correlation of positive staining of TMAs with decreased survival. The protein signature identified in this study has the potential to help distinguish aggressive from indolent prostate cancer.

Contact: Fernando Garcia-Marques, fjgarcia@stanford.edu

P04.05

Proteomic profiling of IDH-mutant gliomas enables prediction of chromosomal copy number variations

Marius Felix (1), Dennis Friedel (1), Uwe Warnken (2), Damian Stichel (1), Christel Herold-Mende (3), Laura Heikaus (4), Andreas von Deimling (1), David E. Reuss (1)

(1) Department of Neuropathology, Institute of Pathology, Heidelberg University Hospital and Clinical Cooperation Unit Neuropathology, German Cancer Research Center, German Consortium for Translational Cancer Research, Heidelberg, Germany

(2) Clinical Cooperation Unit Neurooncology, German Cancer Consortium, German Cancer Research Center, Heidelberg, Germany

(3) Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany

(4) Bruker Daltonick GmbH, Bremen, Germany

Recurrent chromosomal copy number variations (CNV) are hallmarks of different types of brain tumors and status determination is an integral part of WHO classification. There is need for a better understanding of the consequences of gains or deletions involving whole chromosomal arms. A prominent example are IDH-mutant gliomas which are separated in two distinct types based on the deletion of chromosomal arms 1p and 19q. Oligodendrogliomas IDH-mutant are 1p/19q co-deleted while astrocytomas IDH-mutant are not. Therefore, determination of 1p/19q is important for prognosis and therapy. Upon demonstrating that the dda-PASEF based analysis of formalin-fixed paraffin embedded (FFPE) tissue highly correlates with fresh frozen tissue we here present a workflow for in-depth proteomic profiling of IDH-mutant gliomas starting with small punches FFPE tissue. We've identified several novel potential protein biomarkers differentially regulated across two independent cohorts of oligodendroglioma and astrocytoma. Most strikingly, we were able to generate virtual copy number variation plots from the proteomic profile which we termed chromosomal protein abundance plots (CPAP). For CPAP we calculate the mean abundances of proteins grouped together by the chromosomal position of their coding genes and compared these to reference cohorts. Preliminary results from comparison of CNV plots from genome wide DNA methylation profiles and CPAP consistently displayed 1p/19q losses in oligodendrogliomas, as well as the majority of other whole chromosomal alterations present in individual tumors. This implies that gene dosage effects of whole chromosomal alterations are mirrored by the relative abundancy of numerous proteins. CPAP therefore, is a promising tool for the differentiation of tumors based on chromosomal copy number variation.

Contact: Laura Heikaus, laura.heikaus@bruker.com

P04.06

The Prostate Cancer in vivo Secretome

Amanda Khoo (1), Zhuyu Qiu (2), Joseph J. Otto (3), Vladimir Ignatchenko (4), Andrew Macklin (4), Katharina Fritsch (1), Meinusha Govindarajan (1), Danny Vesprini (5), Julius O. Nyalwidhe (6), Stanley Liu (5), O. John Semmes (6), Paul C. Boutros (2), Thomas Kislinger (7)

(1) Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

(2) Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles, California, USA

(3) Leroy T. Canoles Jr. Cancer Research Center, Eastern Virginia Medical School, Norfolk, Virginia, USA

(4) Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada

(5) Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

(6) Department of Microbiology and Molecular Cell Biology, Eastern Virginia Medical School, Norfolk, Virginia, USA

(7) 3Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada

Prostate cancer (PCa) is the most common cancer in men and can be detected early through screening of asymptomatic men. Most PCa are indolent at time of diagnosis and current prognostic protocols do not accurately predict disease aggression and clinical outcome, limiting optimal patient management. Elevated serum prostate-specific antigen (PSA) levels could be indicative of cancer or benign prostatic conditions, while needle biopsies are invasive and undersample the prostate, resulting in uncertainty of cancer grading. We hypothesize that urine collected after a digital rectal exam (post-DRE urine) contains proteins that serve as prognostic biomarkers and will allow for non-invasive PCa risk stratification. We have acquired proteomics data from 1,254 samples which include tumor tissue, adjacent-normal tissue, post-DRE urine, prostatic fluid, and urine-derived extracellular vesicles (EV). These samples were collected from a cohort 656 treatment-naïve patients that span the entire risk spectrum of prostate cancer (low, intermediate and high-risk). Data was acquired on various Orbitrap instruments (HF, Fusion, Fusion Lumos) in data-dependent acquisition mode at 2 sites. We detected a total of 9,220 proteins, 4,359 of which were detected in fluids (urine and prostatic fluid), and 5,322 in urine EVs. Using 66 matched samples, we confirmed that our sample preparation workflow is reproducible across sites (median R^2=0.75), processing replicates (median R^2=0.86) and technical replicates (median R^2=0.87). We developed a pipeline to select for prostate-derived proteins from the background of urine proteins from kidney and bladder by leveraging RNA-seq data from TCGA and GTEX. Using these data from prostatic fluid and urine datasets, we built a biomarker discovery pipeline to identify peptides in urine that can distinguish between low and intermediate/high-risk

patients, as well as benign and cancer patients. Finally, we characterized small and large EVs from matched urines (n=155) and identified protein biomarkers that are contained in EVs.

Contact: Amanda Khoo, amanda.khoo@mail.utoronto.ca

P04.07

Proteomic features of multiparametric magnetic resonance imaging visibility in prostate cancer

Amanda Khoo (1), Taylor Y. Sadun (2), Lydia Y. Liu (1), Kathleen Houlahan (1), Vladimir Ignatchenko (3), Aydin Pooli (2), Steven S. Raman (2), Anthony E. Sisk Jr. (2), Paul C. Boutros (2), Robert E. Reiter (2), Thomas Kislinger (3)

(1) Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

(2) Jonsson Comprehensive Cancer Center, University of California, Los Angeles, California, USA

(3) 3Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada

Multiparametric magnetic resonance imaging (mpMRI) has improved the identification of clinically significant prostate cancers (PCa), reducing under- and over-treatment of treatment-naïve men. However, approximately 20% of primary prostate tumors are invisible to mpMRI. Our collaborators have shown that specific genomic features drive MRI visibility of individual tumors, with MRI-visible tumors harboring features of tumor aggressiveness (Houlahan et al., 2019, European Urology). We complement this dataset with proteomics analysis of 40 matched tumors and adjacent-normal regions from intermediate-risk patients (Gleason score 3+4), building on the hypothesis that aggressive molecular hallmarks contribute to MRI visibility. Tumor foci and adjacent-normal regions from 10 µm thick whole-mount FFPE prostate sections were annotated by a urologic pathologist, and consecutive sections were macrodissected for proteomics. Sections were processed for mass spectrometry using two complementary methods (S-Trap and trifluoroethanol [TFE]-based extraction) to maximize proteome coverage. Samples were analyzed by shotgun proteomics using a 4-hour nano-flow UPLC gradient coupled to a Q-Exactive HF mass spectrometer operating in DDA mode. Raw files were searched with MaxQuant. We quantified 6,797 proteins, 4,656 of which were common to both extraction methods. Of these, 1,475 proteins were associated with features of tumor aggressiveness, such as genomic instability, hypoxia, and intraductal carcinoma/cribriform architecture (IDC/CA), a histomorphological growth pattern associated with adverse outcome. Proteins associated with these features were enriched in MRI-visible tumors compared to MRI-invisible tumors, and enriched in tumor regions compared to normal-adjacent regions. Pathway analysis revealed that MRI-visible tumors and tumor regions were enriched in androgen response pathway, while MRI-invisible tumors and tumor-adjacent normal were enriched in myogenesis and epithelial-mesenchymal transition pathways. Altogether, the proteomic data lends support to the hypothesis that MRI-visible tumors have more aggressive features, and that MRI-invisible tumors closely resemble tumor-adjacent normal.

Contact: Amanda Khoo, amanda.khoo@mail.utoronto.ca

P04.08

Label-Free Proteomics Analysis of Prostate Cancer Tissue Samples on a New Orbitrap Mass Spectrometer with High-Field Asymmetric Waveform Ion Mobility Spectrometry

Daniel Lopez-Ferrer (1), Abel Bermudez (2), Fernando Garcia-Marques (2), Sharon Pitteri (2)

(1) Thermo Fisher Scientific, San Jose, CA, USA

(2) Stanford University, Stanford, CA, USA

Introduction

LC-MS-based proteomics is a powerful technique for identification and quantification of proteins in a complex biological samples. Such analysis needs to provide robustness to analyze thousands of samples while providing optimal proteome coverage and quantitation performance. Here we demonstrate a large proteomics study of human tissue samples using the new Orbitrap Exploris 480 MS coupled to a High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) interface. The sensitivity and performance of the Orbitrap Exploris 480 MS was evaluated while analyzing 22 cancer tissue samples and with 22 matched adjacent benign tissues in addition to analyzing ten different cancer tissues with QA/QC runs throughout the whole analysis.

Methods

Human tissue samples were homogenized with 800 µL of lysis buffer consisting of 12.5 mM Tris pH8, 0.5 mM EDTA, and 7.5 M urea. A micro-bicichoninic acid (BCA) protein assay was performed to quantify proteins. 25 ug of total protein was digested with trypsin. The resulting tryptic peptides were desalted using C18 ZipTip pipette tips and analyzed by 120 min gradients on an EASY-nLC 1200 coupled to an Orbitrap Exploris 480 MS with FAIMS Pro interface. Label free, quantitation performance was evaluated to demonstrate instrument sensitivity and methods reproducibility across this large cancer cohort study.

Preliminary data

The performance of this new benchtop mass spectrometer was evaluated in a data-dependent acquisition (DDA) for sample injection amounts of 2 ug across 250 sample injections. This instrument sensitivity enables identification of over 5000 protein groups with high reproducibility across replicates.

Novel aspect

Robust and sensitive Orbitrap Exploris 480 MS, providing throughput and sensitivity needed for label-free proteomics analysis in large cohort studies.

Contact: Daniel Lopez-Ferrer, daniel.lopezferrer@thermofisher.com

P04.09

Oncogenic signaling drives histone lysine lactylation via the Warburg effect and MOF1 lactyltransferase activity

Di Zhang (1), Lu Yang (1), Ruilong Liu (2), Jinjun Gao (1), Kevin Huang (3), Ziyuan Li (4), He Huang (2), Yingming Zhao (1)

(1) Ben May Department for Cancer Research, The University of Chicago, Chicago, Illinois, USA

(2) Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, P.R. China

(3) St. Mark's School in Southborough, Southborough, Massachusetts, USA

(4) Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, USA

We recently reported the discovery of lysine lactylation (Kla), a novel protein posttranslational modification occurring on histones that alters gene expression in immune cells via the Warburg effect-associate lactate production. Despite the Warburg effect being typically observed in cancer, our understanding of histone Kla in this disease remains unknown. Here, we report that diverse oncogenic signals can promote histone Kla through Warburg effect driven lactate production. We showed that histone Kla in cancer cells is established primarily by the concentration of intracellular lactate. Using quantitative proteomics, we find that BRAFV600E drives a greater than 2-fold increase of Kla on 6 sites on histone H3 and H4 with little to no change observed for Kac. Comparative proteomic analysis reveals that the BRAVF600E mutation-driven histone Kla dynamics are distinct from histone Kla dynamics that are driven by hypoxia. Lastly, we report the discovery of MOF1 as a novel lactyltransferase and validate its lactyltransferase activity in cancer cells. Our work reports the hierarchical regulation of histone Kla via a novel MAPK-Warburg-MOF1 axis, suggesting that histone Kla is a non-stochastic epigenetic mark that is driven by oncogenic signaling.

Contact: Mathew Perez-Neut, mperezneut@uchicago.edu

P04.10

Proteomic Profiling of Microsatellite Instability in Colorectal Cancer

Fernando Tobias (1), Emily Sekera (1), Heather Hampel (2), Amanda Hummon (3)

(1) Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio, United States of America

(2) Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio, United States of America

(3) Department of Chemistry and Biochemistry, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio, United States of America

Colorectal cancer (CRC) is the third most common cancer diagnosed in both men and women in the United States. It is estimated for 2021, over 140,000 new cases of CRC will be diagnosed and will claim over 50,000 lives. Most CRC cases occur through environmental factors leading to mutations. However, a subset of cases is correlated to genetic influences. Many CRC cases are characterized as either microsatellite instable (MSI) or microsatellite stable (MSS). Microsatellites are short repetitive units of DNA sequences that occur throughout the genome, which are prone to insertions and deletions during DNA replication. Within MSI colorectal cancers, cases can be further categorized into either germline mutations, double somatic mutations, or epigenetic silencing. The most common MSI variant is through germline mutations known as Lynch Syndrome. Patients with Lynch Syndrome develop cancerous polyps at a higher rate and a younger age than other forms of CRC. Despite these classifications as MSS or MSI, downstream biochemical effects are still poorly understood. A quantitative proteomics study was previously conducted on 375 cell lines of the Cancer Cell Line Encyclopedia, which found multiple protein complexes altered that monitor DNA and RNA mutations or problems in translation. In a blinded study, we are profiling the protein expressions from fresh frozen tissue samples of twenty-four patients representing four cohorts with CRC. Each patient is represented by a normal mucosa and a tumor sample. The detection of differences in expressed proteins in the tumor samples by data-independent acquisition (DIA) mass spectrometry will help to elucidate the molecular mechanisms of MSI. Furthermore, it will establish a baseline profile for a more extensive and larger cohort analysis of the same CRC variants.

Contact: Fernando Tobias, tobias.62@osu.edu

P04.11

PDX models reflect the proteome landscape of pediatric acute lymphoblastic leukemia but divert in select pathways

Anuli Uzozie (1), C. James Lim (2), Christopher A. Maxwell (2), Gregor S. D. Reid (2), Philipp F. Lange (1)

(1) Department of Pathology, University of British Columbia, Vancouver, BC, Canada, Vancouver, BC, Canada

(2) Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada, Vancouver, BC, Canada

Murine xenografts of pediatric leukemia accurately recapitulate genomic aberrations. How this translates to the functional capacity of cells remains unclear. Here, we studied global protein abundance, phosphorylation, and protein maturation by proteolytic processing in 11 pediatric B- and T- cell ALL patients and 19 corresponding xenografts.

Xenograft models were generated for each pediatric patient leukemia. Mass spectrometry-based methods were used to investigate global protein abundance, protein phosphorylation, and limited proteolysis in paired patient and xenografted pediatric acute B- and T- cell lymphocytic leukemia, as well as in pediatric leukemia cell lines. Targeted next-generation sequencing was utilized to examine genetic abnormalities in patients and in corresponding xenografts. Bioinformatic and statistical analysis were performed to identify functional mechanisms associated with proteins and protein post-translational modifications.

Overall, we found xenograft proteomes to be most equivalent with their patient of origin. Protein level differences that stratified disease subtypes at diagnostic and relapse stages were largely recapitulated in xenografts. As expected, PDXs lacked multiple human leukocyte antigens and complement proteins. We found increased expression of cell cycle proteins indicating a high proliferative capacity of xenografted cells. Structural genomic changes and mutations were reflected at the protein level in patients. In contrast, the post-translational modification landscape was shaped by leukemia type and host and only to a limited degree by the patient of origin. Of 201 known pediatric oncogenic drivers and drug-targetable proteins, the KMT2 protein family showed consistently high variability between patient and corresponding xenografts. Comprehensive N terminomics revealed deregulated proteolytic processing in leukemic cells, in particular from caspase-driven cleavages found in patient cells.

Genomic and host factors shape protein and post-translational modification landscapes differently. This study highlights select areas of diverging biology while confirming murine patient-derived xenografts as a generally accurate model system for B- and T-ALL pediatric leukemia.

Contact: Anuli Uzozie, auzozie@bcchr.ca

P04.12

The Proteome of the Microbiome in Peripheral Fluids: Application to Breast Cancer Detection

Ngoc Vuong (1), Rayan Alhammad (1), Weidong Zhou (1), Neha Shetty (2), Jennifer May (3), Alessandra Luchini (1), Claudius Mueller (1), Virginia Espina (1), Lance Liotta (1), Richard Hoefer (3)

(1) Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, VA, United States

(2) College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States

(3) Sentara Dorothy G. Hoefer Comprehensive Breast Center, Newport News, VA, United States

Proteomics holds the promise to yield functional information on the human microbiome, thus complementing taxonomic abundance information captured by genomic analysis. Sample complexity, including hundreds of species and hundreds of thousands of proteins, and bioinformatics challenges related to the large size and poor accuracy of protein databases, have hampered the progress of microbiome proteome analysis. In this study, we have applied affinity enrichment, mass spectrometry analysis and a novel bioinformatics pipeline to analyze the microbiome proteome in saliva and plasma specimens collected from N=151 breast cancer patients and controls. Peptide specificity was defined by sequence identity with an organism at the genus or species level using a sliding homology threshold, and taxonomic verification for related clades. We identified 11,505 peptides and 2203 genera of microbes. Calcium oxalate, hydroxyapatite and carbonate metabolizing bacteria (e.g., Bacillus, Lactobacillus and Bacteroides) were differently represented in cases and controls, suggesting a correlation with the higher carbonate content found in microcalcifications of women with malignant breast lesions when compared to benign controls. This study supports the notion that, although blood has been traditionally considered a sterile environment, plasma contains microbiota-derived peptides, possibly shed into the circulation by microbes hosted in different organ systems.

Contact: Ngoc Vuong, nvuong3@gmu.edu

P04.13

Combined Surface Proteome and Phosphoproteomics Reveals Novel Therapeutic Targets for Hepatocellular Carcinoma

Shao Huan Samuel Weng (1), Hsiang-En Hsu (1), Junmin Wang (1), Raghothama Chaerkady (1), Lisa Cazares (1), Elaine Hurt (2), Sonja Hess (1)

(1) APDE, AstraZeneca, Gaithersburg, MD, USA

(2) Early Oncology, AstraZeneca, Gaithersburg, MD, USA

Hepatocellular carcinoma (HCC) is the most common primary liver cancer, and the third most common cause of cancer related deaths worldwide. Limited therapeutic options are available for advanced HCC and resistance to sorafenib, the only currently available treatment, has been reported. Therefore, the discovery of novel therapeutic targets is crucial to improve the survival of patients with HCC. To address this challenge, we combined the enrichment of the surface proteome of HCC tissues with advanced phospho-proteomics to facilitate the identification of target antigens for therapeutic development and to deconvolute the complex biological pathways which promote HCC growth and metastasis. We systematically applied different surface protein extraction techniques and automated phosphopeptide enrichment via a Kingfisher Flex robotic device using iron-immobilized metal affinity beads. The performance of 3 different membrane protein isolation techniques; ultracentrifugation, Mem-PER™ Plus Membrane Protein Extraction Kit, and Minute™ Total Protein Extraction Kit were compared to enrich for surface proteins in pooled HCC and normal liver tissues. Overall, the largest number of membrane proteins was identified from the Minute™ Total Protein Extraction Kit. From 30 mg of tissue samples for surface protein enrichment, we identified ~3,300 total proteins where ~900 proteins were associated with plasma membrane and ~260 proteins matched with the stringent surface protein database SURFY. After Fe-IMAC phosphopeptide enrichment from 50 mg of tissue, we identified a total of 13,000 phosphosites in HCC tissue samples; 97 proteins exhibited significantly altered levels of phosphorylation when compared to normal tissues. Several kinase-signaling pathways involved in metabolism and cell signaling were found to be dysregulated in HCC patient samples. This includes RhoA, mTOR, ERK, MAPK and fatty acid beta-oxidation pathways. These findings suggested that multifaceted proteomics approaches can aid in revealing candidate drug targets for HCC and help in designing combination of drugs for therapeutic approaches.

Contact: Shao Huan Samuel Weng, samuel.weng@astrazeneca.com

Cardiovascular Proteomics

P05.01

Top-Down Proteomics of Myofilaments in Swine Hearts throughout Postnatal Development

Timothy Aballo (1), Elizabeth Bayne (2), Wuqiang Zhu (3), Meng Zhao (3), Ahmed I. Mahmoud (1), Jianyi Zhang (3), Ying Ge (4)

(1) Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, United States of America
(2) Department of Chemistry, University of Wisconsin-Madison, Madison, WI, United States of America
(3) Department of Biomedical Engineering, University of Alabama at Birmingham, Birmingham, AL, United States of America
(4) Department of Cell and Regenerative Biology, Department of Chemistry, University of Wisconsin-Madison, Madison, WI, United States of America

The neonatal swine heart possesses an endogenous ability to regenerate after injury, but this regenerative capacity is lost shortly after birth. As the cardiac myofilaments play a key role in the proliferative capacity of cardiomyocytes, it is imperative to study how the myofilament sub-proteome changes throughout postnatal development. Herein, we employ quantitative top-down mass spectrometry to analyze myofilament proteoforms, the

functionally diverse protein species that arise from a single gene due to mRNA splicing and post-translation modification (PTM), at different stages of postnatal development in the swine heart.

In this study, we observed a decrease in expression of slow skeletal troponin I (ssTnI) and an increase in the expression of cardiac troponin I (cTnI) throughout postnatal development. Similarly, we detected a decrease in the expression of myosin light chain 1 and 2 atrial isoforms (MLC-1a and MLC-2a) coupled with an increase in ventricular isoform (MLC-1v and MLC-2v) expression. As the isoform switches of ssTnI to cTnI, of MLC-1a to MLC-1v, and of MLC-2a to MLC-2v have been previously identified as markers of cardiomyocyte maturation in other model systems, these data confirmed this method for assessment of swine heart maturation. Interestingly, we also detected a decrease in the relative abundance of cardiac troponin T isoform 2 (cTnT2); decreased expression of cTnT2 has been implicated in reduced myofibril calcium sensitivity, which could implicate sarcomere rigidity and the proliferative potential of cardiomyocytes. Additionally, we observed a decrease in the relative abundance of phosphorylated α-Tropomyosin (α-Tpm), consistent with a previous study that revealed a decrease in α-Tpm phosphorylation during the maturation of human embryonic stem cell-derived cardiomyocytes.

Overall, we identified altered expression and PTM of important myofilament regulatory proteins throughout postnatal swine heart development. Using this, we hope to further investigate the molecular mechanisms governing the proliferative capacity of cardiomyocytes during endogenous heart regeneration.

Contact: Timothy Aballo, aballo@wisc.edu

P05.02

CellSurfer Platform for semi-automated cell surface N-glycoprotein profiling of human primary cells reveals chamber-specific cardiomyocyte surface maps

Linda Berg Luecke (1), Rebekah Gundry (2)

(1) Department of Biochemistry, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

(2) CardiOmics Program, Center for Heart and Vascular Research; Division of Cardiovascular Medicine; and Department of Cellular and Integrative Physiology, University of Nebraska Medical Center, Omaha, Nebraska, USA

In the heart, cell surface glycoproteins in cardiomyocytes (CM) are essential for sustaining normal cardiac function by facilitating proper conduction of action potentials and subsequent contraction of the myocardium. Improper function and/or expression of cell surface proteins can result in life-threating disorders and sudden cardiac death. Therefore, a variety of cell surface receptors have been targeted in pharmacological therapies for common cardiovascular diseases (e.g. hypertension, angina, arrhythmias, and heart failure). Despite the critical roles that cell surface glycoproteins play in cardiac biology and disease, a detailed cell type- or chamber-resolved view of the cell surface proteome of normal adult human heart does not exist yet.

Here, we report the development of a new analytical platform, CellSurfer, which integrates efficient sample handling, data independent acquisition, and streamlined data analysis workflows for rapid, quantitative discovery of cell surface N-glycoproteome. CellSurfer includes a semi-automated sample preparation workflow, Microscale Cell Surface Capture (uCSC) that enables discovery of >400 N-glycosylated surface proteins from less than <1 mg total cellular proteins (1-10 million cells).

Integrating CellSurfer with an optimized strategy for isolating intact CM from human heart tissue resulted in the generation of the first chamber-specific map of the cell surface N-glycoproteome of adult human CM. Overall, >900 cell surface N-glycoproteins were identified, including transmembrane, GPI-anchored, and extracellular matrix proteins. Differential analysis of CM from each major myocardial chamber revealed

previously unreported differences between left and right sides. These data will enhance our understanding of the distinct phenotypic fingerprint of CM that reside in each of the four chambers of the human heart and represent the first major step towards a comprehensive, cell-type, subtype, and chamber-resolved reference map of cell surface N-glycoproteins in the adult human heart and reveal new potential targets for immunophenotyping, in vivo imaging, and drug delivery.

Contact: Linda Berg Luecke, lbergluecke@mcw.edu

P05.03

Quantitative Analysis Of Cell Type and Chamber Specific Cardiomyocytes

Chase Castro (1), Melinda Wojtkiewicz (1), Rebekah Gundry (1)

(1) University of Nebraska Medical Center, Omaha, Nebraska, USA

Cardiomyocytes (CM) comprise one-third of the cells in the human heart and play a critical role in sustaining normal cardiac function by facilitating proper contraction and allowing the heart to pump blood throughout the body. Previous proteomics approaches have focused on whole heart tissue, which is comprised of a combination of various cell types and a significant abundance of extracellular matrix proteins. Despite being the foremost cell type in the heart, detailed proteomics data does not exist for isolated human CM.

In this study, we applied state-of-the-art mass spectrometry approaches to discover and quantify proteins of isolated human cardiomyocytes from the left ventricle, right ventricle, and apex. We compared cell type specific results to whole tissue lysate of matched donors and the resulting peptides were utilized in a data independent acquisition on a Thermo 480 Exploris, which allows rapid, quantitative discovery of proteins of interest. Whole tissue lysates from matched donors were processed identically allowing for direct comparison of the isolated cardiomyocytes to whole tissue lysates.

To date, we have identified >7000 proteins, including those with known relevance to cardiomyocytes (e.g. TNNI, TPM1, TPM3). Analyses reveal differences among chambers and highlight the expanded depth of coverage when using isolated cardiomyocytes in comparison to whole tissue. Overall, this work allowed for the advancement of a cardiomyocyte isolation protocol and revealed new cell type specific molecular targets.

Contact: Chase Castro, ccastro@unmc.edu

P05.04

Ultrasensitive Capillary Electrophoresis Ion Mobility Mass Spectrometry for Quantifying Angiotensin Peptides in the Brain

Kellen DeLaney (1), Sam Choi (1), Zhe Yu (2), Paul Marvar (2), Peter Nemes (1)

(1) Department of Chemistry & Biochemistry, University of Maryland College Park, College Park, MD, United States

(2) Department of Pharmacology and Physiology, The George Washington University, Washington, DC, United States

The renin-angiotensin system (RAS) is one of the prominent pathways implicated in cardiovascular homeostasis and pathophysiology. However, little is currently known about how these peptides are expressed in critical tissue in the brain. The major limitation in studying these peptides has been a lack of sufficiently sensitive analytical technologies capable of quantifying the angiotensin (Ang) peptides involved. Analyzing the expression of Ang peptides is challenging due to their low abundance and the limited sample amounts of relevant tissue regions. In our previous studies, we have developed a capillary electrophoresis (CE) mass spectrometry (MS) method with sufficient sensitivity (~700 zmol) to detect these peptides in mouse brain tissue regions. However, it remains to be determined whether these peptides originate in the brain tissue or are transported by the circulatory system. This study improves upon our previous CE-MS method to improve sensitivity in order to detect and quantify Ang peptides in perfused tissue (i.e. in the absence of blood). An on-capillary pre-concentration method was developed to increase the amount of sample that can be analyzed in a single run by approximately 7-fold, enabling lower overall concentrations to be detected. With the developed method, limits of concentration of approximately 30 zmol were achieved on a Bruker timsTOF Pro mass spectrometer, a 15-fold improvement over our previous design. This method was used to detect Ang peptides in samples of 3 pooled brain tissue of perfused mice, resulting in detectable signals of all 6 Ang peptides critical to the RAS pathway. These results provide a better understanding of both how peptides are synthesized and transported in the nervous system and the role of RAS peptides in critical regulatory brain tissue regions for cardiovascular homeostasis. Furthermore, the high sensitivity of this method can readily be applied to studying peptides involved in other critical pathways.

Contact: Kellen DeLaney, delaney1@umd.edu

P05.05

Cardiac sex disparities are established via post-transcriptional regulation

Josiah Hutton (1), Wei Shi (2), Xinlei Sheng (1), Kerry Dorr (2), Haley Davies (2), Tia Andrade (2), Todd Grecco (1), Yutaka Hashimoto (1), Joel Federspiel (1), Zachary Robbe (2), Xuqi Chen (3), Arthur Arnold (3), Ileana Cristea (1), Frank Conlon (2)

(1) Department of Molecular Biology, Princeton University, Princeton, NJ, United States

(2) Department of Biology and Genetics, McAllister Heart Institute, UNC-Chapel Hill, Chapel Hill, NC, United States

(3) Department of Integrative Biology & Physiology, University of California, Los Angeles, CA, United States

Male and female disease states differ in their prevalence, treatment responses, and survival rates. In cardiac disease, women almost uniformly fare far worse than men. Though sex plays a critical role in cardiac disease, the mechanisms underlying sex differences in cardiac homeostasis and disease remain unexplained. Here, we reveal sex-specific cardiac transcriptomes and proteomes and show that cardiac sex differences are predominately controlled via post-transcriptional mechanisms. Specifically, using RNA-Seq and either tandem mass tags (TMT) or label-free mass spectrometry quantification, we analyzed hearts dissected from the Collaborative Cross (CC), Four Core Genotype (FCG), and embryonic day 9.5 (E9.5) mouse model systems. Bayesian analysis of the resulting datasets allowed determining cardiac sex differences at the transcript and protein levels. In the context of the global genetic diversity of the CC model, a surrogate for human diversity, we characterized differential sex-specific enriched cardiac proteins. Utilizing the CORUM database and Gene Ontology, we determined that these differentially regulated proteins were associated with the differential regulation of protein complexes and pathways associated with biological sex processes. For example, metabolic pathways exhibited male heart bias, and the insulin induced pathway protein complex, a known pathway involved in female biased cardiac disease, was significantly upregulated in female hearts. Using the FCG model system, we show further sex differences in cardiac protein expression are established by both hormonal and sex chromosomal mechanisms. We determined the onset of sex-biased protein expression and discovered that sex disparities in heart tissue occur at the earliest stages of heart development, during the period preceding mammalian gonadal development. This may explain why congenital heart disease, a leading cause of death whose origin is often developmental, is sex biased. Our results reveal the molecular foundations for the differences in cardiac tissue that underlie sex disparities in health, disease, and treatment outcomes.

Contact: Josiah Hutton, josiahh@princeton.edu

P05.06

Small-Volume Lipoprotein(a) Isolation Guide for Clinical Research

Paul Mueller (1), Sara Rosario (1), Paige Bergstrom (1), Elisabeth Yerkes (1), Nathalie Pamir (1)

(1) Knight Cardiovascular Institute Oregon Health & Science University, Portland, OR, US

Lipoprotein(a) [Lp(a)] is an independent risk factor for cardiovascular disease, and, despite its similarity to low-density lipoprotein (LDL), it does not respond to LDL cholesterol lowering therapies. Lp(a) is only found in humans and non-human primates; thus, isolation from human plasma is the most accessible method to study Lp(a) biology. Previous reproducible isolation techniques require large volumes of plasma (>13mL) further compounding the difficulty in studying Lp(a). We demonstrate a small-volume isolation technique capable of purifying Lp(a) from as little as 0.4mL of human plasma suitable for clinical and translational research.

Plasma was collected from whole blood from donors (n=9) presenting to the Center for Preventive Cardiology clinic at the Oregon Health & Science University with circulating Lp(a) ≥ 50mg/dL. Plasma (0.4mL) was subjected to a two-step density-gradient ultracentrifugation method to isolate purified lipoprotein fractions for Lp(a), LDL and VLDL. HDL was also isolated using established methods. DDA LC-MS/MS using Thermo Q-Exactive HF was performed on each lipoprotein fraction and a comparative analysis was performed to identify proteomic qualities unique to Lp(a).

Our small-volume Lp(a) isolation technique yielded an average concentration of 264 ± 82.3 ug/mL which totaled 61.5 ± 17.2 ug in 244 ± 54.4uL and correlated positively with donor Lp(a) plasma concentrations (r^2=0.75; $p<0.01$). Lipoprotein fraction purity was confirmed visually using cholesterol gel electrophoresis. Proteome analysis revealed a total of 32 proteins on Lp(a) including APOB, LPA, and PLEK, of which LPA was unique to Lp(a). Pathway analyses identified Lp(a) proteins belonging to a diverse set of biological pathways including lipid binding and transport, cholesterol metabolism, complement and coagulation cascade, and various immunoglobulin heavy chain proteins.

Our novel Lp(a) isolation technique coupled with DDA LC-MS/MS reveals a diverse proteome on Lp(a) capable of mediating a broad spectrum of biological functions.

Contact: Paul Mueller, muellerp@ohsu.edu

P05.07

The Effects of Fish Oil Supplementation on High-density Lipoprotein Efflux Capacity, Proteome, Lipidome and Inflammatory Effects

Sara Rosario (1), Paul Mueller (1), Melissa Heard (1), Paige Bergstrom (1), Rayna Gasik (1), Nathalie Pamir (1)

(1) Center for Preventive Cardiology Oregon Health & Science University, Portland, Oregon, United States

Interest has grown in using fish oil treatment in cardiovascular disease (CVD) prevention. High dose eicosapentaenoic acid (EPA) has become an FDA approved treatment for high CVD risk patients. Although the mechanism is unknown, fish oil supplementation alters plasma lipoproteins. We hypothesize that the CVD risk protection is associated with beneficial changes in HDL biology. Healthy subjects (N=7) took fish oil supplementation for 30 days (1,125mg EPA and 875mg docosahexaenoic acid (DHA)). Serum was collected at baseline, at 30 days, and again 30 days after the last dose. Ultracentrifugation isolated HDL's sterol efflux and anti-inflammatory capacity, proteome, and lipidome were measured. Anti-inflammatory capacity was determined by incubating subject HDL with an LPS-treated human macrophage cell line. Proteome was measured by DDA LC-MS/MS with N15[APOA1] (the major structural protein of HDL) spiked HDL using Thermo Q-Exactive HF. Lipidome was measured by UPLC-QTOF on a Sciex 5600 TripleTOF instrument. Fish oil supplementation did not significantly change isolated HDL efflux, but did improve efflux mediated by APOB depleted serum (1.783% p=0.001). HDL's sterol efflux capacity inversely correlated its ability to stimulate IL10 (r^2=0.74 p=0.01) and IL6 (r^2=0.63 p=0.03) expression. We identified fourteen proteins that have significantly ($p<0.05$) changed following fish oil supplementation of which IGHA1, IGHG1, APOD and PLTP were down regulated ($P<0.05$). Thirty day fish oil supplementation increased 5 cholesterol ester (CE) species in HDL of all subjects, CE 18:2 and CE 20:5 increased by 54% (p=0.037) and 483% (p=0.002) respectively. Changes in HDL function track with changes in the HDL proteome and translate into changes in its lipidome, which might alleviate cardiovascular risk burden.

Contact: Sara Rosario, rosarios@ohsu.edu

Cellular Signaling and Systems Biology

P06.01

Proteo-Metabolomic Study of the Spemann's Organizer In the Vertebrate (Frog) Embryo

Aparna B. Baxi (1), Vi M. Quach (2), Jie Li (2), Peter Nemes (2)

(1) DepaDepartment of Chemistry and Biochemistry, University of Maryland, College Park, College Park, MD, USA

(2) Department of Chemistry and Biochemistry,University of Maryland, College Park, College Park, MD, USA

The Spemann's organizer (SO) in the frog (Xenopus laevis) embryo is an important signaling center that guides the induction of the nervous system. Our current understanding of biomolecules in this tissue has been limited to genes, transcripts, and some proteins. Recent advances in mass spectrometry (MS) have highlighted the complex correlation of transcript and protein abundances during embryonic development. Moreover, advances in small molecule analysis have revealed non-canonical roles in cell signaling, raising the question how metabolites and proteins coordinate development of the organizer.

Here, we perform dual proteo-metabolomic characterization of this tissue in the early developing embryo. Early embryonic precursors of the SO were fluorescently labeled to guide dissection of the tissue during gastrulation. The labeled SO was dissected and collected along with the remaining tissue from the embryo. The samples were yolk-depleted and processed for bottom-up proteomics. The resulting peptides were barcoded (tandem mass tags), reversed-phase fractionated at high pH, and the fractions analyzed by low-pH nano-LC/MS with multi-notch MS3 (Orbitrap). Ultimately, ~4,700 protein groups were identified (<1% FDR). Gene ontology and pathway analyses of the proteomic differences between the SO and the rest of the embryo revealed an enrichment of mitochondrial proteins in the SO. To further understand the implication of enriched mitochondrial proteins in the SO, we developed an LC-MS based targeted metabolomic approach to quantify metabolites associated with cellular respiration (timsTOF Pro). The quantitative data on proteins and metabolites produced in this critical signaling center presents a potential to advance our basic molecular understanding of vertebrate embryonic patterning.

Contact: Aparna Baxi, abaxi1@gwu.edu

P06.02

DAGBagM: Learning directed acyclic graphs of mixed variables with an application to identify prognostic protein biomarkers in ovarian cancer

Shrabanti Chowdhury (1), Ru Wang (2), Qing Yu (3), Catherine Huntoon (4), Larry Karnitz (4), Scott Kaufmann (4), Steven P. Gygi (3), Michael J. Birrer (5), Amanda G. Paulovich (6), Jie Peng (2), Pei Wang (1)

(1) Icahn School of Medicine at Mount Sinai, New York, NY, USA

(2) University of California Davis, Davis, CA, USA

(3) Harvard Medical School, Boston, MA, USA

(4) Mayo Clinic, Rochester, MN, USA

(5) University of Arkansas for Medical Sciences, Little Rock, AR, USA

(6) Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Inferring gene or protein regulatory network based on -omics profiles is commonly pursued in biological and medical studies. Specifically, directed acyclic graph (DAG) constructed based on proteogenomic data has recently been used to screen for causal biomarkers of clinical outcomes. Despite initial promises shown in these efforts, there remains unsolved challenges in DAG learning to jointly model clinical outcome variables, which often take binary values, and biomarker measurements, which usually are continuous variables. To that direction, we propose a new tool, DAGBagM, to learn directed acyclic graphs with both continuous and binary nodes (mixed variables) through bootstrap aggregation. By using appropriate models for continuous and binary variables, DAGBagM allows for either type of nodes to be parents or children nodes in the learned graph. DAGBagM also employs a bootstrap aggregating strategy to reduce false positives and to achieve better estimation accuracy. Moreover, in addition to reduce false positives, the aggregation procedure also provides a flexible framework to robustly incorporate prior information on edges for DAG reconstruction. As shown by simulation studies, DAGBagM performs better in identifying edges between continuous and binary nodes, as compared to commonly used strategies of either treating binary variables as continuous or discretizing continuous variables. Moreover, DAGBagM outperforms several popular DAG structure learning algorithms including the score-based hill climbing algorithm (HC), the constraint-based PC-algorithm (PC-alg), and the hybrid method max-min hill climbing (MMHC). When we apply DAGBagM to proteomics datasets from ovarian cancer studies, we identify potential prognostic protein biomarkers in ovarian cancer. The HC implementation in the R package DAGBagM is much faster than that in a widely used DAG learning R package bnlearn.

Contact: Shrabanti Chowdhury, shrabanti.chowdhury@mssm.edu

P06.04

PARP1 inhibitors trigger innate immunity via PARP1 trapping-induced DNA damage response

Chiho Kim (1), Xu-Dong Wang (1), Yonghao Yu (1)

(1) Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, US

It is being increasingly appreciated that the immunomodulatory functions of PARP1 inhibitors (PARPi) underlie their clinical activities in various BRCA-mutated tumors. PARPi possess both PARP1 inhibition and PARP1 trapping activities. The relative contribution of these two mechanisms toward PARPi-induced innate immune signaling, however, is poorly understood. We find that the presence of the PARP1 protein with uncompromised DNA-binding activities is required for PARPi-induced innate immune response. The activation of cGAS-STING signaling induced by various PARPi closely depends on their PARP1 trapping activities. Finally, we show that a small molecule PARP1 degrader blocks the enzymatic activity of PARP1 without eliciting PARP1 trapping or cGAS-STING activation. Our findings thus identify PARP1 trapping as a major contributor of the immunomodulatory functions of PARPi. Although PARPi-induced innate immunity is highly desirable in human malignancies, the ability of 'non-trapping' PARP1 degraders to avoid the activation of innate immune response could be useful in non-oncological diseases.

Contact: Chiho Kim, Chiho.Kim@UTSouthwestern.edu

P06.05

Proteome Dynamics Analysis Identified Hypoxia-mediated DNA Damage Response Pathways in Cancer

Ang Luo (1), Haiping Ouyang (1), Yao Gong (1), Yue Chen (1)

(1) Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN, USA

Mechanistic understanding of hypoxia-responsive signaling pathways provides important insights into oxygen- and metabolic-dependent cellular phenotypes in diseases. Using SILAC-based quantitative proteomics, we provided a quantitative map identifying over 6300 protein groups in response to hypoxia in prostate cancer cells and identified both canonical and novel cellular networks dynamically regulated under hypoxia. Particularly, we identified SDE2, a DNA stress response modulator, that was significantly downregulated by hypoxia, independent of HIF transcriptional activity. Mechanistically, hypoxia treatment promoted SDE2 polyubiquitination and degradation. Such regulation is independent of previously identified Arg/N-end rule proteolysis or the ubiquitin E3 ligase, CDT2. Depletion of SDE2 increased cellular sensitivity to DNA damage and inhibited cell proliferation. Interestingly, either SDE2 depletion or hypoxia treatment potentiated DNA-damage induced PCNA monoubiquitination, a key step for translesion DNA synthesis. Furthermore, knockdown of SDE2 desensitized, while overexpression of SDE2 protected the hypoxia-mediated regulation of PCNA monoubiquitination upon DNA damage. Our large-scale ubiquitination proteomics study further confirmed hypoxia-regulated PCNA monoubiquitination and DNA damage response signaling in cancer cells. Taken together, our quantitative proteomics and biochemical study revealed diverse hypoxia-responsive pathways that strongly associated with tumorigenesis under hypoxic tumor microenvironment and identified the functional roles of SDE2 and hypoxia in regulating DNA damage-induced PCNA monoubiquitination, suggesting a possible link between hypoxic microenvironment and the activation of error-prone DNA repair pathway in tumor cells.

Contact: Haiping Ouyang, ouyan137@umn.edu

P06.06

Cell Type-Selective Secretome Profiling In Vivo

Nicholas Riley (1), Wei Wei (2), Andrew Yang (3), Joon Kim (2), Stephanie Terrell (2), Veronica Li (2), Marta Garcia-Contreras (2), Carolyn Bertozzi (1), Jonathan Long (2)

(1) Department of Chemistry and ChEM-H, Stanford University, Stanford, CA, USA

(2) Department of Pathology and ChEM-H, Stanford University, Stanford, CA, USA

(3) Department of Bioengineering, Stanford University, Stanford, CA, USA

Secreted polypeptides are a fundamental axis of intercellular and endocrine communication. However, a global understanding of the composition and dynamics of cellular secretomes in intact mammalian organisms has been lacking. Here, we introduce a proximity biotinylation strategy that enables labeling, detection and enrichment of secreted polypeptides in a cell type-selective manner in mice. We generate a proteomic atlas of hepatocyte, myocyte, pericyte and myeloid cell secretomes by direct purification of biotinylated secreted proteins from blood plasma. Our secretome dataset validates known cell type-protein pairs, reveals secreted polypeptides that distinguish between cell types and identifies new cellular sources for classical plasma proteins. Lastly, we uncover a dynamic and previously undescribed nutrient-dependent reprogramming of the hepatocyte secretome characterized by the increased unconventional secretion of the cytosolic enzyme betaine–homocysteine S-methyltransferase (BHMT). This secretome profiling strategy enables dynamic and cell type-specific dissection of the plasma proteome and the secreted polypeptides that mediate intercellular signaling.

Contact: Nicholas Riley, nmriley@stanford.edu

Chemical Proteomics and Drug Discovery

P07.01

Label free quantitative pharmacoproteomics enabled the discovery of direct thrombin inhibitors by monitoring changes in the protein expression profiles in human platelets

Cristina Clement (1), Anna Babinska (2), Janet Gonzalez (3)

(1) Radiation Oncology, Weill Cornell Medicine, NY, NY, USA

(2) Medicine, University of New York, Downstate Medical Center, Brooklyn, NY, USA

(3) Natural Sciences, LaGuardia Community College, CUNY, Queens, NY, USA

Pharmacoproteomics uses advanced proteomic technologies for promoting drug discovery and development by highlighting protein expression profiles of diverse cellular and molecular pathways in response to different drug treatment. The research presented herein highlights the discovery of novel anticoagulant tetrapeptides, anti-thrombin direct inhibitors (DTI), aided by the development of label free quantitative proteomics analysis of changes in the protein expression profiles in human platelets. We hypothesized that a quantitative proteomics analysis would reveal that the DTIs treatment would induce a significant down-regulation of integrin mediated signaling triggered by protease activated receptors (PAR 1 and 4) in platelets, known to be the major pathway involved in the thrombin-mediated platelets aggregation. The proteomics platform employed the nanoLC-ESI MS/MS sequencing of tryptic/Glu-C/Lys-C generated peptides from whole lysates of DTIs treated and untreated human platelets from healthy donors. We employed a Q-Exactive quadrupole orbitrap mass spectrometer coupled with the comparative label free quantification (LFQ) analysis to further monitor the changes in the landscape of protein expression profiles. A combination of bioinformatics approaches empowered by the ingenuity pathway analysis (IPA), and curated databases containing platelets proteomes, including the “Adhesome”, “Exocarta”, “Reactome” and “PlateletWeb” was used to highlight the molecular and cellular pathways regulated by thrombin-activated platelets, in presence and absence of DTIs, including dabigatran, a FDA-approved anticoagulant. Our hypothesis was confirmed by the bioinformatics analysis which predicted that many proteins involved in ILK and integrin signaling pathways, the actin-mediated cell signaling and cellular movement, including rhoA and rhoGDI signaling were at least two-fold down-regulated ($p<0.05$, ANOVA) in the DTIs and dabigatran-treated platelets than in the control, untreated, and thrombin activated platelets. This quantitative proteomics analysis in turn supported the discovery of novel tetrapeptides DTIs which act as potent inhibitors of thrombin-activated platelets aggregation that can be used in the treatment of acute coronary disease (ACD).

Contact: Cristina Clement, ccc4002@med.cornell.edu

P07.02

Selective Modulation of ER Proteostasis to Inhibit RNA Virus Infection

Lars Plate (1), Katherine M. Almasy (2), Jonathan P. Davies (3), Eli F. McDonald (2)

(1) Departments of Chemistry and Biological Sciences, Vanderbilt University, Nashville, TN, USA

(2) Department of Chemistry, Vanderbilt University, Nashville, TN, USA

(3) Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

RNA viruses continue to pose a grave global health threat as evident from the current COVID-19 pandemic. Besides coronaviruses (CoV), the most common tropical and subtropical transmitted diseases are caused by flaviviruses, such as Dengue, Zika, West Nile, Yellow Fever, and Hepatitis C virus. Development of effective antiviral therapies that target virus proteins is often challenged by high mutation rates, and these agents are usually only selective for specific virus families. This motivates efforts to find alternative therapeutic paradigms. Here, we focus on host-targeted therapeutics by expanding our understanding of common host processes critical for virus infections. Comparative interactomics with tandem mass tag-based quantitative proteomics is a powerful approach to identify and sensitively compare the enrichment of shared and unique host cell dependencies exploited by virus proteins. We will detail our findings from profiling the host cell interactions of several nonstructural proteins from SARS and non-pathogenic CoV strains. We identified pan-strain interactions with mitochondria-associated endoplasmic reticulum (ER) membrane factors and ER proteostasis pathways. As RNA virus replication and assembly frequently occurs at the ER membrane, viruses extensively remodel this organelle and the ER proteostasis network. We exploited these shared host-cell dependencies taking advantage of recently developed small-molecule ER proteostasis regulator compounds. We discovered that such compounds are broadly effective at inhibiting infection with multiple Dengue and Zika virus strains leading to an assembly defect of mature, infectious virions. Using chemoproteomics for target identification, we found that one active compound (147) targets reactive cysteines and perturbs ER redox signaling processes. Ongoing efforts are characterizing which molecular targets are responsible for inhibiting flavivirus infection. Furthermore, we are testing 147 and related ER proteostasis regulators against other RNA viruses, including CoV strains and SARS-CoV-2. These results open up a broadly-applicable therapeutic strategy to target essential host cell processes to impair viral infections.

Contact: Lars Plate, lars.plate@vanderbilt.edu

Computational Proteomics and Data Science

P08.01 HLA Peptide De novo Sequencing with GPU-powered Deep Learning Algorithm
Xin Chen, Rui Qiao, Xiaodong Wei, Shengying Pan, Lei Xin, Baozhen Shen

P08.02 Peptide Correlation Analysis (PeCorA) Reveals Differential Proteoform Regulation
Maria Dermit, Jesse Meyer

P08.03 mokapot: Fast and Flexible Semi-Supervised Learning for Peptide Detection
William Fondrie, William Noble

P08.04 Predicting peptide CCS distributions with a Mixture Density Network
Patrick Garrett, Robin Park, Titus Jung, Casimir Bamberger, John Yates

P08.05 Predicting peptide CCS distributions with a Mixture Density Network
Patrick Garrett, Robin Park, Titus Jung, Casimir Bamberger, John Yates

P08.06 Predicting novel interactions between kinases and substrates using graph representation learning
Sachin Gavali, Karen Ross, Julie Cowart, Chuming Chen, Cathy Wu

P08.07 Modeling nano-bio interactions of 37 engineered nanoparticles that enable deep plasma proteomics studies at unprecedented scale
Daniel Hornburg, Tristan Brown, Shadi Ferdosi, Behzad Tangeysh, Michael Figa, Eltaher Elgierari, Kateryna Riedesel, Jessica Chu, Max Mahoney, Hongwei Xia, Craig Stolarczyk, Theodore Platt, Martin Goldberg, Asim Siddiqui, Juan Cruz Cuevas, John E. Blume, Omid C. Farokhzad

P08.08 Improving power while controlling the false discovery rate when only a subset of peptides are relevant
Andy Lin, Deanna Plubell, Uri Keich, William Noble

P08.09 FragPipe: A Graphical Interface for Fast and Flexible Proteomics Data Analysis
Alexey Nesvizhskii, Dmitry Avtonomov, Fengchao Yu, Guo Ci Teo, Felipe Da Veiga Leprevost, Sarah Haynes, Hui-Yin Chang, Daniel Geiszler, Daniel Polasky

P08.10 Research on The Human Proteome Reaches a Major Milestone: >90% of Predicted Human Proteins Now Credibly Detected, According to the HUPO Human Proteome Project
Gilbert S. Omenn, Lydie Lane, Christopher M. Overall, Ileana M. Cristea, Fernando J. Corrales, Cecilia Lindskog, Young-Ki Paik, Jennifer E. Van Eyk, Siqi Liu, Stephen R. Pennington, Michael P. Snyder, Mark S. Baker, Nuno Bandeira, Ruedi Aebersold, Robert L. Moritz, Eric W. Deutsch

P08.11 Super.Complex v3.0: A Supervised Machine Learning Pipeline for Molecular Complex Detection in Protein-interaction Networks
Meghana Palukuri, Edward Marcotte

P08.12 Combining the Accurate Mass and Time Tag and Ion Mobility to Find Missing Peptides
Robin Park, Sven Brehmer, Titus Jung, Patrick Garrett, Vijayaraja Gnanasambandan, Marc-Antoine Beauvais, Casimir Bamberger, Hyunsoo Kim, Christopher Adams, Dennis Trede, John Yates, Rohan Thakur

P08.13 Combining the Accurate Mass and Time Tag and Ion Mobility to Find Missing Peptides

Robin Park, Sven Brehmer, Titus Jung, Patrick Garrett, Vijayaraja Gnanasambandan, Marc-Antoine Beauvais, Casimir Bamberger, Hyunsoo Kim, Christopher Adams, Dennis Trede, John Yates, Rohan Thakur

P08.14 CCS-centric protein identification

Robin Park, Patrick Garrett, Sven Brehmer, Titus Jung, Vijayaraja Gnanasambandan, Marc-Antoine Beauvais, Sebastian Wehner, Lennard Pfennig, Hyunsoo Kim, Christopher Adams, Dennis Trede, John Yates, Rohan Thakur

P08.15 DeepNovo AB: Next-generation antibody sequencing technology

Rui Qiao, Lei Xin, Xiaodong Wei, Wendy Sun, Hui Li, Xin Chen, Jordanna Newington, Baozhen Shan, Jonathan Krieger, Kyle Hoffman

P08.16 Integration of MSstatsTMT into Proteome Discoverer software 2.5 for statistical analysis of complex TMT experiments

David Horn, Ting Huang, Meena Choi, Olga Vitek, Rosa Viner, Aaron Robitaille, Frank Berg, Kai Fritzemeier, Carmen Paschke

P08.17 A streamlined workflow for the quality control, interpretation and analysis of timsTOF Pro data: towards large-scale clinical proteomics

Xianming Liu, Chengpin Shen, Xiaoxian Du, Xiaoqing Wang, Hebin Liu, Ruoyu Wu

P08.18 CCS-centric protein identification

Robin Park, Tharan Srikumar, Patrick Garrett, Sven Brehmer, Titus Jung, Sebastian Wehner, Lennard Pfennig, Hyunsoo Kim, Christopher Adams, Dennis Trede, John Yates, III

P08.19 MSstats Ecosystem: a Family of R/Bioconductor Packages for Detecting Differential Protein Abundance in Mass Spectrometry Data

Mateusz Staniak, Ting Huang, Devon Kohler, Tsung-Heng Tsai, Ajeya Kempegowda, Dhaval Mohandas, Meena Choi, Olga Vitek

P08.20 Proteomics Data Commons (PDC): A Node in NCI's Cancer Research Data Commons Dedicated to Cancer Proteomics

Ratna Thangudu, Paul Rudnick, Michael Holck, Deepak Singhal, Michael MacCoss, Nathan Edwards, Karen Ketchum, Sudha Venkatachari, John Otridge, Henry Rodriguez, Erika Kim, Anand Basu

P08.21 Improving retention time prediction for modified peptides using multi-step transfer learning

Bo Wen, Kai Li, Wen Jiang, Bing Zhang

P08.22 Features of human disease-associated mutations related to structure, stability, and binding

Jaie Woodard, Chengxin Zhang, Yang Zhang

P08.23 False Discovery Rate-Controlled Match-Between-Runs Enables Accurate and Sensitive Label Free Quantification

Fengchao Yu, Sarah Haynes, Alexey Nesvizhskii

P08.01

HLA Peptide De novo Sequencing with GPU-powered Deep Learning Algorithm

Xin Chen (1), Rui Qiao (1), Xiaodong Wei (1), Shengying Pan (1), Lei Xin (1), Baozhen Shen (1)

(1) Bioinformatics Solutions Inc, Waterloo, Ontario, Canada

Human leukocyte antigen (HLA) peptides are generally short peptides that are brought to the cell surface by major histocompatibility complex (MHC) proteins. These peptides play an important role in the human immune system. The identification of HLA peptides is a key technology that has implicated many applications such as cancer immunotherapy and peptide vaccines. In mass spectrometry, there are three well-established approaches to identifying peptides: database search, spectral library search, and de novo sequencing. As HLA peptides are not enzyme-digested, a database search engine would incur an unfeasibly large search space. Further, existing database engines are not designed for HLA peptides and hence their results are biased in favor of tryptic peptides. Spectral library search, on the other hand, suffers from a lack of comprehensive HLA peptide spectral libraries. In this work, we propose a deep learning based de novo peptide sequencing approach to quickly and accurately identifying HLA peptides as well as a post-analysis protocol on top of MHC binding prediction tools from IEDB. Our experimental results on melanoma patients' HLA-I peptide mass spectrometry (MS) data show that our approach can expand the predicted immunopeptidomes by 5-15%. In the meantime, several interesting mutated tumor specific antigens (also called neoantigens) were detected directly from the MS data, which we believe is not possible with database search or spectral library search.

Contact: Xin Chen, xchen@bioinfor.com

P08.02

Peptide Correlation Analysis (PeCorA) Reveals Differential Proteoform Regulation

Maria Dermit (1), Jesse Meyer (2)

(1) QMUL, UK

(2) − Department of Biochemistry, US

Shotgun proteomics techniques infer the presence and quantity of proteins using peptide proxies produced by cleavage of the proteome with a protease. Most protein quantitation strategies assume that multiple peptides derived from a protein will behave quantitatively similar across treatment groups, but this assumption may be false due to (1) heterogeneous proteoforms and (2) technical artifacts. Here we describe a strategy called peptide correlation analysis (PeCorA) that detects quantitative disagreements between peptides mapped to the same protein. PeCorA fits linear models to assess whether a peptide's change across treatment groups differs from all other peptides assigned to the same protein. PeCorA revealed that ∼15% of proteins in a mouse microglia stress data set contain at least one discordant peptide. Inspection of the discordant peptides shows the utility of PeCorA for the direct and indirect detection of regulated post-translational modifications (PTMs) and also for the discovery of poorly quantified peptides. The exclusion of poorly quantified peptides before protein quantity summarization decreased falsepositives in a benchmark data set. Finally, PeCorA suggests that the inactive isoform of prothrombin, a coagulation cascade protease, is more abundant in plasma from COVID-19 patients relative to non-COVID-19 controls. PeCorA is freely available as an R package that works with arbitrary tables of quantified peptides.

Contact: Maria Dermit, maria.dermit@qmul.ac.uk

P08.03

mokapot: Fast and Flexible Semi-Supervised Learning for Peptide Detection

William Fondrie (1), William Noble (1)

(1) University of Washington Department of Genome Sciences, Seattle, WA, USA

The accurate assignment of peptides to mass spectra is an essential task in a proteomics experiment. This task is most often accomplished using a database search engine, which produces one or more scores that quantify the quality of a peptide-spectrum match (PSM). Machine learning methods have played a pivotal role in boosting the sensitivity of peptide detection by aggregating multiple scores and properties of a PSM---features---into a single score in an adaptable and unbiased manner. One such method, Percolator, introduced a semi-supervised algorithm to iteratively learn a linear support vector machine (SVM) and has proven to be highly effective in practice. Our new tool, mokapot, is a fast and extensible Python implementation of the same semi-supervised algorithm, but expanded to provide immense flexibility for highly customized analyses. Here, we present the benefits of this flexibility in two vignettes. First, we found that the ability of mokapot to use non-linear models and compute grouped confidence estimates improved the detection of modified peptides from an open modification search of RNA-cross-linked proteins: A non-linear gradient boosting model increased the number of modified PSMs (15%), peptides (19%), and proteins (11%) detected at a 1% false discovery rate (FDR) over a linear SVM. We then improved the power and consistency of peptide detection in a single-cell proteomics study by employing a new joint modeling approach. The joint modeling approach resulted in consistent gains in the detected PSMs, peptides, and proteins at 1% FDR over analyzing each experiment independently. Furthermore, joint modeling resulted in a more consistent set of discovered peptides and proteins across experiments in the dataset, yielding less missing data. With its immense flexibility, mokapot brings the power of the Percolator algorithm to a wider array of proteomics studies.

Contact: William Fondrie, wfondrie@uw.edu

P08.04

Predicting peptide CCS distributions with a Mixture Density Network

Patrick Garrett (1), Robin Park (2), Titus Jung (1), Casimir Bamberger (1), John Yates (1)

(1) The Scripps Research Institute, La Jolla, CA, United States

(2) Bruker Daltonics, La Jolla, CA, United States

Database searching and spectral library matching are the leading methods for quantifying and identifying peptides in mass spectrometry based proteomics. These methods rely on comparing measured spectra to theoretical peptides from known proteomes. With the addition of new instruments capable of measuring a peptides CCS (collisional cross section), a new dimension, based on an ions mobility, can be utilized to significantly improve these approaches, but without an accurate method of prediction, the full capability of database searching and spectral library

matching were limited. In this paper we describe a novel machine learning model named PepMDN (peptide mixture density network) , based on applying the most recent advances in natural language processing to this problem. PepMDN utilizes an encoder/decoder based structure which generates a feature vector from a peptide sequence and its charge state, then decodes this feature vector into a predicted CCS distribution. PepMDN 's decoder layer utilizes an MDN, which models the distribution of data through training many overlapping Gaussian distributions, and is capable of predicting the spread of the target values. This is critical as unique peptide sequences can have a multimodal distribution of CCS values, which is not appropriate for traditional machine learning methods. Previous attempts to train models to predict CCS values relied on calculating a median or most common value. While these models demonstrate great capabilities, they lose valuable information regarding the effects that the peptide sequence has on its CCS distribution. We demonstrate that by incorporating the predictions from PepMDN into the database searching application (ProLuCID), we were able to identify more total peptides, and obtain a lower FDR (false discovery rate).

Contact: Patrick Garrett, pgarrett@scripp.edu

P08.05

Predicting peptide CCS distributions with a Mixture Density Network

Patrick Garrett (1), Robin Park (2), Titus Jung (1), Casimir Bamberger (1), John Yates (1)

(1) Yates Lab / TSRI, La Jolla, CA, US

(2) Bruker Daltonics, San Diego, CA, US

Database searching and spectral library matching are the leading methods for quantifying and identifying peptides in mass spectrometry-based proteomics. These methods rely on comparing measured spectra to theoretical peptides from known proteomes. A recent breakthrough of 4D proteomics by the timsTOF Pro adds an extra dimension of separation, providing a solution to accurately clarifying a significant number of spectra with CCS (collisional cross-section) values, which can be utilized to significantly improve these approaches, but without an accurate method of prediction, the full capability of database searching, and spectral library matching are limited. In this paper we describe a novel machine learning model named PepMDN (peptide mixture density network), based on applying the most recent advances in natural language processing to this problem. PepMDN utilizes an encoder/decoder-based structure which generates a feature vector from a peptide sequence and its charge state, then decodes this feature vector into a predicted CCS distribution. PepMDN 's decoder layer utilizes an MDN, which models the distribution of data through training many overlapping Gaussian distributions and is capable of predicting the spread of the target values. This is critical as unique peptide sequences can have a multimodal distribution of CCS values, which is not appropriate for traditional machine learning methods. Previous attempts to train models to predict CCS values relied on calculating a median or most common value. While these models demonstrate great capabilities, they lose valuable information regarding the effects that the peptide sequence has on its CCS distribution. We demonstrate that by incorporating the predictions from PepMDN into the database searching application (ProLuCID), we were able to identify more total peptides, and obtain a lower FDR (false discovery rate).

Contact: Patrick Garrett, pgarrett@scripps.edu

P08.06

Predicting novel interactions between kinases and substrates using graph representation learning

Sachin Gavali (1), Karen Ross (2), Julie Cowart (1), Chuming Chen (1), Cathy Wu (1)

(1) Data Science Institute, University of Delaware, Newark, Delaware, USA

(2) Georgetown University Medical Center, Washington, District of Columbia, USA

The proteins in the human phosphoproteome constitute a vast network of interacting kinases and substrates. Among these proteins, many of the known phosphorylated substrates have limited kinase information. Identifying new kinase-substrate interactions requires an experimental approach which is time consuming and expensive. To alleviate this, we use graph representation learning approaches to build predictive models. For this study, we use the data in iPTMnet to build an interaction network which consists of 3419 proteins and 8591 interactions. We then reserve 10% of the interactions for validation, 10% for testing and use the remaining 80% for training. In the first approach, we utilized the node2vec algorithm to learn a vector representation of the network. We then used these representations to train a random forest model which achieved an F1 score of 0.79 and ROC-AUC of 84%. Since the node2vec algorithm takes into account only the connectivity of proteins, it cannot make any predictions for proteins with no connections. Hence, in the second approach we combine the network information with the subcellular locations of the proteins and use GraphSage algorithm to learn a vector representation of the resulting network. This model achieved an F1 score of 0.70 and ROC-AUC of 83%. To further evaluate the biological feasibility of the models for studying understudied proteins we utilized them to identify kinases for 162 substrates from “Illuminating the Druggable Genome” project. The node2vec and GraphSage models were able to identify novel kinases for 12 and 92 substrates respectively. As expected, the GraphSage model in spite of being slightly less performant, was able to identify more interactions due to contextual information about the proteins. This work has demonstrated that it is possible to exploit the connectivity of proteins while including contextual information to predict novel kinase-substrate relationships using machine learning methods.

Contact: Sachin Gavali, saching@udel.edu

P08.07

Modeling nano-bio interactions of 37 engineered nanoparticles that enable deep plasma proteomics studies at unprecedented scale

Daniel Hornburg (1), Tristan Brown (1), Shadi Ferdosi (1), Behzad Tangeysh (1), Michael Figa (1), Eltaher Elgierari (1), Kateryna Riedesel (1), Jessica Chu (1), Max Mahoney (1), Hongwei Xia (1), Craig Stolarczyk (1), Theodore Platt (1), Martin Goldberg (1), Asim Siddiqui (1), Juan Cruz Cuevas (1), John E. Blume (1), Omid C. Farokhzad (1)

(1) Seer Inc., Foster City, CA, USA

Introducing a nanoparticle (NP) into a biofluid such as blood plasma leads to the formation of a selective, specific, and reproducible protein corona at the nano-bio interface driven by the relationship between protein-NP affinity, protein abundance and protein-protein interactions. We previously demonstrated that this process, incorporated within the Proteograph™ proteomics platform, offers superior profiling performance in terms of depth, breadth, precision, and throughput compared to conventional workflows. NP physicochemical design possibilities are essentially infinite and can be tailored to enhance and differentiate protein selectivity. Here we investigate the relationship between nanoparticle chemical functionalization and corona formation building linear mixed-effects models that could enhance future NP design. We investigated a set of 37 engineered nanoparticles with specific physicochemical properties processed with Proteograph interrogating a pooled plasma sample. Proteomics data were acquired using 30-minute LC runs with an Orbitrap Lumos. MaxQuant raw data processing identified more than 1,500 protein groups at 1% protein and peptide FDR. Developing machine learning (linear mixed-effects) models we identified significant relationships between physicochemical NP properties (including zeta potential, amine, and carboxy functionalization) and differential abundance of individual proteins and protein classes within NP corona. For example, 23% of the abundance of C-reactive protein (CRP) in a protein corona was associated with NP zeta potential, and 22% could be allocated to polymeric and sugar surface functionalization. In contrast, we observed the abundance of plasma kallikrein (KLKB1) to be unaffected

by NP zeta potential but more than 50% driven by sugar functionalization. Our results suggest that we can model the relationship between NP surface functionalization and specific proteins or protein classes in complex biological samples and use this information to guide future NP design to further increase the utility of the Proteograph platform in proteomics research and biomarker discovery.

Contact: Daniel Hornburg, dhornburg@seer.bio

P08.08

Improving power while controlling the false discovery rate when only a subset of peptides are relevant

Andy Lin (1), Deanna Plubell (1), Uri Keich (2), William Noble (1)

(1) Department of Genome Sciences, University of Washington, Seattle, WA, USA

(2) School of Mathematics and Statistics, University of Sydney, Sydney, NSW, Australia

The standard proteomics database search strategy involves searching spectra against a peptide database and estimating the false discovery rate (FDR) of the resulting set of peptide-spectrum matches. Generally, experimental spectra are searched against a database containing all peptides reasonably expected to be found in the sample. However, in certain scenarios, not all the peptides present in the database are relevant to the hypothesis being asked. For example, when proteomics is used to study the biology of Plasmodium falciparum, the causative agent of malaria, spectra are typically searched against a database containing both the human and P. falciparum proteome even though human peptides are irrelevant to Plasmodium biology. This distinction between relevant and irrelevant database peptides is currently unaccounted for in the standard FDR estimation process. Recently, two methods were proposed to address this problem: subset-search and all-sub. We compared these two methods to a previously developed method called group-FDR. We show that both new methods fail to control the FDR. For subset-search, this failure is due to the presence of “neighbor" peptides, which are defined as irrelevant peptides with a similar precursor mass and fragmentation spectrum as a relevant peptide. Not considering neighbors compromises the FDR estimate because a spectrum generated by an irrelevant peptide can incorrectly match well to a relevant peptide. Therefore, we have developed a new method, "filter then subset-neighbor search" (FSNS), that accounts for neighbor peptides. We show evidence that FSNS properly controls the FDR when neighbors are present and that FSNS outperforms group-FDR, the only other method able to control the FDR relative to a subset of relevant

Contact: Andy Lin, linandy@uw.edu

P08.09

FragPipe: A Graphical Interface for Fast and Flexible Proteomics Data Analysis

Alexey Nesvizhskii (1), Dmitry Avtonomov (2), Fengchao Yu (2), Guo Ci Teo (2), Felipe Da Veiga Leprevost (2), Sarah Haynes (2), Hui-Yin Chang (3), Daniel Geiszler (2), Daniel Polasky (2)

(1) Department of Pathology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, United States

(2) Department of Pathology, University of Michigan, Ann Arbor, MI, United States

(3) Department of Biomedical Science and Engineering, National Central University; Department of Pathology, University of Michigan, Ann Arbor, MI, United States

FragPipe is a multi-platform proteomics analysis software that accommodates a variety of common workflows through a graphical interface. Users can choose from complete pre-configured workflows (including routine database search with label free or TMT/iTRAQ quantification; O- or N-glycopeptide, HLA/peptidome, and open searches; and differential phosphorylation site analysis) which are ready to run, just drag and drop LC-MS data files to the FragPipe window and specify the output location. FragPipe brings together a number of software packages that users can run alone or consecutively. Fast database searching can be performed with MSFragger. DIA-Umpire enables pre-processing of data-independent acquisition spectra. Crystal-C can be used for post-search cleanup of open search results. Peptide validation and protein inference are performed with PeptideProphet and ProteinProphet, respectively, both distributed via the Philosopher toolkit, which also performs isobaric quantification and generates multi-level reports. PTM-Shepherd characterizes post-translational modifications (PTMs) from open searches, and PTM-Prophet can be used for localization of PTMs from closed searches. Label-free quantification can be accomplished with IonQuant, and TMT or iTRAQ data can be analyzed with TMT-Integrator. FragPipe supports direct reading of native file formats from Thermo (.raw, including FAIMS) and Bruker timsTOF (.d), removing the need for file conversion prior to analysis. Input files can be organized into fractions, replicates, and experimental groups for better reporting and compatibility with MSstats for differential label-free quantification analysis. For isobaric labeling-based workflows, TMT-Integrator can be used to generate accurate, normalized quantification reports at multiple levels: peptide-spectrum match, modification site, peptide, protein, and gene. FragPipe enables users to customize their workflows with minimal installation and an easy-to-use interface. FragPipe is well suited to tasks such as high throughput peptide and protein identification, post-translational modification discovery through open-search, PTM profiling and localization, label-free or isobaric labelling-based quantification, peptidome workflows, and spectral library generation.

Contact: Alexey Nesvizhskii, nesvi@umich.edu

P08.10

Research on The Human Proteome Reaches a Major Milestone: >90% of Predicted Human Proteins Now Credibly Detected, According to the HUPO Human Proteome Project

Gilbert S. Omenn (1), Lydie Lane (2), Christopher M. Overall (3), Ileana M. Cristea (4), Fernando J. Corrales (5), Cecilia Lindskog (6), Young-Ki Paik (7), Jennifer E. Van Eyk (8), Siqi Liu (9), Stephen R. Pennington (10), Michael P. Snyder (11), Mark S. Baker (12), Nuno Bandeira (13), Ruedi Aebersold (14), Robert L. Moritz (15), Eric W. Deutsch (15)

(1) University of Michigan, Ann Arbor, Michigan, United States

(2) CALIPHO Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

(3) University of British Columbia, Vancouver, Canada

(4) Princeton University, Princeton, New Jersey, United States

(5) Centro Nactional de Biotecnologia, Madrid, Spain

(6) Uppsala University, Uppsala, Sweden

(7) Yonsei Proteome Research Center, Seoul, Korea

(8) Cedars Sinai, Los Angeles, California, United States

(9) BGI Group, Shenzhen, China

(10) University College Dublin, Dublin, Ireland

(11) Stanford University, Stanford, California, United States

(12) Macquarie University, Macquarie Park, Australia

(13) University of California, San Diego, La Jolla, California, United States

(14) ETH-Zurich and University of Zurich, Zurich, Switzerland

(15) Institute for Systems Biology, Seattle, Washington, United States

According to the Decadal High-Stringency Blueprint of the Human Proteome1 and the 2020 Metrics of the HUPO Human Proteome Project (HPP)2, expression has now been detected at the protein level for >90% of the 19,773 predicted proteins coded in the human genome. The HPP annually reports progress made throughout the world toward credibly identifying and characterizing the complete human protein parts list and promoting proteomics as an integral part of multi-omics studies in medicine and the life sciences. NeXtProt release 2020-01 classified 17,874 proteins as PE1, having strong protein-level evidence, up 180 from 17,694 one year earlier, under HPP Guidelines v3.03. These represent 90.4% of the 19,773 predicted coding genes (all PE1,2,3,4 proteins). Conversely, the number of neXtProt PE2,3,4 proteins, the “missing proteins” (MPs), was reduced by 230 from 2129 to 1899 since the neXtProt 2019-01 release. PeptideAtlas is the primary source of uniform re-analysis of raw mass spectrometry data for neXtProt, supplemented this year with extensive data from MassIVE. PeptideAtlas 2020-01 added 362 canonical proteins between 2019 and 2020 and MassIVE contributed 84 more; many converted PE1 entries based on non-MS evidence to the MS-based subgroup. Further progress on MPs is expected. The 19 Biology and Disease-driven B/D-HPP teams continue to identify driver proteins that underlie disease states, characterize regulatory mechanisms controlling the functions of these proteins, their proteoforms, and their interactions, and define transitions from correlation to co-expression to causal networks after system perturbations. The Human Protein Atlas/Antibody Pillar published Blood, Brain, and Metabolic Atlases. The Mass Spectrometry pillar conducted a 22-laboratory analysis of 96 phosphopeptides4.
1Adhikari et al, Nat Commun, 2020 Oct 16;11:5301. PMID:33067450. 2Omenn et al, J Proteome Res, 2020 Oct 19. PMID:329331287.
3Deutsch et al, J Proteome Res 2019 Dec; PMID:31599596. 4Hoopman et al, J Proteome Res, 2020 Dec (in press).

Contact: Gilbert Omenn, gomenn@umich.edu

P08.11

Super.Complex v3.0: A Supervised Machine Learning Pipeline for Molecular Complex Detection in Protein-interaction Networks

Meghana Palukuri (1), Edward Marcotte (2)

(1) Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas, USA

(2) Institute for Cell and Molecular Biology, University of Texas at Austin, Austin, Texas, USA

Protein complexes can be computationally mined from protein-interaction networks with community detection methods, accelerating biological experiments by providing candidates for previously unknown protein complexes. Previous methods in community detection comprise a majority of unsupervised clustering strategies and semi-supervised machine learning methods, which work on the assumption that communities are dense subgraphs in a network - which is not always true, especially in the case of protein complexes which exhibit diverse topologies on the protein-interaction network. Further, the few existing supervised machine learning methods can be improved - (i) in terms of accuracy, by using better-suited machine learning models, and (ii) in terms of efficiency, by using parallel algorithms, which would improve upon previous serial algorithms that had been applied to small networks. In this work, we present Super.Complex, a distributed supervised machine learning pipeline for community detection in networks, which learns a community fitness function from known communities using an auto-ML method and uses the fitness function to detect

new communities in the network. The heuristic local search algorithm to find maximally scoring communities uses an epsilon-greedy criterion with additional criteria such as iterative-simulated annealing or pseudo-metropolis. This is designed as an embarrassingly parallel process and we provide an implementation for a compute cluster that scales up to large networks that can fit on disk memory of a single compute node. Comparing a set of predicted complexes and known complexes satisfactorily is an outstanding issue, for which we propose 3 new evaluation measures including a 'net F1 score'. Application of Super.Complex on hu.MAP, a human protein-interaction network with ~7k nodes and ~50k edges yields 1361 protein complexes with a net F1 score of 0.7, perfectly recalling 61 known literature-curated complexes from CORUM. Of the predicted complexes, 234 are potentially linked to COVID-19 via direct interaction with a SARS-COV2 protein.

Contact: Meghana Palukuri, meghana@oden.utexas.edu

P08.12

Combining the Accurate Mass and Time Tag and Ion Mobility to Find Missing Peptides

Robin Park (1), Sven Brehmer (1), Titus Jung (2), Patrick Garrett (2), Vijayaraja Gnanasambandan (1), Marc-Antoine Beauvais (1), Casimir Bamberger (2), Hyunsoo Kim (2), Christopher Adams (1), Dennis Trede (1), John Yates (2), Rohan Thakur (1)

(1) Bruker Daltonics, San Diego, CA, United States

(2) Scripps Research, La Jolla, CA, United States

An accurate mass and time (AMT) tag approach for proteomic analyses has been developed to validate peptide identifications using mass measurement accuracy and retention time. Because of the complexity of the proteome, ambiguities in peptide assignments can occur when two peptides are assigned to the same mass value or when PTM peptides are identified with low localization scores. In this study, we applied high throughput 4-D proteomics strategy to the AMT tag by adding the ion mobility dimension (CCS-centric AMT tag), which plays a critical role in clarifying such ambiguities. The timsTOF Pro and timsTOF fleX instruments provide four-dimensional data space called 4D-Proteomics with CCS values and parallel accumulation serial fragmentation (PASEF) to improve ion utilization efficiency and data acquisition speed. In this study, we utilized CCS-centric AMT for finding missing peptides when we compared multiple experiments. We acquired the spectra on a timsTOF PRO from 500ng of digested Hela samples. The search space included all tryptic peptide candidates that fell within the precursor mass tolerance window with three miscleavage constraints. Parallel Database Search Engine in Real-time (PaSER) performed the database search while the instrument was acquiring spectra on the UniProt human database. Census reconstructed chromatograms for each identified peptide. We built a CCS-centric AMT library storing accurate precursor mass, retention time, charge states, and CCS values. When peptides are not identified in all the relevant replicates, Census went through spectra, searching them using information from the library to detect missing peptides. To increase accuracy for finding peptide precursors, we compared CCS values from library and target spectra. For peptide abundance, we applied smoothing and calculated Pearson product-moment correlation coefficient comparing theoretical and experimental isotope distributions.

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

P08.13

Combining the Accurate Mass and Time Tag and Ion Mobility to Find Missing Peptides

Robin Park (1), Sven Brehmer (1), Titus Jung (2), Patrick Garrett (2), Vijayaraja Gnanasambandan (1), Marc-Antoine Beauvais (1), Casimir Bamberger (2), Hyunsoo Kim (2), Christopher Adams (1), Dennis Trede (1), John Yates (2), Rohan Thakur (1)

(1) Bruker Daltonics, San Diego, CA, United States

(2) Scripps Research, La Jolla, CA, United States

An accurate mass and time (AMT) tag approach for proteomic analyses has been developed to validate peptide identifications using mass measurement accuracy and retention time. Because of the complexity of the proteome, ambiguities in peptide assignments can occur when two peptides are assigned to the same mass value or when PTM peptides are identified with low localization scores. In this study, we applied high throughput 4-D proteomics strategy to the AMT tag by adding the ion mobility dimension (CCS-centric AMT tag), which plays a critical role in clarifying such ambiguities. The timsTOF Pro and timsTOF fleX instruments provide four-dimensional data space called 4D-Proteomics with CCS values and parallel accumulation serial fragmentation (PASEF) to improve ion utilization efficiency and data acquisition speed. In this study, we utilized CCS-centric AMT for finding missing peptides when we compared multiple experiments. We acquired the spectra on a timsTOF PRO from 500ng of digested Hela samples. The search space included all tryptic peptide candidates that fell within the precursor mass tolerance window with three miscleavage constraints. Parallel Database Search Engine in Real-time (PaSER) performed the database search while the instrument was acquiring spectra on the UniProt human database. Census reconstructed chromatograms for each identified peptide. We built a CCS-centric AMT library storing accurate precursor mass, retention time, charge states, and CCS values. When peptides are not identified in all the relevant replicates, Census went through spectra, searching them using information from the library to detect missing peptides. To increase accuracy for finding peptide precursors, we compared CCS values from library and target spectra. For peptide abundance, we applied smoothing and calculated Pearson product-moment correlation coefficient comparing theoretical and experimental isotope distributions.

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

P08.14

CCS-centric protein identification

Robin Park (1), Patrick Garrett (2), Sven Brehmer (1), Titus Jung (2), Vijayaraja Gnanasambandan (1), Marc-Antoine Beauvais (1), Sebastian Wehner (1), Lennard Pfennig (1), Hyunsoo Kim (2), Christopher Adams (1), Dennis Trede (1), John Yates (2), Rohan Thakur (1)

(1) Bruker Daltonics, San Diego, CA, United States

(2) Scripps Research, San Diego, CA, United States

The database search algorithms to identify proteins have advanced since SEQUEST was first introduced in 1994. The search engines calculate scores by comparing experimental spectra from MS/MS data against in-silico spectra derived from protein databases. Even though a lot of efforts have been invested in improving protein identification, still a large number of spectra cannot be identified. One of the obstacles is spectrum ambiguity. For example, more than one peptide candidate from the protein database can often have the same precursor mass and similar fragment ion patterns generating competing PSM search scores.

A recent breakthrough of 4D proteomics by the timsTOF Pro adds an extra dimension of separation, providing a solution to accurately clarifying a significant number of spectra with CCS (collisional cross-section) values. To address the spectrum ambiguity, we have developed a deep learning model to predict retention time, fragment ion intensities using bidirectional long-short term memory (LSTM) recurrent neural networks, and CCS values using PepMDN (peptide mixture density network) model. IP2GPU search engine calculates an ion mobility-based search score called IMScores, p-values derived the CCS distribution based on peptide candidates. Next, we added IMScores as an additional metric for DTASelect to

classify decoy and target peptides and calculate FDR using quadratic discriminant analysis. The unique CCS prediction model and IMScore we developed can be used to improve both protein identification and quantification.

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

P08.15

DeepNovo AB: Next-generation antibody sequencing technology

Rui Qiao (1), Lei Xin (1), Xiaodong Wei (1), Wendy Sun (1), Hui Li (1), Xin Chen (1), Jordanna Newington (1), Baozhen Shan (1), Jonathan Krieger (1), Kyle Hoffman (1)

(1) Bioinformatics Solutions Incorporation, Waterloo, ON, Canada

Antibody protein sequencing by LC-MS showed potential application in decoding immune complexity. Deep learning-based algorithms have shown superior performance in peptide de novo sequencing on both accuracy and speed. In this work we propose DeepNovo AB, an antibody protein sequencing technology that optimized LC-MS experimental setting and de novo peptide sequencing engine. In DeepNovo AB, separate de novo models are trained for each LC-MS setting, including the combination of enzyme and fragmentation method. The predicted de novo peptides with MS/MS data are fed into a de Bruijn graph-based assembler to reconstruct the original antibody sequences. Our experiment results show that the fine-tuned experimental specific de novo models significantly outperform general peptide de novo sequencing algorithms on antibody datasets by a large margin. Including multi-enzyme sample digestion, LC-MS/MS analysis, and automatic sequencing, an antibody sequence can be derived in under 48 hours. The data analysis step takes only about 60 - 90 minutes.

Contact: Rui Qiao, rqiao@bioinfor.com

P08.16

Integration of MSstatsTMT into Proteome Discoverer software 2.5 for statistical analysis of complex TMT experiments

David Horn (1), Ting Huang (2), Meena Choi (2), Olga Vitek (2), Rosa Viner (1), Aaron Robitaille (1), Frank Berg (1), Kai Fritzemeier (1), Carmen Paschke (1)

(1) Thermo Fisher Scientific, San Jose, CA, USA

(2) Northeastern University, Boston, MA, USA

MSstats is an R package for statistical relative quantification of proteins and peptides in mass spectrometry-based proteomics, which has become a standard tool for statistical analysis of label-free quantification data. Recently, the MSstatsTMT library was released to support quantitation of multiplexing experiments using Tandem Mass Tag™ (TMT™) isobaric reagents. These R libraries support direct import of peptide-spectrum match

(PSM) tables from Thermo Scientific™ Proteome Discoverer™ software, but they require a manual step to export the list of peptide-spectrum matches. Furthermore, the end user must create an annotation file that maps quantitation channels from the TMT reporter ions with the various study factors, which is a time consuming and error prone process. Here, we demonstrate how the Proteome Discoverer software 2.5 Scripting Node can automatically call an R script that sends study factor annotations and PSM tables to the MSstatsTMT library. The Proteome Discoverer software subsequently reimports the results for visualization and bioinformatic analysis. We demonstrate this scripting workflow with Proteome Discoverer software on three different TMT datasets of increasing complexity; First, we evaluated a single plex experiment using the Thermo Scientific™ Pierce™ TMT11plex Yeast Digest standard. Secondly, two replicate plexes of the same yeast triple knockout standard were analyzed. Thirdly, we analyzed three replicate TMT11plex experiments showing the effect of SARS-CoV-2 infection in human cells. Additionally, we show integration of the MSstatsTMT statistical results combine with the enrichment analysis function in Proteome Discoverer software to aid the downstream bioinformatic analysis of each of these datasets. In summary, we show that MSstatsTMT library can be access through a scripting node in Proteome Discoverer software to improve the data analysis and quantitation of TMT multiplexing experiments.

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

P08.17

A streamlined workflow for the quality control, interpretation and analysis of timsTOF Pro data: towards large-scale clinical proteomics

Xianming Liu (1), Chengpin Shen (2), Xiaoxian Du (1), Xiaoqing Wang (2), Hebin Liu (2), Ruoyu Wu (1)

(1) Bruker Daltonics, Shanghai, Shanghai, China

(2) Shanghai Omicsolution Co., Ltd, Shanghai, Shanghai, China

The importance of quality control (QC) has been long acknowledged, especially in the large-scale proteomics researches, but most of them only utilized protein quantitation sheet produced by search engine. Such kind of QC lacks comprehensive LC and MS level information that may influence dataset quality in all aspects. Here, we introduce a web based proteomics data interpretation platform, OmicsCloud, to provide a streamlined MS data processing solution. We use SpectroMine, a DDA/DIA/PRM search platform from Biognosys Inc. as QC module to evaluate LC, MS/MS data quality. All the factors that affect the data quality like missed cleavage, contaminants, chromatographic tailing, non-specific cleavage, spray instability, cross species interference are evaluated. Furthermore, based on the Pulsar engine in Spectronaut/Spectromine, large-scale DIA/DDA data quantitation together with missing value imputation, batch effect correction, project-wide normalization and etc. are provided. Several species and organ specific libraries were generated by Bruker timsTOF Pro platform, which are set as reference libraries to evaluate user's proteomics data coverage depth. Based on OmicBOX software, we established an automatic proteome sequence updating and annotation pipeline to provide an up-to-date bioinformatics analysis, including gene ontology, KEGG, interpro, COG/KOG, STRING and etc. In conclusion, a standardized MS data QC, large-scale DIA quantitation and bioinformatics platform called OmicsCloud can tremendously facilitates clinical proteomics research.

Contact: Chengpin Shen, issac.shen@omicsolution.com

P08.18

CCS-centric protein identification

Robin Park (1), Tharan Srikumar (1), Patrick Garrett (2), Sven Brehmer (3), Titus Jung (2), Sebastian Wehner (3), Lennard Pfennig (3), Hyunsoo Kim (2), Christopher Adams (1), Dennis Trede (3), John Yates, III (2)

(1) Bruker Scientific LLC, La Jolla, CA, USA

(2) The Scripps Research Institute, La Jolla, CA, USA

(3) Bruker Gmbh, Bremen, Germany

Database search algorithms to identify proteins have advanced since SEQUEST was first introduced in 1994. Search engines calculate scores by comparing experimental spectra from MS/MS data against in-silico spectra derived from protein databases. Even though, a lot of effort has been invested in improving protein identification, a large number of spectra still cannot be identified. One of the obstacles is spectrum ambiguity. For example, more than one peptide candidate from the protein database can often have the same precursor mass and similar fragment ion patterns generating competing PSM search scores.

The recent breakthrough of 4D-Proteomics by the timsTOF Pro adds an extra dimension of information, providing a solution to accurately clarify a significant number of spectra with CCS (collisional cross-section) values. To address the spectrum ambiguity, we have developed a deep learning model to predict retention time, fragment ion intensities, using bidirectional long-short term memory (LSTM) recurrent neural networks, and CCS values, using PepMDN (peptide mixture density network) models. IP2GPU search engine calculates an ion mobility-based search score called IMScores, and derives p-values from the CCS distribution based of peptide candidates. Next, we added IMScores as an additional metric for DTASelect to classify decoy and target peptides and calculate FDR using quadratic discriminant analysis. The unique CCS prediction model and IMScore we developed can be used to improve both protein identification and quantification.

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

P08.19

MSstats Ecosystem: a Family of R/Bioconductor Packages for Detecting Differential Protein Abundance in Mass Spectrometry Data

Mateusz Staniak (1), Ting Huang (2), Devon Kohler (2), Tsung-Heng Tsai (3), Ajeya Kempegowda (2), Dhaval Mohandas (2), Meena Choi (2), Olga Vitek (2)

(1) Institute of Mathematics, University of Wrocław, Wrocław, Poland

(2) Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA

(3) Kent State University, Kent, Kent, OH, USA

Quantitative mass spectrometry (MS)-based proteomics is an important technology for biological and clinical research. A typical application is the detection of differentially abundant proteins across conditions. To this end, modern proteomic workflows employ a variety of acquisition modes and labeling strategies, resulting in large and diverse datasets that are then processed with a wide range of computational tools.

The MSstats Ecosystem, including MSstatsConvert 1.0, MSstats 4.0, MSstatsTMT 2.0, MSstatsPTM 1.0 aims to support statistical analyses of data generated by these diverse workflows. Implemented open source in R/Bioconductor, these packages take as input identified and quantified spectral peaks and perform data preprocessing and model-based differential analysis. They offer tools for visualization, quality control and sample size estimation.

We will present the latest developments in the MSstats ecosystem. A recent update to MSstats and MSstatsTMT, along with a newly released MSstatsConvert package introduces a new, modular software design. This design increases interoperability of MSstats with MS data processing tools, and ensures maximum code reusability. It increases the efficiency of data transformations, reducing the required computation time by up to 50% depending on a task. MSstats 4.0 offers general tools for transforming MS datasets into a unified format suitable for downstream statistical analyses. It allows the user to perform data transformations such as filtering or normalization with methods implemented in MSstats, as well as with methods implemented outside of the MSstats ecosystem. The format can be easily wrapped to create GUI or to parallelize code execution. The MSstats ecosystem is routinely tested on over 30 datasets with diverse properties, and supports public submission of analysis scripts and results to repository MassIVE.quant. We will present the design principles of MSstats 4.0. and demonstrate its applications and performance with case studies of published proteomics datasets.

Contact: Mateusz Staniak, staniak@math.uni.wroc.pl

P08.20

Proteomics Data Commons (PDC): A Node in NCI's Cancer Research Data Commons Dedicated to Cancer Proteomics

Ratna Thangudu (1), Paul Rudnick (2), Michael Holck (1), Deepak Singhal (1), Michael MacCoss (3), Nathan Edwards (4), Karen Ketchum (1), Sudha Venkatachari (1), John Otridge (5), Henry Rodriguez (6), Erika Kim (6), Anand Basu (1)

(1) ESAC, Inc., Rockville, MD, USA

(2) Spectragen Informatics LLC, Bainbridge Island, WA, USA

(3) University of Washington, Seattle, WA, USA

(4) Georgetown University, Washington, DC, USA

(5) Frederick Natl. Laboratory, Rockville, MD, USA

(6) Natl. Cancer Institute, Bethesda, MD, USA

The objective of the National Cancer Institute's Proteomic Data Commons (PDC) is to make cancer-related proteomic datasets easily accessible to the public. As a domain-specific repository within the Cancer Research Data Commons (CRDC), the vision for the PDC is to provide researchers with the ability to find and analyze mass spectrometry-based proteomic data across a wide variety of tumor types. The CRDC framework facilitates multi-omic integration in support of precision medicine through interoperability with other resources. The PDC provides an intuitive interface for researchers to search, visualize, and analyze protein expression (through their mapped genes) across diverse studies. Other features include building and exploring pan-cancer cohorts using highly curated, clinical metadata, and comprehensively viewing a study without needing to download the data. Through a robust and extensible data model, rich data dictionaries, and an application programming interface (API), PDC facilitates interoperability with NCI Cloud Resources (e.g., Seven Bridges Cancer Genomic Cloud) for efficient access to cloud computation. All of the data in PDC are harmonized through a common data analysis pipeline, removing the data analysis variable, and enabling comparisons across datasets. When available, PDC simplifies access to sample-matched data available at other resources, such as the Genomic Data Commons and the Cancer Imaging Archive. PDC will also leverage other emerging CRDC efforts, such as harmonized data models, to support cross-domain analysis of large datasets. In this presentation, we will highlight the various use cases and the complete feature set. The resource is available at https://pdc.cancer.gov. Questions and comments can be sent to nci.pdc.help@esacinc.com

Contact: Ratna Thangudu, ratna.thangudu@esacinc.com

P08.21

Improving retention time prediction for modified peptides using multi-step transfer learning

Bo Wen (1), Kai Li (1), Wen Jiang (1), Bing Zhang (1)

(1) Lester and Sue Smith Breast Center, Houston, TX, USA

In bottom-up proteomics, peptide mixtures are separated via a liquid chromatography (LC) system before they are injected into the coupled mass spectrometer. The retention time (RT) of a peptide recorded by the system is an intrinsic feature of the peptide and has many applications. RTs of non-modified peptides can be accurately predicted using deep learning algorithms such as AutoRT. However, RT prediction for modified peptides remains challenging due to limited training sample size and additional complexity in encoding modified peptides.

In order to improve the accuracy of RT prediction for modified peptides, we proposed a multi-step transfer learning strategy based on the AutoRT framework (https://github.com/bzhanglab/AutoRT). Firstly, we trained a general base model using a large public dataset derived from a global proteome experiment. Next, for each type of PTM, we used a large public dataset of the specific PTM type to fine-tune the base model to create a PTM-specific base model. Finally, we fine-tuned the PTM-specific base model to produce an experiment-specific model using a PTM dataset of interest, which could predict RTs of peptides under the same LC condition. We evaluated the performance of the new method by comparing it with a few different training strategies and previous methods on a few PTM datasets. Our preliminary data showed that the new method outperformed all other methods compared in this study, and the improvement was the most pronounced when applied to small datasets.

In summary, a multi-step transfer learning strategy was developed to predict RTs for modified peptides. The performance of RT prediction for modified peptides was comparable to that for non-modified peptides. Works are underway to demonstrate the utility of the accurately predicted RTs of modified peptides in the analysis of data from PTM profiling experiments and the design of parallel reaction monitoring-mass spectrometry experiments.

Contact: Bo Wen, bo.wen@bcm.edu

P08.22

Features of human disease-associated mutations related to structure, stability, and binding

Jaie Woodard (1), Chengxin Zhang (1), Yang Zhang (1)

(1) University of Michigan, Ann Arbor, Michigan, USA

Numerous human diseases result from missense genetic mutations. At the same time, many such mutations are known to be benign and part of normal variation. We present results based on mapping over 20K known mutations from the Humsavar database to x-ray and NMR structures. For each mutation, we compute stability change predictions using the program EvoEF, which estimates the change in folding free energy upon mutation, based on thermodynamic data from a large dataset. Decision tree analysis suggests that mutations with delta-delta-G greater than about 2 kcal/mol tend to be pathogenic; such a low value suggests that kinetic, rather than purely thermodynamic, stability may be under selection in the cellular environment. We organize our results into a publicly-available database called ADDRESS (Annotated Database of Disease-RElated Sequences and

Structures). We are determining additional features, including ligand-binding prediction using the program COACH, with the goal of developing a high-quality pathogenicity prediction tool incorporating protein structural data.

Contact: Jaie Woodard, jcwood@umich.edu

P08.23

False Discovery Rate-Controlled Match-Between-Runs Enables Accurate and Sensitive Label Free Quantification

Fengchao Yu (1), Sarah Haynes (1), Alexey Nesvizhskii (1)

(1) Department of Pathology, University of Michigan, Ann Arbor, Michigan, United States

Label free quantification is commonly used to measure peptides and proteins in bottom-up proteomics. Due to the nature of data-dependent acquisition, missing quantification values are unavoidable. These missing values weaken the power of label-free proteomics studies to uncover true quantitative differences between biological samples or experimental conditions. Match-between-runs (MBR) has become a popular approach to mitigate the missing value problem, where peptides identified by tandem mass spectra in one run are transferred to another by inference based on m/z, charge state, retention time, and ion mobility when applicable. Though tolerances are used to ensure such transferred identifications are reasonably located and meet certain quality thresholds, little work has been done to evaluate the statistical confidence of MBR. Here, we present a mixture model-based approach to estimate the false discovery rate (FDR) of peptide and protein identification transfer, which we implement in the label-free quantification tool IonQuant. We use a two-organism dataset from an Orbitrap Fusion Lumos mass spectrometer to demonstrate that FDR-controlled MBR has a lower false positive rate compared to MaxQuant. Using four HeLa cell lysate replicates from a timsTOF Pro mass spectrometer, we show that our approach has high accuracy (lower coefficient of variation) and sensitivity (more quantified proteins). Finally, we demonstrate that our tool works well with single-cell data and FAIMS-LC-MS/MS (high-field asymmetric waveform ion mobility spectrometry coupled with liquid chromatography tandem mass spectrometry) data.

Contact: Fengchao Yu, yufe@umich.edu

Data Analysis and Visualization

P09.01

Impact of spectral and chromatogram libraries on data-independent acquisition workflows: Human case studies for neurodegenerative disease and aging

Joanna Bons (1), Sandip Kumar Patel (1), Kizito-Tshitoko Tshilenge (1), Lindsay Pino (1), Nathan Basisty (1), Cameron Wehrfritz (1), Judith Campisi (2), Lisa M. Ellerby (1), Birgit Schilling (1)

(1) The Buck Institute for Research on Aging, Novato, CA, United States

(2) The Buck Institute for Research on Aging & Lawrence Berkeley Laboratory, University of California, Novato & Berkeley, CA, United States

Data-independent acquisition (DIA) is gaining more and more interest for label-free proteome-wide quantification as it provides high accuracy and reproducibility. Spectral libraries are generally used for data extraction and interpretation, and can be built with or without offline sample fractionation in data-dependent acquisition (DDA) mode. Additionally, more recent promising alternatives have been developed to cope with limitations of these experiment-derived spectral libraries, such as the use of predicted spectral libraries, hybrid DDA-DIA libraries, and chromatogram libraries. As these impact protein identification and quantification, the content and quality of spectral libraries are crucial.

In this study, we benchmarked the qualitative and quantitative performances of library-based DIA workflows, using sample-specific libraries (generated using FAIMS or high pH reversed-phase fractionation), a deep pan-Human library, in silico-generated libraries, hybrid spectral libraries, and chromatogram libraries. They were also compared to a spectral library-free workflow (directDIA). This evaluation was applied to two independent human studies, investigating medium spiny neurons in Huntington's disease (Tshilenge et al.) and studying plasma exosomes during aging (Patel et al.), analyzed on a TripleTOF 6600, Orbitrap Tribrids and an Exploris 480 platform operating in DIA mode. We investigated the number of peptides and proteins identified and quantified in the DIA data sets using the different library-based and library-free workflows. Moreover, their quantitative

performances were compared in terms of reproducibility, number of quantified peptides/proteins, and number of significantly changing proteins as well as their ratios.

This work aims at providing deep insights into how generation, specificity and comprehensiveness of the spectral and chromatogram libraries can influence DIA data processing workflows and performances to discover potential candidate biomarkers and highlight PTM patterns, while offering a better understanding of the fundamental mechanisms of neurodegeneration and aging.

This work was supported by the National Institute on Aging (U01 AG060906, PI: Schilling).

Contact: Joanna Bons, jbons@buckinstitute.org

P09.02

Missing data imputation for efficient label-free proteomics in non-human primates

Zeeshan Hamid (1), Cun Li (2), Peter Nathanielsz (2), Laura Cox (3), Michael Olivier (3)

(1) Center for Precision Medicine,Wake Forest University School of Medicine, Winston-Salem, NC, United States

(2) Southwest National Primate Research Center, San Antonio, TX, United States

(3) Center for Precision Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, United States

Reliable and effective label-free proteomics analyses are not only dependent on the method of data acquisition in the mass spectrometer, but equally dependent on the downstream data processing, including software tools, query database, data normalization and imputation. Label-free proteomics becomes particularly challenging if the query database of a model organism is not comprehensively annotated as in the case of non-human primate species, usually resulting in higher fraction of missing data points. Leaving missing data unchecked without imputation can limit downstream analyses and biological interpretation of the data. In this study, we compared various imputation approaches for label-free proteomics in non-human primates to maximize quantitative data information.

Our study included 45 brain frontal cortex samples from baboons in the age range of 6-23 years (22-80 human equivalent years). Samples were homogenized and protein extracts were reduced, alkylated and digested using trypsin. Resulting peptides were separated on a C18 easy spray column and analyzed on Fusion Lumos mass spectrometer.

Raw data analysis using MetaMorpheus software identified 2852 proteins. Only 112 proteins were identified with no missing values in all 45 samples. For improving quantitative data information and fill in the missing data points, we tested 20 different imputation methods on our dataset. Individual values were manually removed and imputed back to check the imputation accuracy of each method. The top performing method (Glmnet Ridge Regression-GRR) was tested further to assess the effect of missing data percentage on the imputation accuracy. Our analysis suggests that imputation methods taking overall data structure into consideration perform better compared to methods which replace the missing data with a constant minimum value. Here, the imputation accuracy is near 90% even if 50% of the data are missing. Overall our study offers a detailed comparative analysis of various imputation approaches for label free proteomics in non-human primates.

Contact: Zeeshan Hamid, zhamid@wakehealth.edu

P09.03

OpenTIMS, TimsPy, and TimsR: Quick and Friendly Access to timsTOF Pro Data

Mateusz Krzysztof Łącki (1), Michał Piotr Startek (2), Sven Brehmer (3), Christina Bell (3), Ute Distler (1), Stefan Tenzer (1)

(1) Institute of Immunology, University Medical Center of the Johannes-Gutenberg University Mainz, Mainz, Rhineland-Palatinate, Germany

(2) Department of Mathematics, Informatics, and Mechanics, University of Warsaw, Warsaw, Poland

(3) Bruker Daltonik GmbH, Bremen, Bremen, Germany

The timsTOF Pro is a novel mass spectrometer recently introduced by Bruker. It uses trapped ion mobility spectrometry (TIMS) together with liquid chromatography (LC) and mass spectrometry (MS) to further extend the platform's analytical capabilities and boost confidence in compounds' characterization. Collected data is stored in the vendor's Tims Data Format (TDF). Stored data includes scans, time of flights, and intensities of centroided peaks per-frame basis. Frames translate into retention times, scans into drift times, and time of flights into mass to charge ratios. Typically, the data occupies only a few gigabytes of storage, opening the possibility to access and process it on a modern laptop. Here we introduce OpenTIMS, TimsPy and TimsR. Open TIMS is a C++ library for low level access to TDF. It is coupled to Python (opentimspy) and to R (opentimsr) that export the data into higher-level data structures. Additionally, TimsPy and TimsR deliver data in user friendly tabular format and enable basic plotting. The outputs can be easily processed using well known data-scientific libraries, like Numpy, Pandas, or data.table. Data can be also converted to a simple vaex-compatible HDF5 format. With vaex, out-of-memory data can be quickly and automatically analyzed even on consumer laptops. OpenTIMS is open-source and works on major OS platforms, such as Windows, Linux, and macOS. The m/z and drift time calibrations are patent protected and closed source, but can be seamlessly integrated into OpenTIMS under separate licence agreements. OpenTIMS, opentimspy and opentimsr are available on github free of charge at https://github.com/michalsta/opentims. Additionally, opentimspy/TimsPy and opentimsr/TimsR are available for download from Pypi and Rcran repositories, respectively. To overcome some of the computational limitations of the Pandas library, OpenTIMS includes a simple script translating Bruker`s data into a vaex-compatible HDF5 for efficient, multithreaded and lazy aggregation of the data using vaex.

Contact: Mateusz Krzysztof Łącki, matteo.lacki@gmail.com

P09.04

MASH Explorer, a Universal, Comprehensive, and User-friendly Software Environment for Top-down Proteomics

Sean J. McIlwain (1), David S. Roberts (2), Jake A. Melby (2), Kent Wenger (3), Eli J. Larson (2), Zhijie Wu (4), Anna Janicek (5), Molly Wetzel (3), Yiran Yan (6), Elizabeth F. Bayne (2), Kyndalanne A. Pike (2), Xiaowen Liu (7), Ruixiang Sun (8), Irene M. Ong (9), Ying Ge (10)

(1) Biostatistics and Medical Informatics, University of Wisconsin - Madison, Madison, WI, United States

(2) Department of Chemistry, University of Wisconsin - Madison, Madison, WI, United States

(3) Department of Cell and Regenerative Biology, University of Wisconsin - Madison, Madison, WI, United States

(4) Department of Chemistry, University of Wisconsin, Madison, Madison, WI, United States

(5) Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, United States

(6) Human Proteomics Program, School of Medicine and Public Health, University of Wisconsin - Madison, Madison, WI, United States

(7) Department of BioHealth Informatics, Indiana University, Indianapolis, IN, United States

(8) National Institute of Biological Sciences, Beijing, 102206, China

(9) Department of Biostatistics and Medical Informatics and Department of Obstetrics and Gynecology, University of Wisconsin - Madison, Madison, WI, United States

(10) Department of Chemistry and Department of Cell and Regenerative Biology and Human Proteomics Program, University of Wisconsin - Madison, Madison, WI, United States

INTRODUCTION

Top-down mass spectrometry (MS)-based proteomics enables comprehensive analysis of proteoforms with molecular specificity to achieve a proteome-wide understanding of protein functions. However, the lack of a universal software tool for top-down proteomics has been increasingly recognized as a major barrier, especially for newcomers. Herein, we present MASH Explorer, a universal, comprehensive, and user-friendly software environment for top-down proteomics. MASH Explorer integrates multiple spectral deconvolution and database search algorithms into a single, universal platform which can process top-down proteomics data from various vendor formats.

METHODS

MASH Explorer is a multithreaded Windows application implemented in C# using the .NET framework within the Visual Studio integrated development environment. The software visual components are provided by Microsoft Office Runtime Support. Importing data obtained from different MS instruments is supported using ProteoWizard, DeconEngine, and vendor-provided libraries. FLASHDeconv and UniDec are the newest additions to the suite of previously available deconvolution engines.

RESULTS

We have developed MASH Explorer as a comprehensive software environment for top-down proteomics (https://labs.wisc.edu/gelab/MASH_Explorer/index.php). The core functions of MASH Explorer include spectral deconvolution, protein identification, proteoform characterization, graphical data output, data validation, and workflow automation. MASH Explorer can process data from multiple vendor formats and features workflows for discovery- and targeted-based top-down proteomics, with newly added FLASHDeconv and UVPD support. Moreover, with the addition of UniDec, MASH Explorer now supports native MS and ion mobility spectrometry (IMS).

CONCLUSIONS

MASH Explorer offers a universal platform for processing top-down MS data. It addresses an urgent need in the rapidly growing top-down proteomics community and is freely available to all users worldwide. With the critical need and tremendous support from the community, we envision that the MASH Explorer software application will play an integral role in advancing top-down proteomics to realize its full potential, including features for handling native MS, IMS, and UVPD data.

Contact: Sean McIlwain, sean.mcilwain@wisc.edu

P09.05

PINE: An Automation Tool to Extract and Visualize Protein-Centric Functional Networks

Niveda Sundararaman (1), James Go (1), Aaron E Robinson (1), José M Mato (2), Shelly C Lu (3), Jennifer Van Eyk (1), Vidya Venkatraman (1)

(1) Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA

(2) CIC bioGUNE, Technology Park of Bizkaia, Bizkaia, Spain

(3) Division of Digestive and Liver Diseases, Cedars-Sinai Medical Center, Los Angeles, CA, USA

Recent surges in mass spectrometry-based proteomics studies demands a concurrent rise in speedy and optimized data processing tools and pipelines. Although several stand-alone bioinformatics databases and tools exist that provide protein-protein interaction and annotation information, we developed Protein Interaction Network Extractor (PINE) with a fully-automated, user-friendly, graphical user interface that has an easily adoptable framework suitable for study of global proteome and post-translational modifications (PTMs), opening the possibility for simple, quick and robust visualization of differentially expressed proteins or PTMs and their interaction or enriched functional networks. PINE also supports overlaying differential expression, statistical significance thresholds and PTM sites on functionally enriched visualization networks to gain insights into proteome-wide regulatory mechanisms as well as how PTMs might regulate PPIs. To illustrate the relevance of the tool, we explore the total proteome and its PTM-associated relationships in two different nonalcoholic steatohepatitis (NASH) mouse models to show different context-specific use cases. The strength of this tool relies in its ability to (1) perform accurate protein identifier mapping to resolve any ambiguity, (2) retrieve interaction data from multiple publicly available protein-protein interaction databases, (3) assimilate these complex networks into meaningful functionally enriched pathways. PINE also supports analysis of multi-omics datasets to highlight their interrelationships and functional impact by building cohesive networks that capture the flow of biological information across the different molecular layers. Ultimately this tool can be used as an extremely powerful approach for novel hypothesis generation towards understanding disease mechanisms and potential therapeutics.

Contact: Niveda Sundararaman, niveda.sundararaman@cshs.org

Glycoproteomics and Glycomics

P10.01

A single-pot ultrasensitive heterogeneity analysis of viral spike N-glycoforms

Sabyasachi Baboo (1), Jolene Diedrich (1), Salvador Martínez-Bartolomé (1), Xiaoning Wang (1), Torben Schiffner (2), William Schief (2), James Paulson (1), John Yates (1)

(1) Molecular Medicine, The Scripps Research Institute, La Jolla, California, United States of America

(2) Immunology and Microbiology, The Scripps Research Institute, La Jolla, California, United States of America

Broadly specific proteases are theoretically capable of completely resolving a protein sequence. We show that given specific conditions, a single broadly specific protease when provided a substrate like heavily N-glycosylated proteins, generates peptides of ideal lengths for LC-MS/MS. Coupled with glycosidases, we cover >95% protein sequence by bottom-up proteomics. We describe a "single pot" method to analyze heterogeneity

in N-glycoforms of viral spikes, which is 50-100 times more sensitive than existing methods. HIV vaccine-design strategies targeting broadly neutralizing antibodies are dependent on heterogeneity of N-glycans decorating HIV Env protein. This novel method will significantly increase sensitivity and throughput for analyzing viral spikes, boosting precision in vaccine design.

Contact: Sabyasachi Baboo, sbaboo@scripps.edu

P10.02

Creating an Analytical Workflow to Assess N-linked Glycan Species and Their Alterations in Alzheimer's Disease Brain Tissue

Karen Butler (1), Jaclyn Kalmar (1), Brendan MacLean (2), Michael MacCoss (2), Edward Fox (3), Thomas Montine (3), David Muddiman (1), Erin Baker (1)

(1) Department of Chemistry, North Carolina State University, Raleigh, North Carolina, United States

(2) Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, United States

(3) Department of Pathology, Stanford University, Stanford, California, United States

N-linked glycosylation is a ubiquitous post-translational modification known to have essential biological roles including the regulation of cellular signaling, protein folding, and inflammation response. Perturbations in N-linked glycosylation have been implicated in a variety of disease states, including neurodegenerative diseases such as Alzheimer's disease (AD). However, the structure-function relationship of each glycan is often difficult to discern given the numerous isomers present in complex samples. Traditional glycan analytical methods involve liquid chromatography separations prior to mass spectrometry assessments (LC-MS). Due to the highly isomeric nature of this heterogeneous class of molecules, LC-MS/MS alone is often insufficient for resolving the glycan isomers present. Thus, in this work we coupled LC, ion mobility spectrometry, and MS (LC-IMS-MS) separations and utilized a novel derivatization strategy to assess glycans in the brain tissue of AD patients. The derivation strategy, termed, Individuality Normalization when Labeling with Isotopic Glycan Hydrazide Tags (INLIGHT™), uses both a natural (NAT) and stable-isotope label (SIL) with a distinctive six Dalton mass difference. By combining INLIGHT™ with LC-IMS-MS, the NAT and SIL labeled glycans are mobility aligned for confident identifications as the incorporation of $^{13}C_6$ does not significantly perturb the glycan structure. Additionally, the INLIGHT™-derivatized glycans exhibit m/z versus IMS trend lines, facilitating their rapid identification in the complex biological samples. The LC, IMS, and MS information garnered from these experiments was also used to create an N-linked glycan library in the open-source data analysis program Skyline, allowing for their rapid identification in the various biological samples studied.

Contact: Karen Butler, kebutler@ncsu.edu

P10.03

GPQuant: An Algorithm for Unidentified Glycopeptide Recovery and Glycosylation Quantification

Shao-Yung Chen (1), Yingwei Hu (1), Hui Zhang (1)

(1) Johns Hopkins University, Department of Pathology, Baltimore, MD, United States

Protein glycosylation accounts for one of the most complex protein modifications, with aberrant glycosylation reported to be associated with different diseases. While multiple attempts have been made to characterize the glycosylation event, there is still a need for an algorithm for accurately quantifying glycosylated peptides, and further search for unidentified glycosylated peptide under a LC-MS/MS experiment. There are several reasons for intact glycopeptides (IGPs) being harder to detect, including IGPs are harder to ionize, the intensity of IGPs are usually suppressed by unmodified peptides, which leads lack of MS2 scans under a Data-Dependent Analysis (DDA) setting. Here, we report an algorithm, GPQuant, which recovers unidentified IGPs by using retention time (RT) of the glycopeptide with different glycan as a reference, comparing the theoretical isotopic distributions based on the molecular composition, and peak integrity using our in-house evaluation system. After recovering the unidentified glycopeptides, GPQuant then proceeds to accurately quantify all the IGPs, identified or previously unidentified. With GPQuant, peptides with different glycoforms are comprehensively described, therefore, a relationship for different glycoforms on the same glycosylation site can be more accurately described. This could potentially benefit future studies regarding glycoproteomics in disease monitoring. Together, GPQuant introduces a novel approach in glycoproteomics. By using unidentified IGP recovery, GPQuant improves the reproducibility of IGPs, while also accurately describe the relative glycan distribution of a glycoprotein under a complex biological sample

Contact: Shao-Yung Chen, schen158@jhu.edu

P10.04

Optimization and application of EThcD-MS based glycoproteomics for identifying specific glycoproteins in different cancer cell lines

Hsiang-En Hsu (1), Junmin Wang (1), Shao Huan Samuel Weng (1), Raghothama Chaerkady (1), Matthew Glover (1), Lisa Cazares (1), Elaine Hurt (2), Sonja Hess (1)

(1) ADPE, AstraZeneca, Gaithersburg, Maryland, United States

(2) Early Oncology, AstraZeneca, Gaithersburg, Maryland, United States

Protein glycosylation has drawn increasing attention for its functions in numerous cellular processes, including cell differentiation and trafficking, protein stability, and signal transduction. Aberrant glycosylation of proteins has been described as a hallmark of cancer. Recently, electron-transfer/higher-energy collision dissociation fragmentation (EThcD) was introduced for improved site-specific glycoproteomic analysis, facilitating broad applications in biology. In this study, we used our optimized workflow to perform an in-depth analysis of the glycoproteome and glycan heterogeneity, associated with cancer progression. We analyzed 10 human cell lines, including 9 carcinoma cell lines (liver, breast, lung, pancreatic, colorectal, cervical, prostate, and leukemia) and HEK293 normal cell line and identified a total of 6,043 proteins. Most importantly, we observed 46 up-regulated and 32 down-regulated cancer specific proteins when compared to normal cells. The glycosylation profile analysis, led to the identification of ~310 glycoproteins in the cancer cell lines tested. Among the glycoproteins identified, 30 and 27 of those exhibited increased fucosylation and sialylation modifications, respectively, when compared to normal cells. Since our current research aim is to identify therapeutic targets for hepatocellular carcinoma (HCC), we focused on these cell types for a detailed comparison. In our global proteomic analysis, 38 proteins were significantly up-regulated in HCC cell lines, among which, components of glycolytic pathways were highly correlated. In contrast, 19 proteins associated with the apoptosis pathway were down-regulated in HCC cells. Next, we performed the site-specific glycosylation analysis. We found 13 fucosylated and 15 sialylated potential candidate proteins in HCC cell lines when compared to normal cells. Additionally, this glycoproteomic analysis revealed many surface proteins specific to each cancer cell type. In future work, this workflow will be implemented to analyze HCC clinical tissue samples and has the potential to provide valuable characterization of surface glycoproteins and glycan profiles to facilitate therapeutic drug discovery.

Contact: Hsiang-En Hsu, carter700507@gmail.com

P10.05

Novel mucin enrichment strategy reveals molecular signatures of cancer in cellular systems and ovarian cancer patient ascites fluid

Stacy Malaker (1), Nicholas Riley (2), Judy Shon (2), Kayvon Pedram (2), Carolyn Bertozzi (2)

(1) Department of Chemistry, Yale University, New Haven, CT, USA

(2) Department of Chemistry and ChEM-H, Stanford University, Stanford, CA, USA

Mucin domains are densely O-glycosylated modular protein domains found in a wide variety of cell surface and secreted proteins. When strung together in “tandem repeats”, mucin domains can form the large structures characteristic of mucin family proteins, such as MUC1 and MUC16. These canonical mucins are known to be key players in a host of human diseases, especially cancer. However, several other glycoproteins are known to contain mucin domains, such as: CD43, PSGL-1, and CD45, and their mucin domains are key to their molecular functions. However, a comprehensive list of mucin-domain glycoproteins does not exist, and this knowledge is critical in unraveling the role of mucin domains in health and disease.

Recently, we characterized a bacterial mucinase, StcE, and used its unique properties to improve analysis of mucin-domain glycoproteins by mass spectrometry. Given StcE’s selectivity for mucins, we reasoned that it could be employed to purify mucins from complex mixtures. Thus, the inactivated point mutant of StcE was conjugated to beads and we demonstrated that we could selectively enrich canonical mucins from HeLa cell lysate. However, we were faced with the issue of how to define a mucin domain. Thus, we developed a “mucin candidate algorithm” in order to calculate which human proteins have a high probability of bearing a mucin domain.

With this program in-hand, we expanded our “mucinomics” platform to several other cell lines and crude ovarian cancer patient ascites fluid. We demonstrated high mucin overlap between ovarian cancer patients, and the enrichment strategy allowed us to detect hundreds of glycopeptides from the mucin proteins. Finally, we used a KRAS dox-inducible system to show which mucins contribute to molecular bulk at the cell surface. Ultimately, we demonstrate this mucinomics platform can be used in a wide variety of systems to define key molecular signatures of cancer.

Contact: Stacy Malaker, stacy.malaker@yale.edu

P10.06

An Improved Search Method for Glycosylation and Other Labile Post-translational Modifications using MSFragger-Glyco

Daniel Polasky (1), Fengchao Yu (1), Guo Ci Teo (1), Alexey Nesvizhskii (2)

(1) University of Michigan, Department of Pathology, Ann Arbor, MI, USA

(2) University of Michigan, Departments of Pathology and Computational Medicine and Bioinformatics, Ann Arbor, MI, USA

Identifying and localizing protein post-translation modifications by mass spectrometry is a critical task complicated by fact that many such modifications are labile and dissociate from peptides during tandem MS. Traditional database search algorithms often struggle to identify peptides following PTM loss as the peptide sequence mass does not match the observed precursor mass from MS1. We recently developed MSFragger-Glyco to address these issues in the analysis of glycoproteomics data. Glycosylation is a ubiquitous and heterogeneous post-translational modification (PTM) used to accomplish a wide variety of critical cellular tasks, and is extremely labile in positive mode MS. We modified mass offset search of MSFragger to check for diagnostic fragment ions, restrict localization of delta masses to specific amino acids or sequons, and consider fragment ions that retain a partial (fragmented) modification. Applying these modifications to glycoproteomics data, we demonstrate an 80% improvement in glycoPSMs annotated from a combined HCD and AI-ETD analysis. In a complex O-glycoproteomics dataset analyzed by HCD alone, we obtain twice as many glycoPSMs as the original analysis using similar search parameters, or more than four times as many when using the speed of MSFragger to expand the search space to include many more possible glycan compositions and peptide sequences.

The adaptations to MSFragger for glyco-searches are generalizable to other labile PTMs by altering the diagnostic ions and partial fragment masses. We show that the degree of improvement for our labile search over traditional search methods scales with how labile a particular modification is. While the greatest improvements in performance are observed in glycoproteomics data, we also demonstrate improved identification rates for ADP-ribosylation and phosphorylation. In quantitative phosphoproteomics data in particular, the presence of isobaric TMT tags increases the probability of phosphate loss, resulting in improved performance of our labile search over traditional methods.

Contact: Daniel Polasky, dpolasky@umich.edu

P10.07

Effective Mass Spectrometry Methods for Global Analysis of Protein Glycosylation in Complex Biological Samples

Ronghu Wu (1)

(1) Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, USA

Protein glycosylation plays essential roles in mammalian cells because it determines protein folding, trafficking and stability. Moreover, it regulates nearly every extracellular activity, including cell-cell communication, cell-matrix interactions and cell immune response. Currently mass spectrometry (MS)-based proteomics provides a unique opportunity to analyze proteins and their modifications on a large scale. However, global and site-specific analysis of glycoproteins is extraordinarily challenging because of the low abundance of many glycoproteins, the heterogeneity of glycans and the complexity of biological samples. We have developed effective MS-based chemical and enzymatic methods to globally analyze glycoproteins in complex biological samples. For example, regarding boronic acid-based methods, normally the interactions between boronic acid and glycan is relatively weak. Therefore, they are not effective to capture low-abundance glycoproteins. We took several steps to make the methods more effective, including testing different boronic acid derivatives, enriching glycopeptides under various conditions and benefiting from the synergistic interactions between boronic acid and glycan. In addition, we have developed integrative methods combining chemical and enzymatic reactions to characterize glycoproteins with a particular and important glycan. Considering the importance of glycoproteins, the newly developed methods will have extensive applications in the biological and biomedical research fields.

Contact: Ronghu Wu, ronghu.wu@chemistry.gatech.edu

P10.08

O-live: An Interactive O-linked Glycosylation Database

Yuanwei Xu (1), Hui Zhang (2)

(1) Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland, United States

(2) Department of Pathology, Baltimore, Maryland, United States

Protein glycosylation takes up more than half of the proteome, yet our understanding of it remain at infancy. The highly versatile nature of glycosylation, especially O-linked glycosylation, in adaption of the dynamic changes in physiological and pathological states, has always sparked the interest of scientists. The lack of unity of O-linked glycosylation database, poses challenges on developing suitable analytical methodologies for O-linked glycoproteomic data analysis.

Owing to the recent discovery of an endoprotease that catalyzes the hydrolysis of peptide bonds at O-linked sites, a solid phased based O-linked glycopeptide enrichment method was established (EXoO). In EXoO, peptides are conjugated onto solid phase, and O-linked glycopeptides are released specifically by the O-linked endoprotease while preserving the glycan, O-linked glycosite and glycopeptide information. Paired with liquid chromatography tandem mass spectrometry (LC MS/MS), the comprehensive characterization of intact O-linked glycopeptide at a large-scale is made possible. In a single experiment using EXoO, over 3,000 O-GalNAcylation sites could be mapped for the first time in complex biological samples.

Along with the increase in intact O-linked glycopeptide data availability, a combined data analysis approach is also applied for the interpretation of these data generated by EXoO, where identifications of the peptide portion are generated using protein database search algorithm (SEQUEST) while the O-linked glycan portion is matched using intact glycopeptide search algorithm (GPQuest). We established a corresponding interactive database specifically for O-linked glycosylated peptides, O-live, where users could get access to the spectrum, O-linked glycans, O-linked glycosylation sites, intact O-linked glycopeptides, and protein information. The O-live serves as O-linked glycoprotein database resource for data analysis algorithms and biological or clinical investigations.

Contact: Yuanwei Xu, yxu105@jh.edu

Metabolomics and Imaging Mass Spectrometry

P11.01

Three-Dimensional (3D) Imaging of Lipids in Skin Tissues with Infrared Matrix-assisted Laser Desorption Electrospray Ionization (MALDESI) Mass Spectrometry

Honxgia Bai (1), Keith Linder (2), David Muddiman (1)

(1) Department of Chemistry, North Carolina State University, Raleigh, NC, US

(2) Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, US

Three-dimensional (3D) mass spectrometry imaging (MSI) has become a growing frontier as it has the potential to provide a 3D representation of analytes in a label-free, untargeted, and chemically specific manner. The most common 3D MSI is accomplished by the reconstruction of 2D MSI from serial cryosections; however, this presents significant challenges in image alignment and registration. An alternative method would be to sequentially image a sample by consecutive ablation events to create a 3D image. In this study, we present the use of infrared matrix-assisted laser desorption electrospray ionization (IR-MALDESI) in ablation-based 3D MSI for analyses of lipids within fresh frozen skin tissue. Depth resolution using different laser energy levels was explored with a confocal laser scanning microscope to establish the imaging parameters for the skin. The lowest and highest laser energy level resulted in a depth resolution of 7 microns and 18 microns, respectfully. A total of 594 lipids were putatively detected and detailed lipids profiles across different skin layers were revealed in a 56-layer 3D imaging experiment. Correlated with histological information, the skin structure was characterized with differential lipids distributions with a lateral resolution of 50 µm and a z resolution of 7 µm.

Contact: Hongxia Bai, hbai3@ncsu.edu

P11.02

A high-throughput global targeted MRM method for serum metabolites using a triple quadrupole MS

Kranthi Chebrolu (1), Robert Proos (1), Darren Dumalao (1), Baljit Ubhi (1), Christie Hunter (1)

(1) Sciex, Redwood City, CA, US

Over 95% of the serum metabolites are well documented but there are not many global targeted methods that could quantify a majority of these compounds. Relative and absolute quantitation of plasma metabolite methods have great utility for many applications including disease research, food Omics, pharmaceutical development, etc. The wide structural diversity and ionization properties of these compounds have posed an immense challenge in developing high throughput, comprehensive global metabolomic methods that can be used for semi quantitative purposes. Here, a broad, high throughput serum metabolite method consisting of over 400 serum metabolites was developed and tested using SRM 1950 plasma sample. The chromatographic separations were performed on a Kinetex F5 (Phenomenex, CA) column using standard reversed- phase mobile phases for reliable separations for a majority of serum analytes. The raw MRM data was acquired on both the SCIEX Triple Quad™ 7500 System QTRAP® Ready and the QTRAP 6500+ System and processed on a SCIEX OS Software 2.0. Development of the MRM methods was done both from standards and also leveraging the HMDB information. The collision energies for all the analytes were optimized to obtain a reliable S/N for relative quantitation. Polarity switching and Scheduled MRM™ Algorithm were used to enable the robust analysis of this large panel of analytes in a single injection. Finally, the sMRM Builder template was also used for method optimization, helping with retention time scheduling and filtering to detected analytes, streamlining method transfer across different sciex platforms and across the labs. This method currently provides semi quantitation of amines, amino acids, purines, organic acids, phenolic metabolites, nucleosides, vitamins and sugars from plasma in 20 min. Significant sensitivity gains were observed when running the method on the SCIEX 7500 System, with S/N gains of 3-5 fold on average in positive and negative ion mode.

Contact: Kranthi Chebrolu, kranthi.chebrolu@sciex.com

P11.03

Acute stress effects: Let's ask metabolomics!

Aikaterini Iliou (1), Angeliki-Maria Vlaikou (2), Markus Nussbaumer (2), Dimitra Benaki (1), Emmanuel Mikros (1), Evangelos Gikas (3), Michaela Filiou (2)

(1) Section of Pharmaceutical Chemistry, Department of Pharmacy, School of Health Sciences, National and Kapodistrian University of Athens, Athens, Greece

(2) Laboratory of Biochemistry, Department of Biological Applications and Technology, School of Health Sciences, University of Ioannina, Ioannina, Greece

(3) Section of Analytical Chemistry, Department of Chemistry, School of Science, National and Kapodistrian University of Athens, Athens, Greece

Psychological stress and stress-related disorders constitute a major health problem worldwide. Although several brain regions have been implicated in emotional processing, little is known about the involvement of cerebellum in stress response mechanisms. In this study, the effects of acute stress exposure to cerebellum metabolome were investigated in mice, using the forced swim test (FST) as an acute stressor. A Nuclear Magnetic Resonance (NMR)-based metabolomics approach was employed to analyze the metabolomic profiles of the cerebellum of stressed and unstressed mice. The statistical analysis showed clear differentiation between the two groups, indicating that cerebellum metabolome is involved in mediating stress responses. Significant alterations were observed for 19 out of the 47 annotated metabolites related to purine/pyrimidine metabolism, amino acids, neurotransmitters and organic acids. Pathway enrichment analysis revealed aspartate metabolism as the most affected pathway, followed by the urea cycle and purine metabolism. We also correlated individual metabolite levels with FST behavioral parameters, with amino acid

neurotransmitters exhibiting a strong association with FST behavioral parameters. The present study shows that acute stress affects the cerebellum metabolomics and underlines the implication of the cerebellum in stress responses.

Keywords: Forced swim test (FST), Cerebellum, Metabolomics, NMR, Acute stress, Mouse

ACKNOWLEDGEMENTS: We thank Dr. Panagiotis Lekkas and the personnel of the University of Ioannina animal facility for expert animal care. This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014- 2020» in the context of the project “How mitochondria regulate stress? Let’s ask metabolomics” (MIS 5047652).

Contact: Aikaterini Iliou, katerinail@pharm.uoa.gr

P11.04

Profiling the Lipidome: Quantification of Over 2000 Lipid Molecular Species in Under 20 Minutes

Mackenzie Pearson (1), Paul Norris (1), Christie Hunter (2), Paul Baker (3)

(1) Applications, SCIEX, Danville, IN, United States

(2) Marketing, SCIEX, Redwood City, CA, United States

(3) Commericial, SCIEX, Birmingham, AL, United States

The inherent complexity of the lipidome means that identifying and quantifying different lipid classes and isobaric species in a single analysis method is difficult, and often requires a high degree of expertise and method development time. The lack of simplified methods and poor quality of quantitative data makes it more challenging to understand the role of lipids in biology. Here, a targeted, relative quantitation, LC-MRM method is presented that provides comprehensive coverage of many of the multiple lipid classes present in complex biological samples.
In this assay, MRMs to quantitate almost 2000 lipid molecular species were combined into a single assay on the LC-MS/MS system. Amino column chemistry was chosen for lipid class separation to minimize isomeric interference. The target list of lipids is comprehensive, covering 19 lipid classes and categories, and MRMs were selected to cover lipids containing fatty acids with 14 to 26 carbons and 0 to 6 double bonds. The sMRM Pro Builder template was used to streamline method optimization, assigning the retention times, and setting detection window size per MRM to enhance the coverage and sensitivity of the method. This optimization improved results quality, especially on low abundant lipids. Lipid internal standards from 20 different classes, with either heavy isotopic labeled lipids or odd chain lipids (Lipidyzer Standards or SPLASH mix) were used. The method is customizable, so new lipid categories, classes and molecular species can be added to the MRMs list and is matrix agnostic. Ten replicates of NIST plasma was used to test the assay on the SCIEX Triple Quad™ 7500 System. Compared to previously established methods, over 20% more lipids could be identified in a third less of the time. In a total 17-minute chromatographic run, 1046 lipid molecular species were quantified from NIST in a single injection with very good reproducibility.

Contact: Mackenzie Pearson, mackenzie.pearson@sciex.com

Multi-omics: Methods and Applications

P12.01

A Novel Multi-omics Strategy for Human Pluripotent Stem Cell-derived Cardiomyocytes

Elizabeth F. Bayne (1), Aaron Simmons (2), David S. Roberts (1), Yanlong Zhu (3), Timothy J. Aballo (4), Sean Palecek (2), Ying Ge (3)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

(2) Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA

(3) Department of Cell and Regenerative Biology and Human Proteomics Program, University of Wisconsin-Madison, Madison, WI, USA

(4) Molecular and Cellular Pharmacology Training Program, Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI, USA

Human pluripotent stem cell-derived cardiomyocytes (hPSC-CMs) show immense promise for precision medicine. However, the molecular underpinnings of hPSC-CM maturation are not well understood, and late-stage hPSC-CMs lack functional properties of adult CMs in vivo. To gain systems-level insight to hPSC-CM biology, we created multi-dimensional profiles of hPSC-CM phenotypes by connecting metabolome and proteome measurements from the same cell culture. Parallel access to biomolecules gives precise insight to the phenotype from which they arise by resolving biological and batch variability. Thus, we developed a novel cell culture-based dual extraction and mass spectrometry (MS)-based methods to quantitate metabolites and proteins from a single well of hPSC-CMs. Our method achieves rapid, reproducible access to the hPSC-CMs metabolome and proteome with no interference to cell culture conditions. Metabolites were extracted using a solvent-based quench technique while simultaneously precipitating proteins. Metabolite samples were analyzed by flow injection to Bruker 12T FT-ICR MS. Dehydrated proteins were resolubilized in a buffer containing MS-compatible surfactant Azo before MS analysis by Bruker timsTOF Pro. We detected 1,228 metabolite features in a single MRMS spectrum, spanning a range of feature intensities over 1000x and requiring only 3.5 min. We detected 3,400 protein groups per sample and achieved high quantitative correlation (0.96-0.98) of peptide spectral intensities. From 10^6 cells, we identified and quantified 259 metabolites and 3,229 proteins. Finally, we mapped proteins and metabolites to biochemical pathways using MetaboAnalyst integrative analysis. In total, 321 pathways were identified and highly populated pathways included ATP generation such as TCA cycle, oxidative phosphorylation, and glycolysis, as well as those involved in cardiac contraction. Overall, we achieved reproducible and broad coverage of complex metabolite and protein

profiles from a single monolayer culture of hPSC-CMs. This multi-omics approach can be used to quantitatively compare differentiation protocols to improve our understanding of hPSC-CM maturation.

Contact: Elizabeth Bayne, ebayne@wisc.edu

P12.02

Confirmatory Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma

David Clark (1), Yize Li (2)

(1) Department of Pathology, The Johns Hopkins University, Baltimore, MD, USA

(2) McDonnell Genome Institute, Washington University in St. Louis, St. Louis, MO, USA

Clear cell renal cell carcinoma (ccRCC) is the most predominant histology of renal cell carcinoma (RCC), accounting for 75% of all diagnosed cases and majority of associated deaths. Previously, the Clinical Proteomics Tumor Analysis Consortium (CPTAC) performed a comprehensive proteogenomic characterization of a cohort of 103 ccRCC tumors and 80 paired normal adjacent tissue (NAT) samples in an effort to understand the impact of genomic alterations on the functional phenotypes that drive ccRCC tumorigenesis. In this report, we sought to verify these initial findings in an independent cohort comprised of 112 RCC tumors and 87 paired NAT samples. We integrated data-independent acquisition (DIA) mass spectrometry-based proteomic and phosphoproteomic analyses with genomic, epigenomic, microRNA, and transcriptomic profiling for tumor and NAT samples to identify the dysregulated cellular mechanism resulting from genomic aberrations. Our results showed many of genomic and transcriptomic characteristics defined in the discovery cohort are robustly maintained in this confirmatory cohort. Further proteomic characterization identified protein profiles associated with disparate genomic alterations associated with ccRCC, as well as protein features specific to the four immune-based subtypes we defined in our previous report: CD8+ Inflamed, CD8- Inflamed, VEGF Immune Desert, and Metabolic Immune Desert. We provided new insight into the impact of FDA-approved kinase inhibitors, prioritized by our previous phosphoproteomic results, on ccRCC cells via the targeting of ERK/MAPK and PI3K/AKT signaling pathways, and G2/M cell cycle regulation. In addition, we explored the degree of intratumor heterogeneity of ccRCC at the genomic, transcriptomic, and proteomic levels, and linked these molecular signatures to observed histopathological features. Overall, this report provides a deeper understanding of functional consequences of the genomic alterations on gene and protein expression in ccRCC, and provides further rationale for integrating multi-level “omics” analyses to characterize the proteogenomic landscape of ccRCC pathobiology.

Contact: David Clark, dclark71@jhmi.edu

P12.04

OpenCMS: a tool to build personalized proteogenomics databases and detect unannotated proteins and possible biomarkers.

Noé Guilloy (1), Sébastien Leblanc (1), Jean-François Jacques (1), Marie Brunet (1), Claude Perreault (2), Xavier Roucou (1)

(1) Université de sherbrooke, departement of Biochemistry and Functional Genomics, Sherbrooke, Quebec, Canada

(2) Université de Montréal, Institut de Recherche en Immunologie et Cancérologie, Montréal, Quebec, Canada

Ribosome profiling studies indicate that up to 40% of initiating ribosomes map to previously non-annotated open reading frames (ORFs), suggesting that a large number of proteins are coded by unconventional coding sequences. In addition to the currently annotated protein-coding sequences, the proteogenomc resource OpenProt annotates all ORFs with a minimum size of 30 codons present in the transcriptome across 10 species. These novel ORFs are termed alternative ORFs and potentially code for alternative proteins (altProts), which are different from currently annotated proteins (refProts). Thus, the OpenProt-derived protein databases allow the detection of altProts by MS-based proteomics.

In Human, the whole OpenProt database contains the sequence of 488,956 altProts in addition to refProts. This represents a challenge for database searching of MS/MS data. Furthermore, common protein databases cannot be used to detect mutated proteins in patients. To address both issues, we implemented a proteogenomic approach, which allows the detection of both altProts and mutated proteins.

We developed a program called Open-Custom-Mass Spectrometry (OpenCMS). OpenCMS uses sample-specific DNA or RNAseq data to identify genomic variants and uses the OpenProt resource to generate custom databases of protein sequences. Each database includes altProts and mutated protein for MS-based proteomics analyses.

We tested OpenCMS with a cohort of 38 patients with acute myeloid or lymphoid leukemia. RNAseq data were generated from extracted myeloid and lymphoblastic cells and used to create personalized protein databases. Proteins were extracted from the same cells, separated by SDS-PAGE, in-gel digested and peptides were analyzed by LC-MS/MS. We used the peptide centric approach pepQuery to ensure that spectra matching to altProts could not be better assigned to any RefProt with any post translational modification. Preliminary results show the expression of 33 altProts and 136 mutated proteins, including 24 mutated altProts in acute lymphoid leukemia patients.

Contact: Noé Guilloy, noe.guilloy@usherbrooke.ca

P12.05

MS-based Proteomics and Metabolomics of MELAS Syndrome

Haorong Li (1), Martine Uittenbogaard (2), Andrea Gropman (3), Anne Chiaramello (2), Ling Hao (4)

(1) Department of Chemistry, George Washington University, District of Columbia, District of Columbia, United States

(2) Department of Anatomy and Cell Biology, George Washington University School of Medicine and Health Sciences, District of Columbia, District of Columbia, United States

(3) Division of Neurogenetics and Developmental Pediatrics, Children's National Medical Center, District of Columbia, District of Columbia, United States

(4) Department of Chemistry, District of Columbia, District of Columbia, United States

The mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes syndrome (MELAS, OMIM 540000) is a progressive neurodegenerative disease that results in devastating multi-organ failure due to chronic energy deficit. MELAS patients harbor a pathogenic mitochondrial variant that affects the mitochondrial oxidative pathway (OXPHOS) responsible for ATP synthesis. These MELAS mitochondrial variants are always maternally inherited, except in the rare case of de novo pathogenic variants. In the absence of curative treatment, MELAS patients only have access to palliative treatment that fails to halt the progression of the symptoms. To design novel therapeutic intervention, it is important to understand the

underpinning pathogenetic mechanisms, as the exact mechanisms underlying the metabolic alterations in MELAS patients remain unknown. Due to a lack of knowledge of MELAS biomarkers, the diagnosis of MELAS patients remains challenging and their prognosis unknown.

Our study aims to unravel the molecular MELAS signature. We used patient-derived dermal fibroblasts (N=4) from a symptomatic MELAS patient harboring a de novo pathogenic variant, m.14453G>A mapping in the MT-ND6 gene encoding the NADH dehydrogenase subunit 6, and her healthy mother. MS-based proteomics and metabolomics were conducted in parallel to achieve a comprehensive understanding of MELAS pathogenesis. We found that MELAS fibroblasts exhibited significantly dysregulated levels in mitochondrial proteins, such as downregulated NADH dehydrogenase subunits. We also discovered significantly altered levels of small molecules from key bioenergetic pathways, such as glycolysis, TCA cycle, and fatty acid oxidation.

In conclusion, our study has revealed novel molecular signatures that may function as MELAS biomarkers and provide insights into the pathogenic mechanism of MELAS. Our findings may also benefit secondary mitochondrial neurological diseases, such as Parkinson's disease and Alzheimer's disease. Our ongoing efforts focus on integrating multi-omics data into pathway analysis and validating metabolite and protein changes using targeted LC-MS assays with a comprehensive cohort of MELAS patients.

Contact: Haorong Li, haorong@gwu.edu

P12.06

Quantitative proteomics analysis of urine samples from COVID-19 patients

Catherine WONG (1), Nan Zhang (1), Shuaixin Gao (1), Wenmin Tian (1), Siyuan Wang (1)

(1) Peking University Health Science Center, Beijing, Beijing, China

The outbreak of COVID-19 has become a worldwide pandemic. The pathogenesis of this infectious disease and how it differs from other drivers of pneumonia is unclear. Here we analyze urine samples from COVID-19 infection cases, healthy donors and non-COVID-19 pneumonia cases using data independent acquisition quantitative proteomics. The molecular changes suggest that immunosuppression and tight junction impairment occur in the early stage of COVID-19 infection. Further subgrouping of COVID-19 patients into moderate and severe types shows that an activated immune response emerges in severely affected patients. We propose a two stage mechanism of pathogenesis for this unusual viral infection. Our data advance our understanding of the clinical features of COVID-19 infections and provide a resource for future mechanistic and therapeutics studies.

Contact: Nan Zhang, nzhang_we@bjmu.edu.cn

Post-Translational Modifications

P13.01 Real-time search of Ubiquitin diGLY modified peptides and PaSER acquisition control

Christopher Adams, Robin Park, Nagarjuna Nagaraj, Sven Brehmer, Matt Willetts, Francesco Pingitore, Hayley Peckham, Kimberley Lee, Tharan Srikumar

P13.02 Relating Metabolic Control with Lysine Acylation in an Exceptional Metabolically Versatile Bacterium

Janine Y. Fu, Robert P. Gunsalus, Rachel R. Ogorzalek Loo, Joseph A. Loo

P13.03 The regulatory enzymes and protein substrates for the lysine β-hydroxybutyrylation pathway

He Huang, Di Zhang, Zhanyun Tang, Cong Yan, Shankang Qi, Philip A Cole, Robert G Roeder, Yingming Zhao

P13.04 Metabolic regulation of gene expression by histone lactylation

Di Zhang, Zhanyun Tang, He Huang, Robert G Roeder, Lev Becker1, Yingming Zhao

P13.05 Discovery of In Situ Deamidation and Citrullination in Human Pancreatic Islets as Potential Neoepitopes for Type 1 Diabetes

Tai-Du Lin, Adam C Swensen, Paul D Piehowski, Ying Zhu, Jing Chen, Martha Campbell-Thompson, Clayton E Mathews, Wei-Jun Qian

P13.06 Systematic discovery of high abundance truncated protein variants

Claire McWhite, Nicolas Gort Frietas, Wisath Sae-Lee, Edward Marcotte

P13.07 FLEXIQuant-LF: A Modification-agnostic Robust Computational Approach to Quantify Protein Modification Extent in Label-Free Proteomics Data

Christoph N Schlaffner, Konstantin Kahnert, Jan Muntel, Pauline Hiort, Ruchi Chauhan, Bernhard Y Renard, Judith A Steen, Hanno Steen

P13.08 The Antiviral Sirtuin 3 Bridges Protein Acetylation to Mitochondrial Integrity and Metabolism during Human Cytomegalovirus Infection

Xinlei Sheng, Ileana Cristea

P13.09 OxRAC Unveils Novel Role of LARP1 Along the Photosynthetic TOR Pathway

Amanda Smythers, Sarah Lowery, Saher Mubarek, Hailey Lewis, Danielle Slone, Derrick Kolling, Leslie Hicks

P13.10 The DNA Sensor cGAS is Decorated by Acetylation and Phosphorylation Modifications in the Context of Immune Signaling

Bokai Song, Todd Greco, Krystal Lum, Caroline Taber, Ileana Cristea

P13.11 Role of posttranslational modification, citrullination in polarization to proinflammatory and anti-inflammatory macrophages

Aneta Stachowicz, Justyna Fert-Bober

P13.12 Unraveling the Histone Code and Its Readers on Decondensed Heterochromatin

Stephanie Stransky, Jennifer Aguilan, Edward Nieves, Simone Sidoli

P13.13 Novel bifunctional material coupled to PASEF enables sensitive extracellular vesicles phosphoproteomic analysis of urine sample

Jie Sun, W. Andy Tao, Xiaoxian Du, xianming Liu

P13.01

Real-time search of Ubiquitin diGLY modified peptides and PaSER acquisition control

Christopher Adams (1), Robin Park (1), Nagarjuna Nagaraj (1), Sven Brehmer (1), Matt Willetts (1), Francesco Pingitore (1), Hayley Peckham (2), Kimberley Lee (2), Tharan Srikumar (1)

(1) Bruker Daltonic, San Jose, CA, USA

(2) Cell Signaling Technologies, Danvers, MA, USA

Introduction: Recently we introduced PaSER (PArallel Search Engine in Real-time) which performs a database search in parallel with the 120Hz sequencing speed of the timsTOF Pro mass spectrometer. Biological processes are influenced by the PTM state of proteins, one such PTM (ubiquitination) drives protein catabolism. Here we show PaSER has the computational power to real time assign ubiquitinated peptides by identification of the remnant GlyGly tag on lysine residues. The search engine is so fast it can even perform this process under semi-tryptic search conditions which often add substantial search time. Methods: The mass spectrometer was a timsTOF Pro (Bruker) equipped with a nanoElute. The sample was a diGLY enriched lysate using a modified protocol of the K-ε-GG kit from Cell Signaling Technologies and a QC sample of K562. Using a “if then that” logic, a user defined metric (number of protein or peptide IDs) is used to inform the Hystar (Bruker) sequence acquisition software. Results: After data acquisition was completed so were the search results, where if the identification of ≥3500 proteins and ≥16000 unique peptide observations in a 21min gradient were achieved then the sequence queue progressed. Greater than 5000 unique diGly remnant peptides were identified in a single shot 21minute gradient from 100ng of peptide load. The observation of diGly positional isomers was also made where co-eluting peptides were only resolved using the collisional cross section (CCS) term. Offline results were identical to the Real-time search results. Conclusions: Real-time search for even the most arduous PTM search parameters are capable using PaSER technology.

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

P13.02

Relating Metabolic Control with Lysine Acylation in an Exceptional Metabolically Versatile Bacterium

Janine Y. Fu (1), Robert P. Gunsalus (2), Rachel R. Ogorzalek Loo (3), Joseph A. Loo (4)

(1) Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, USA

(2) Department of Microbiology, Immunology, and Molecular Genetics, UCLA-DOE Institute, University of California, Los Angeles, Los Angeles, CA, USA

(3) David Geffen School of Medicine, Department of Biological Chemistry, UCLA-DOE Institute, University of California, Los Angeles, Los Angeles, CA, USA

(4) Department of Chemistry and Biochemistry, UCLA-DOE Institute, University of California, Los Angeles, Los Angeles, CA, USA

Post-translational modification (PTM) of proteins is a conserved strategy used to efficiently control biological mechanisms in response to changing cellular environments. One type of modification, lysine acylation, has been implicated in metabolic pathways across many systems. Elevated metabolic intermediates, such as acyl-CoA, have been shown to modify enzymes and affect their activity, highlighting the relationship between protein acylation and metabolic regulation. Here, we investigate the metabolically versatile α-proteobacterium Rhodopseudomonas palustris proteome, focusing on identifying acyl-PTMs across this model system's diverse metabolic modes. R. palustris is unique among organisms in its ability to switch among four drastically different metabolic states and by mapping acyl-PTMs in this model system, we can further advance our knowledge of how these PTMs reflect the cellular metabolic state and directly modulate catabolism. To characterize the range of acyl-PTMs in this system, we utilized tandem mass spectrometry (LC-MS/MS) on an Orbitrap instrument and improved current methods for identifying these modifications, gaining unbiased proteomic depth. In our approach we utilize unique diagnostic ions that arise from these PTMs during our spectral assignments, allowing more confident targeting of acyl modifications in complex proteomic datasets. By incorporating these diagnostic ions in our analyses, we have confidently identified a wide range of acyl-PTMs, expanding beyond acetylation. For instance, we have characterized propionyl-, glutaryl-, and succinyl-lysine and have mapped these modifications across 8 diverse carbon sources and 2 metabolic modes. The type of acyl modification on a given peptide differs depending on the cell's growth condition, demonstrating how these PTMs can act as a marker of the cell's metabolic state. We have also found that the degree of modification shifts under different cellular conditions; for a given site, there is decreased acylation when cells are grown in more reduced environments.

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

P13.03

The regulatory enzymes and protein substrates for the lysine β-hydroxybutyrylation pathway

He Huang (1), Di Zhang (2), Zhanyun Tang (3), Cong Yan (1), Shankang Qi (1), Philip A Cole (4), Robert G Roeder (3), Yingming Zhao (2)

(1) Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, Shanghai, China

(2) Ben May Department for Cancer Research, The University of Chicago, Chicago, Illinois, USA

(3) Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York, USA

(4) Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA

Metabolism-mediated epigenetic changes represent an adapted mechanism for cellular signaling, in which lysine acetylation and methylation have been the historical focus of interest. We recently discovered a β-hydroxybutyrate-mediated epigenetic pathway that couples metabolism to gene expression. However, its regulatory enzymes and substrate proteins remain unknown, hindering its functional study. Here we report that the acyltransferase p300 can catalyze the enzymatic addition of β-hydroxybutyrate to lysine (Kbhb), while HDAC1 and HDAC2 enzymatically remove Kbhb, both in vitro and in vivo. We demonstrate that p300-dependent histone Kbhb can directly mediate in vitro transcription. Moreover, a

comprehensive analysis of Kbhb substrates in mammalian cells has identified 3248 Kbhb sites on 1397 substrate proteins. The dependence of histone Kbhb on p300 argues that enzyme-catalyzed acylation is the major mechanism for nuclear Kbhb. Our study thus reveals key regulatory elements for the Kbhb pathway, laying a foundation for studying its roles in diverse cellular processes.

Contact: He Haung, hhuang@simm.ac.cn

P13.04

Metabolic regulation of gene expression by histone lactylation

Di Zhang (1), Zhanyun Tang (2), He Huang (3), Robert G Roeder (2), Lev Becker1 (4), Yingming Zhao (1)

(1) Ben May Department for Cancer Research, The University of Chicago, Chicago, Illinois, USA

(2) Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York, USA

(3) Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, Shanghai, China

(4) Ben May Department for Cancer Research, The University of Chicago University of Chicago, Medicine Comprehensive Cancer Center, Chicago, Illinois, USA

The Warburg effect, originally describing augmented lactogenesis in cancer, is associated with diverse cellular processes such as angiogenesis, hypoxia, macrophage polarization, and T-cell activation. This phenomenon is intimately linked with multiple diseases including neoplasia, sepsis, and autoimmune diseases. Lactate, a compound generated during Warburg effect, is widely known as an energy source and metabolic byproduct. However, its non-metabolic functions in physiology and disease remain unknown. Here we report lactate-derived histone lysine lactylation as a new epigenetic modification and demonstrate that histone lactylation directly stimulates gene transcription from chromatin. In total, we identify 28 lactylation sites on core histones in human and mouse cells. Hypoxia and bacterial challenges induce production of lactate through glycolysis that in turn serves as precursor for stimulating histone lactylation. Using bacterially exposed M1 macrophages as a model system, we demonstrate that histone lactylation has different temporal dynamics from acetylation. In the late phase of M1 macrophage polarization, elevated histone lactylation induces homeostatic genes involved in wound healing including arginase 1. Collectively, our results suggest the presence of an endogenous “lactate clock” in bacterially challenged M1 macrophages that turns on gene expression to promote homeostasis. Histone lactylation thus represents a new avenue for understanding the functions of lactate and its role in diverse pathophysiological conditions, including infection and cancer.

Contact: He Huang, hhuang@simm.ac.cn

P13.05

Discovery of In Situ Deamidation and Citrullination in Human Pancreatic Islets as Potential Neoepitopes for Type 1 Diabetes

Tai-Du Lin (1), Adam C Swensen (1), Paul D Piehowski (1), Ying Zhu (2), Jing Chen (3), Martha Campbell-Thompson (3), Clayton E Mathews (3), Wei-Jun Qian (1)

(1) Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, United States

(2) Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, United States

(3) 3Department of Pathology, Immunology, and Laboratory Medicine, Gainesville, FL, United States

Proteins with abnormal deamidation/citrullination may act as neoepitopes which trigger the break of immune tolerance during type 1 diabetes (T1D) progression. Although mass spectrometry (MS)-based proteomics is capable of identifying these modifications, the resulting 0.9840 Da mass shift is often identified incorrectly due to the near overlap with naturally-occurring heavy-isotope-containing versions of the peptides. Recently, we had reported an improved algorithm, "Dual-search Delta Score strategy", for more confident identification of these modifications from a typical global shotgun proteomics dataset. Herein, we have combined nanoscale proteomics and Dual-search Delta Score strategy to identify deamidation and citrullination site in laser micro-dissected (LMD) human islet sections from pre-symptomatic autoantibody positive (AAb+) subjects and age/sex matched controls (n=6) as well as one T1D donor. In total, 1614 unique deamidated/citrullinated peptides from 432 unique proteins were identified. In addition, we applied the A-score as a probability-based algorithm for determining the PTM sites. Overall, the exact localization of 651 N-deamidated, 52 Q-deamidated, and 12 R-citrullinated peptides were identified (A-score >19). Interestingly, 639 deamidation/citrullination sites identified in our human islet dataset were reported as autoantigens in Immune Epitope Database (IEDB) including the known antigenic proteins, insulin and 78 kDa glucose-regulated protein (GRP78). Furthermore, 5 deamidated/citrullinated peptides were selected for further confirmation using synthetic peptides. The deamidated epitope, EAEDLQVGQ, of insulin had been confirmed as an important autoantigen in T1D patient was confirmation in our study. Our results not only demonstrate the confident identification of in situ deamidated/citrullinated proteins and confirmation the deamidated/citrullinated T1D autoantigens but also provide the potential deamidated/citrullinated T1D neoantigens.

Contact: Tai-Du Lin, tai-tu.lin@pnnl.gov

P13.06

Systematic discovery of high abundance truncated protein variants

Claire McWhite (1), Nicolas Gort Frietas (2), Wisath Sae-Lee (3), Edward Marcotte (3)

(1) Princeton University, Princeton, NJ, USA

(2) Harvard University, Boston, MA, USA

(3) The University of Texas at Austin, Austin, TX, USA

Proteolysis is an on-demand way to achieve rapid control over a protein's function. Short forms of proteins often adopt regulatory roles, including competing for interaction partners, and can have altered solubility or localization. We surveyed elution patterns of proteins in high-throughput fractionation experiments to detect short proteoforms that are found in high abundance, but do not correspond to annotated splice variants. We recover both known and novel short proteoforms in non-apoptotic human cell culture and blood, and plant species. A subset of short variants are found conserved across lineages, suggesting that they may date to their last common ancestor 1.5 BYA or potentially older. Several novel proteoforms have been shown recombinantly to have distinct function from the full-length protein, but have not previously been shown to exist endogenously in cells. High-throughput, multispecies data thus allows the discovery and validation of conserved, abundant, stable, short proteoforms.

Contact: Claire McWhite, cmcwhite@princeton.edu

P13.07

FLEXIQuant-LF: A Modification-agnostic Robust Computational Approach to Quantify Protein Modification Extent in Label-Free Proteomics Data

Christoph N Schlaffner (1), Konstantin Kahnert (2), Jan Muntel (2), Pauline Hiort (2), Ruchi Chauhan (1), Bernhard Y Renard (3), Judith A Steen (1), Hanno Steen (2)

(1) F.M. Kirby Neurobiology Center, Boston Children's Hospital, Boston, MA, USA

(2) Department of Pathology, Boston Children's Hospital, Boston, MA, USA

(3) Data Analytics and Computational Statistics, Hasso-Plattner-Institute, University of Potsdam, Potsdam, Germany

Computational methods employing searches with wide precursor error tolerances alongside improvements in LC-MS/MS methods and technology have resulted in the identification of thousands of post-translational modifications (PTMs) in a single experiment. While different PTMs now can readily be identified, quantifying these has become more challenging. Consistent quantification of individual post-translational modifications can only be achieved by specialized enrichment methods or targeted MS approaches requiring a priori knowledge of the modification in question. Here we introduce FLEXIQuant-LF as a robust computational method to identify differentially modified peptides and quantify the modification extent changes between samples from label-free LC-MS/MS data. FLEXIQuant-LF builds on the principle of our previous experimental method, FLEXIQuant, which employs a heavy isotope labeled unmodified full-length protein standard to quantify the modification extent of peptides. Our label-free method uses unmodified peptides to fit a robust linear regression between samples of interest and a reference to quantify the modification extent. Simulated modification data show that our method identifies differentially modified peptides with high sensitivity and precision. We also applied FLEXIQuant-LF to data-independent acquisition (DIA) data of the well studies anaphase promoting complex/cyclosome (APC/C) during mitosis. We identify differentially modified peptides and verify their modification through data dependent acquisition (DDA) data as well as literature. In summary, our results highlight that the ability to identify differentially modified peptides and quantify the modification extent of FLEXIQuant-LF is on par with the expensive heavy isotope labeled method. The use of a reference sample instead of labeled reference proteins allows FLEXIQuant-LF to be universally applicable post hoc to any protein robustly identified and quantified across samples. Our method therefore can help improve our understanding of peptide resolved PTM dynamics and guide researchers to study differential, rare, and potentially novel PTMs in more targeted approached in future experiments.

Contact: Christoph Schlaffner, christoph.schlaffner@childrens.harvard.edu

P13.08

The Antiviral Sirtuin 3 Bridges Protein Acetylation to Mitochondrial Integrity and Metabolism during Human Cytomegalovirus Infection

Xinlei Sheng (1), Ileana Cristea (1)

(1) Molecular Biology, Princeton University, Princeton, New Jersey, United States

Regulation of mitochondrial structure and function is a central component of infection with viruses, including human cytomegalovirus (HCMV), as a virus means to modulate cellular metabolism and immune responses. Here, we link the activity of the mitochondrial deacetylase SIRT3 and global mitochondrial acetylation status to host antiviral responses via regulation of both mitochondrial structural integrity and metabolism during HCMV infection. We establish that SIRT3 deacetylase activity is essential for suppressing virus production, and that SIRT3 maintains mitochondrial pH and membrane potential during infection. By defining the temporal dynamics of SIRT3-substrate interactions during infection, and overlaying acetylome and proteome information, we find altered SIRT3 associations with the mitochondrial fusion factor OPA1 and acetyl-CoA acyltransferase 2 (ACAA2), concomitant with changes in their acetylation levels. Using mutagenesis, microscopy, and virology assays, we determine OPA1 regulates mitochondrial morphology of infected cells and inhibits HCMV production. OPA1 acetylation status modulates these functions, as acetylation of K931, an established SIRT3-regulated site, dictates its capability to control mitochondrial morphology and restrict viral production. In addition, we establish K834 as a site regulated by SIRT3. Control of SIRT3 protein levels or enzymatic activity is sufficient for regulating mitochondrial filamentous structure. Lastly, we establish a virus restriction function for ACAA2, an enzyme regulating fatty acid beta-oxidation. Altogether, we highlight SIRT3 activity as a regulatory hub for mitochondrial acetylation, integrity, and metabolism during HCMV infection, and point to global acetylation as a reflection of mitochondrial health.

Contact: Xinlei Sheng, xsheng1990@gmail.com

P13.09

OxRAC Unveils Novel Role of LARP1 Along the Photosynthetic TOR Pathway

Amanda Smythers (1), Sarah Lowery (1), Saher Mubarek (1), Hailey Lewis (1), Danielle Slone (2), Derrick Kolling (2), Leslie Hicks (1)

(1) University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

(2) Marshall University, Huntington, WV, United States

The target of rapamycin (TOR) kinase is a master regulator implicated in nutritional sensing, protein translation, and autophagy, among others. In Chlamydomonas reinhardtii, a unicellular green alga, TOR inhibition results in increased triacylglycerol accumulation, suggesting that TOR or a downstream target is responsible for switching between protein biosynthesis and neutral lipids. Despite the desire to increase neutral lipid production for biofuel applications, the phototrophic TOR pathway is not well characterized. Although TOR is a Ser/Thr kinase that causes rapid intracellular phosphorylation in response to external stimuli, TOR inhibition also generates a large increase in intracellular ROS. Our previous work has revealed that TOR relies heavily on oxidative signaling, showing the crosstalk between phospho- and redox modifications are imperative to the metabolic response. Our previous work has also indicated that La-related protein 1 (LARP1) is a TOR target and involved in propagating the intracellular response following inhibition. In mammalian systems, LARP1 interacts directly with mTOR and regulates protein synthesis by binding and stabilizing mRNA before recruiting polysomes to initiate translation. LARP1 is essential in mammalian systems, and deficiencies lead to inhibited growth, defects, and increased apoptosis. However, LARP1 deficiency in C. reinhardtii does not cause the same phenotype, indicating a distinct uncharacterized role in phototrophic organisms. This study seeks to understand this role by using a LARP1 knockout mutant to compare the protein oxidation following inhibition with the parent strain. Using oxidized resin assisted capture to enrich for reversibly oxidized proteins, >4,000 oxidized cysteine thiols were observed and quantified. Simultaneously, the strains were assessed for physiological changes, including biochemical composition, cell size, and photosynthetic electron fluxes. Together, these data support LARP1 as a novel regulator of photosynthesis along the TOR pathway. The differential delineation of redox-controlled processes in LARP1 knockouts provides a framework for further TOR pathway characterization in phototrophic eukaryotes.

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

P13.10

The DNA Sensor cGAS is Decorated by Acetylation and Phosphorylation Modifications in the Context of Immune Signaling

Bokai Song (1), Todd Greco (1), Krystal Lum (1), Caroline Taber (1), Ileana Cristea (1)

(1) Molecular Biology Department, Princeton University, Princeton, NJ, USA

Cyclic GMP-AMP synthase (cGAS) is a pattern-recognition receptor of the mammalian innate immune system that is recognized as a main cytosolic sensor of pathogenic or damaged DNA. cGAS DNA binding initiates catalytic production of cyclic GMP-AMP, which activates the STING-TBK1-IRF3 signaling axis to induce cytokine expression. Post-translational modification (PTM) has started to be recognized as a critical component of cGAS regulation, yet the extent of these modifications remains unclear. Here, we report the identification and functional analysis of cGAS phosphorylations and acetylations in several cell types under basal and immune-stimulated conditions. cGAS was enriched by immunoaffinity purification from human primary fibroblasts prior to and after infection with herpes simplex virus type 1 (HSV-1), as well as from immune-stimulated STING-HEK293T cells. Six phosphorylations and eight acetylations were detected, of which eight PTMs were not previously documented. PTMs were validated by parallel reaction monitoring (PRM) mass spectrometry in fibroblasts, HEK293T cells, and THP-1 macrophage-like cells. To assess the functional relevance of each PTM, we generated a series of single-point cGAS mutations that prevented phosphorylation and acetylation or that mimicked the modification state. cGAS-dependent apoptotic and immune signaling activities were then assessed for each mutation. Our results show that acetyl-mimic mutations at Lys384 and Lys414 inhibit the ability of cGAS to induce apoptosis. In contrast, the Lys198 acetyl-mimic mutation increased cGAS-dependent interferon signaling when compared with the unmodified charge-mimic. Moreover, targeted PRM quantification showed that Lys198 acetylation is decreased upon infections with two herpesviruses—HSV-1 and human cytomegalovirus (HCMV), highlighting this residue as a regulatory point during virus infection. By determining the protein interactomes of the cGAS mutants following HSV-1 infection, several associations were identified that can contribute to the functional difference of Lys198 acetylated cGAS.

Contact: Bokai Song, bokais@princeton.edu

P13.11

Role of posttranslational modification, citrullination in polarization to proinflammatory and anti-inflammatory macrophages

Aneta Stachowicz (1), Justyna Fert-Bober (1)

(1) Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA

Introduction Macrophages play a pivotal role in the development of low-grade chronic inflammation present in many civilization diseases. They can be classified according to two main phenotypes: “proinflammatory” M1 macrophages, which are responsible for the clearance of pathogens and “anti-inflammatory” M2 macrophages, that are involved in resolution of inflammation and tissue repair. In recent years, citrullination, an irreversible enzymatic posttranslational modification by which arginine is converted to citrulline by the Protein Arginine Deiminase (PAD) family of enzymes has gained much attention due to its involvement in various physiologic and pathologic conditions. The loss of positively charged arginine can alter protein structure and function, affects protein-protein interactions, leads to proteolysis as well as acts as auto-antigen. Methods The aim of our study

was to investigate the role of PAD enzymes and its downstream effects, such as protein citrullination in polarization of macrophages to M1 and M2. We used human monocytic cell line THP-1 differentiated by PMA to macrophages. The cells were further differentiated to M1 and M2 macrophages by LPS and IL-4, respectively, and treated with PAD inhibitor – BB-Cl-amidine. Then we measured mRNA and protein expression of different PAD isoforms. To identify protein citrullination sites we applied mass spectrometry DIA method and hypercitrullinated library approach. Results PAD2 and PAD4 were the most abundant isoforms in THP-1 macrophages with PAD2 being downregulated in anti-inflammatory macrophages and PAD4 being upregulated in proinflammatory macrophages. Moreover, treatment with PAD inhibitor decreased expression of proinflammatory markers (IL-6, TNF-α) in M1 macrophages measured by ELISA and real-time PCR. Proteomic approach identified different patterns of protein citrullination in M1 and M2 macrophages. Conclusions Taken together our results indicate different expression of PAD enzymes as well as specific pattern of protein citrullination in proinflammatory and anti-inflammatory macrophages. The exact functional consequences of the revealed alterations require further investigations.

Contact: Aneta Stachowicz, aneta.stachowicz@uj.edu.pl

P13.12

Unraveling the Histone Code and Its Readers on Decondensed Heterochromatin

Stephanie Stransky (1), Jennifer Aguilan (2), Edward Nieves (1), Simone Sidoli (1)

(1) Biochemistry, Albert Einstein College of Medicine, New York, NY, US

(2) Pathology, Albert Einstein College of Medicine, New York, NY, US

Chromatin is composed of DNA and interacting proteins. Histones are the most abundant protein family on chromatin; their post-translational modifications (PTMs) model DNA readout. While histone PTMs are normally grouped in activating and silencing transcription, we previously demonstrated that a large percentage of histones is co-decorated by active and repressive marks (Sidoli et al. Anal Chem 2015, Epigenetics & Chromatin 2017). This hybrid chromatin is thought to be the result of cell stress since quiescent healthy cells have higher levels of compact heterochromatin (Sen et al. Molecular Cell 2019). To model physiological chromatin, we utilize an innovative 3D cell culture system to grow >1 mm large spheroids surviving for long periods of time (tested up to 80 days). This has been a milestone for our group, as canonical cell cultures are in an excessive proliferation state compared to primary solid tissues, preventing a proper modeling of DNA compaction. As well, our isotopic labeling techniques would have prohibitive costs if we utilized complex organisms. By utilizing sodium butyrate (HDAC inhibitor) we modeled an increase of accessible heterochromatin, demonstrated by the stable levels of silencing histone PTMs and an increase in abundance of hybrid codes such as H3K9me3K14ac. By using a peptide pull-down strategy, we identified the readers of hybrid histone codes. The most enriched proteins are involved in minimizing spurious RNA transcription, indicating that hybrid active-silencing histone PTM codes decorate DNA regions with anomalous transcription. Validation experiments are currently being performed to identify components of open and compact chromatin; these experiments include by separating chromatin domains by size after cross-linking, and mass spectrometry-based analysis of DNA modifications within those domains. This will demonstrate the role of open chromatin modified by silencing histone marks, and its impact on disease etiology.

Contact: Stephanie Stransky, stephanie.stransky@gmail.com

P13.13

Novel bifunctional material coupled to PASEF enables sensitive extracellular vesicles phosphoproteomic analysis of urine sample

Jie Sun (1), W. Andy Tao (2), Xiaoxian Du (3), xianming Liu (3)

(1) Southeast University, Nanjing, Jiangsu, China

(2) Purdue University, West Lafayette, Indiana, America

(3) Bruker Daltonics, Shanghai, Shanghai, China

Extracellular vesicles (EVs) have emerged as important carriers for intercellular communication and biological sources for diagnosis and therapeutics. Low efficiency in EV isolation from biofluids, however, severely restricts their downstream characterization and analysis. Here we introduced a novel strategy for EV isolation from urine for prostate cancer diagnosis using bi-functionalized magnetic beads through high affinity Ti(IV) ions and the insertion of a phospholipid derivative, 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, into EV membrane, synergistically. We demonstrated its efficient isolation of EVs from urine samples with low contamination, high recovery (>80%), and short separation time (within 1 h), resulting in the identification of 36,262 unique EV peptides corresponding to 3,302 unique proteins, and 3,233 unique phosphopeptides representing 1,098 unique phosphoproteins, using only 100 µL and 5 mL urine samples, respectively. Coupled with trapped ion mobility spectrometry (TIMS) and parallel accumulation-serial fragmentation (PASEF) for phosphosite specific resolution, quantitative phosphoproteomics of urine samples from prostate cancer patients and healthy individuals revealed 121 up-regulated phosphoproteins in cancer patients contrast to the healthy group. These particular advantages indicate that the novel bifunctional material enables sensitive EV phosphoproteomics analysis for non-invasive biomarker screening and cancer early diagnosis.

Contact: Jie Sun, 230189808@seu.edu.cn

P13.14

Analysis of Pancreatic Extracellular Matrix Proteins via Electrostatic Repulsion-Hydrophilic Interaction Chromatography Coupled with Mass Spectrometry

Dylan Nicholas Tabang (1), Yusi Cui (1), Daniel Tremmel (2), Megan Ford (3), Zihui Li (1), Sara Sackett (2), Jon Odorico (2), Lingjun Li (4)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States

(2) Department of Surgery, Division of Transplantation, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin, United States

(3) Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, Wisconsin, United States

(4) School of Pharmacy & Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, United States

The extracellular matrix (ECM) is a non-cellular network containing macromolecules which provide scaffolding for cells and participates in various biochemical processes, including cellular differentiation and maturation. Abnormalities in the ECM have been implicated in various diseases, including the desmoplastic reaction found in pancreatic cancer which leads to fibrous growth in the ECM. Studying the post-translational modifications (PTMs) of ECM proteins may provide valuable insight into the relationship of disease pathology and protein structure and function. Better understanding of the PTMs found in human pancreatic tissues may also lead to better outcomes in tissue engineering toward improving organ transplantation.
To analyze the PTM profile of the pancreas ECM, we compared pancreatic tissue from cadaveric donors (N=4) in native and decellularized states. These samples were analyzed using a novel spin-tip based enrichment method to identify PTMs of ECM proteins. Samples were trypsin digested and electrostatic repulsion-hydrophilic interaction chromatography was utilized to enrich and separate glycosylated and phosphorylated peptides. These peptides were then analyzed using liquid chromatography-tandem mass spectrometry. Nearly 10-fold more proteins were identified in the native tissues compared to the decellularized. Approximately 3000 glycopeptides and 80 phosphopeptides were identified in the native tissues. Among all eight samples, 138 core ECM and ECM-associated proteins were identified, with 46 of these proteins relatively quantified using precursor ion area. Among all identified glycoproteins, 195 distinct glycan compositions were identified, with a majority being sialylated or fucosylated. A deeper analysis into PTM heterogeneity on different ECM protein regions was conducted. These findings include high glycan microheterogeneity found on collagen alpha-1(V) chain, which plays roles in both insulin binding and pancreatic islet organogenesis.

Overall, this work demonstrates our ability to process and identify PTMs found in pancreatic tissue and its ECM and lays the groundwork for investigating pathologic differences in disease states.

Contact: Dylan Nicholas Tabang, tabang@wisc.edu

P13.15

Characterization of the Substrate Profile of Tyro3 using Phosphoproteomics

Naomi Widstrom (1), Laurie Parker (1)

(1) University of Minnesota, Biochemistry, Molecular Biology, and Biophysics, Minneapolis, MN, USA

Tyro3 is a poorly studied receptor tyrosine kinase with potential to be a target for cancer therapeutics. Tyro3 belongs to the TAM family of kinases, which are best known for immune signaling resolution and promoting wound healing in innate immune cells. However, the individual role of Tyro3 is not well understood. Tyro3 is overexpressed or ectopically expressed in a number of different cancer types, and is associated with cancer survival, migration, and chemoresistance. This suggests that Tyro3 may be a target for cancer therapeutics to block these key pathways. In order to study the functional role of Tyro3 in cancer as well as evaluate Tyro3 as a potential target for cancer therapeutics, we need to develop tools to study its activity. To do this, we first determined the substrate profile of Tyro3 using phosphoproteomics. In brief, we added active Tyro3 to a naturally derived peptide library, allowing the kinase to preferentially phosphorylate substrates and identified the sequences using mass spectrometry. Using this substrate profile, we designed synthetic peptide substrates to assess Tyro3 activity. We verified the top two scoring synthetic substrates were phosphorylated by Tyro3 using in vitro kinase assays. Additionally, we discovered these synthetic substrates were not phosphorylated by Tyro3 family members Axl and Mer. We aim to continue investigating the specificity of our Tyro3 synthetic substrates against other kinases using both in vitro kinase assays and cell-based assays. In addition, we determined the substrate profile of Axl and plan to determine the profile of Mer to begin examining the differences in substrate preferences between the TAM family members. These studies will help elucidate Tyro3's biological role, potentially providing insight into Tyro3 signaling in cancer, as well as provide a readout method for drug development and testing.

Contact: Naomi Widstrom, widst023@umn.edu

P13.17

Identifying HDAC3 as a delactylase using quantitative proteomics

Di Zhang (1), Jinjun Gao (1), Yingming Zhao (1)

(1) Ben May Department for Cancer Research, The University of Chicago, Chicago, IL, USA

We have recently reported lysine lactylation (Kla) as a new type of histone post-translational modification (Nature, 574:575-580, 2019), which can be stimulated by the Warburg effect-derived lactate. We show that histone lactylation is different from acetylation in terms of structural change, biochemical regulation and epigenetic functions. This discovery offers a new opportunity to understand function of the Warburg effect and glycolysis in normal physiology as well as in diseases. Nevertheless, the key regulatory elements for this pathway, “erasers”, “writers”, and “readers”, remain unknown, representing a major bottleneck to study its biology. In this abstract, we report the identification of HDAC3 as a delactylase. We used quantitative mass spectrometry to identify its target sites on core histones.

Contact: Di Zhang, dizhang2020@outlook.com

Protein Complexes and Interactomics

P14.01

Spatiotemporal virus-host interactions viewed through the lens of Inter-ViSTA

Katelyn Cook (1), Joel Federspiel (1), Michelle Kennedy (1), Samvida Venkatesh (1), Clayton Otter (1), William Hofstadter (1), Ileana Cristea (1)

(1) Princeton University, Princeton, NJ, USA

All viruses rely on organelles to infect, replicate, and spread between host cells. Accordingly, generating virus-host protein interactomes is a major goal of virology studies, yet effectively interpreting these datasets remains computationally challenging. To address this, we developed the web-accessible platform Inter-ViSTA (Interaction Visualization in Space and Time Analysis), which acts as a research accelerator tool by enabling users to quickly build animated protein interaction networks that automatically integrate information on protein abundances, functions, complexes, subcellular localizations, and changes across time. We used Inter-ViSTA to examine protein dynamics during infection with the critical human pathogen cytomegalovirus (HCMV), a betaherpesvirus, focusing on the viral anti-apoptotic protein pUL37x1. With quantitative mass spectrometry and Inter-ViSTA analysis, we defined the pUL37x1 protein interactome across the five-day cycle of HCMV infection. We uncover temporally-controlled associations that underlie the ability of pUL37x1 to remodel the mitochondria and peroxisome during infection, and use reciprocal isolations, microscopy, and genetic manipulations to functionally characterize these interactions. We find that pUL37x1 engages core mitochondrial structural proteins throughout infection, including the highly conserved MICOS complex, which we determine is required for infection. We also examine interactions with apoptotic and calcium-sensing proteins, revealing interactions that reflect the ability of pUL37x1 to simultaneously manipulate mitochondrial structure, antiviral signaling, and metabolism. Finally, we investigate associations through which pUL37x1 activates the peroxisome fission pathway to control peroxisome shape. Altogether, our work introduces Inter-ViSTA as a computational platform for analyzing and visualizing dynamic protein interactions, uncovering mechanisms that underly the viral ability to manipulate the balance of mitochondria and peroxisome fission-fusion for virus production.

Contact: Katelyn Cook, katelync@princeton.edu

P14.02

DNA-PK and IFI16 Coordinate an Intrinsic Immune Response to Herpesvirus Infection

Joshua L. Justice (1), Michelle A. Kennedy (1), Josiah E. Hutton (1), Dawei Liu (1), Bokai Song (1), Ileana M. Cristea (1)

(1) Department of Molecular Biology, Princeton University, Princeton, NJ, USA

Rapid and dynamically shifting protein-protein interactions underlie the capacity of cells to recognize viral infections and then ignite a downstream cascade leading to innate immune signaling. Simultaneously, viruses that share an ancient history of coevolution with their hosts have developed reciprocal methods to suppress or hijack immune response proteins—underscoring the biological complexity of host-pathogen interactions during immune responses. Accordingly, conventional approaches to study the protein-protein interactions that are downstream of innate immune sensing fail to capture the global scope and dynamic behavior of protein complexes as they are altered through space and time during infection. To overcome this barrier, we have applied Thermal Proteome Profiling Mass Spectrometry (TPP-MS) method to systemically characterize protein interactions that coordinate the innate immune response to Herpes Simplex Virus 1 (HSV-1) infection in primary human fibroblasts. Further, by advancing the power of TPP-MS to infer associations de novo and globally track these interactions we developed a time-resolved portrait of cellular and viral protein associations with IFI16, a nuclear DNA sensor that serves as a central platform for HSV-1 immune responses. Our TPP-MS analysis, along with high-resolution microscopy and molecular virology, linked IFI16 sensing of viral DNA in the nuclear periphery to the master DNA damage response (DDR) regulatory kinase, DNA-PK—which we show was necessary for the antiviral and inflammatory response to infection. Finally, phospho-peptide enrichment and MS analysis of DNA-PK substrates uncovered that IFI16 is targeted by the DDR kinase after both DNA damage and viral infection. Functional analysis of this DDR-dependent IFI16 phosphorylation revealed the specific modified site required for IFI16-driven cytokine responses. Altogether, our study represents the first cell-wide characterization of the global dynamics of PPIs during HSV-1 infection and uncovers a missing link in the immune signaling pathway that places IFI16 and DNA-PK central to innate immunity against herpesvirus.

Contact: Joshua Justice, jjustice@princeton.edu

P14.03

A Study of the NPC1 Interactome Using Mass Spectrometry

Thu Nguyen (1), Daisy Edmison (2), Swetha Gowrishankar (2), Stephanie Cologna (1)

(1) Department of Chemistry, University of Illinois at Chicago, Chicago, IL, USA

(2) Department of Anatomy and Cell Biology, University of Illinois at Chicago, Chicago, IL, USA

Niemann-Pick disease, type C (NPC1) is an autosomal recessive, neurodegenerative, lysosomal storage disorder, caused by mutations in the NPC1 gene. As a result, lysosomal accumulation of unesterified cholesterol and glycosphingolipids are observed, which leads to a myriad of defects including progressive neurodegeneration of cerebellar Purkinje neurons. Currently, there is no FDA-approved therapy for NPC1. While much effort has been focused on understanding the disease mechanisms, much less has been revealed about the NPC1 protein itself beyond a role in cholesterol transport. In the current study, we aim to investigate protein-protein interactions of NPC1 to provide a deeper understanding of this protein's function.

An initial study was done using the Balb/c Npc1^{nih} (Npc1 null) mouse model. Cortical tissues of 9-week NPC1 null and control mice were collected, lysed, and protein quantification was obtained. Immunoprecipitation was performed using the Thermo Co-IP kit and the Abcam NPC1 primary antibody (ab134113). Eluents were digested using the S-Trap method, followed by nanoflow liquid chromatography-tandem mass spectrometry (LC-MS/MS) detection with an Agilent 1260 LC coupled with a Thermo QExactive mass spectrometer. Proteins were identified using Proteome Discoverer equipped with the SEQUEST search engine. Identifications were filtered to include proteins with >2 unique peptides, and presence in all 4 biological replicates of only wild-type animals. We observed 39 novel potential interactors of NPC1 in cortical tissues, including proteins that had been known to be essential to neuron function such as synaptophysin, microtubule-associated protein 6, and disks large homolog 2. Current efforts are focused on using immunostaining of cortical neurons to examine the co-localization of NPC1 and potential interactors. In addition, due to the high number of cytoskeletal proteins as possible interactors, we hypothesize NPC1 may interact with the lysosomal transport protein scaffold. Thus, we aim to study the morphology and distribution of lysosomes in neurons lacking NPC1.

Contact: Thu Nguyen, tnguy214@uic.edu

P14.04

A comprehensive map of protein complexes in mature human erythrocytes

Wisath Sae-Lee (1), Eric Verbeke (1), Claire McWhite (2), Ophelia Papoulas (1), Pierre Havugimana (3), Andrew Emili (3), Edward Marcotte (1)

(1) Molecular Biosciences, University of Texas at Austin, Austin, TX, USA

(2) Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA

(3) Center for Network Systems Biology, Boston University School of Medicine, Boston, MA, USA

Erythrocytes (red blood cells; RBCs) are the simplest primary human cells, lacking nuclei and all major organelles. Despite their simplicity, RBCs dynamically change cellular morphology and physiology throughout their journey in the body. These cellular dynamics are mediated by protein assemblies whose complete picture in RBCs is still unknown. While hemoglobin accounts for 98% of expressed RBC proteins, the full proteome includes >1,000 distinct proteins in which the roles of many remain elusive. In this study, we first identified a comprehensive RBC proteome of 1,202 proteins (1% FDR) using machine learning from quantitative mass spectrometry and RNA-seq on RBCs and other blood cell types. We then determined the stable protein complexes in mature RBCs, based on mass spectrometry of 1944 native biochemical fractions of hemoglobin-depleted hemolysate and detergent solubilized membrane protein complexes. Our data reveal an RBC interactome dominated by protein homeostasis, redox biology, cytoskeletal dynamics, and carbon-metabolism. We confirmed many protein complexes through chemical cross-linking and electron microscopy. As the first near-complete interactome for any primary human cell type, our map of RBC protein complexes provides a better understanding of the unique constraints of RBC function and serves as a comprehensive resource for future research.

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

P14.05

Characterization of Native Protein Complexes via In Situ Chemical Cross-linking and Mass Spectrometry (CHEM-MS)

Yuefan Wang (1), Yingwei Hu (1), Naseruddin Höti (1), Hui Zhang (1)

(1) pathology, Baltimore, MD, USA

Every organism has limited genes and identical genome throughout the whole life. However, proteins as gene products carrying out dynamic and diverse biological functions depend on protein complexes, which play central roles in physiological and pathological processes. The protein complexes dynamically assemble and dissemble based on expression of different proteins and could provide mechanistic insights into organization of biological system. Characterization of protein complexes is critical to understanding the biological and clinical significances of protein functions such as cancer and infectious disease related pathways and has become one of the major efforts of functional proteomics. Accepting the fact that a large-scale characterization of protein complexes remains difficult, protein co-aggregation or co-elution coupled with size exclusion chromatography (SEC) and mass spectrometry strategies are the comprehensive, antibody or gene editing free technologies, such as cellular thermal shift assay (CETSA) and label free co-elution-based SEC-MS method are used to characterize global protein interactome. These approaches have been successfully used in cell or tissue lysates of protein extracts. It has been challenging to determine which of these protein complexes are formed in situ in cells prior to cells or tissues are lysed. Herein, we describe an approach to preserve in situ protein complexes using covalent chemical crosslinks followed by size exclusion chromatography (SEC) and mass spectrometry (CHEM-MS). More importantly, this method enables the characterization of in situ protein complexes in any kind of sample, which enables to apply in clinical research.

Contact: Yuefan Wang, ywang298@jhmi.edu

Proteomics in Immunology and Infectious Disease (COVID-19)

P15.01

Protein Responses in Children with Multisystem Inflammatory Syndrome developed after SARS-CoV-2 infection

Rebecca Porrit (1), Aleksandra Binek (1), Angela McArdle (1), Alejandro Rivas (1), Blandine Chazarin (1), Danica-Mae Manalo (1), Vidya Venkatraman (1), Magali Noval Rivas (1), Moshe Arditi* (1), Jennifer Van Eyk* (1)

(1) Cedars-Sinai Medical Center, Los Angeles, CA, USA

Background

Within few months of the onset of SARS-CoV-2 pandemic, novel multisystem inflammatory syndrome in children (MIS-C) emerged. In this rare but very severe hyperinflammatory syndrome children present with high fevers and number of symptoms such as pink eye, rash, enlarged lymph nodes, coronary artery dilation and in most severe cases cardiovascular shock and multiple organ failure.

Methods

We carried out proteomics on native and top abundant protein depleted serum obtained from MIS-C patients (MIS-C, n=25), pediatric healthy controls (HC, n=20) and pediatric patients diagnosed with Kawasaki disease (KD, n=7). After Top14 abundant protein depletion (High Select) samples were processed in the same manner as the native serum, using Beckman-Coulter i7 automated workstation protocol and measured by data independent acquisition (DIA) unbiased discovery workflow on Thermo Fisher Exploris 480 mass spectrometer (MS) coupled to Ultimate 3000 liquid chromatography (LC) system using a 60 min long gradient time. We processed DIA raw files using our in-house built pipeline designed on Optra systems, which integrates OpenSwath search engine workflow and mapDIA statistical analysis.

Results

Combined datasets yielded 680 serum proteins and were subjected to unsupervised statistical analysis. Data showed an intriguing segregation of severe MIS-C patients into two subpopulations (i. MIS-C severe cases only, and ii. MISC severe, mild and KD), where first was defined by a cluster of proteins involved in humoral response, while the mixed population showed signatures mainly involved DNA damage, hypoxia, ROS signaling and broad spectrum of innate immunity.

Conclusion

In conclusion, our proteomic study allowed us to observe an intriguing segregation of MIS-C patients. One population of severe-only cases was clearly defined by immunoglobulin components and the other by complex cellular signaling responses to inflammatory stress.

Contact: Aleksandra Binek, bineka@cshs.org

P15.02

A longitudinal proteomic and diagnostic map characterizes COVID-19 progression and predicts outcome

Vadim Demichev (1)

(1) The Francis Crick Institute, London, UK

COVID-19 is highly variable in its clinical presentation, ranging from asymptomatic infection to severe organ damage and death. There is an urgent need for predictive markers that can guide clinical decision-making, monitor the effects of therapies and point to novel therapeutic targets. Addressing this need, we characterize the time-dependent progression of COVID-19 through different stages of the disease by measuring 86 accredited diagnostic parameters and plasma proteomes at 687 sampling points, in a cohort of 139 patients during hospitalization.

We report that the time-resolved patient molecular phenotypes reflect an initial spike in the systemic inflammatory response, which is then gradually alleviated and followed by a protein signature indicative of tissue repair, metabolic reconstitution and immunomodulation, in both mildly and severely ill patients. In contrast, many organ function markers do not show consistent alleviation and exhibit high heterogeneity between patient trajectories.

As age is the most significant risk factor in COVID-19, we further looked for plasma protein markers characteristic of age-specific response to the disease. Among the markers identified, 20 are not associated with age in the general population and reveal increased inflammation and lipoprotein dysregulation in older COVID-19 patients.

We show that the early host response is predictive for the disease trajectory and gives rise to proteomic and diagnostic marker signatures that classify the need for supplemental oxygen therapy and mechanical ventilation (168 markers), and that predict the time to recovery of mildly ill patients (40 markers). In severely ill patients, the molecular phenotype of the early host response predicts survival, in two independent cohorts and weeks before outcome.

Contact: Vadim Demichev, vadim.demichev@gmail.com

P15.03

Analyzing the Proteomes Associated with Viral Genomes to Identify Host Restriction Factors

Joseph Dybas (1), Krystal Lum (1), Eui Tae Kim (2), Katarzyna Kulej (1), Emigdio Reyes (1), Benjamin Garcia (3), Matthew Weitzman (1)

(1) Division of Protective Immunity and Division of Cancer Pathobiology, The Children's Hospital of Philadelphia, Philadelphia, PA, United States

(2) Department of Microbiology and Immunology, Jeju National University School of Medicine, Jeju, Republic of Korea

(3) Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, United States

The host intrinsic immune response to viral infection comprises antiviral proteins, known as restriction factors, that target viral genomes and inhibit infection. On the other hand, viruses have evolved mechanisms to counteract the host responses, including hijacking the cellular ubiquitination machinery to redirect protein modifications by ubiquitin. Viral-mediated ubiquitination and degradation of host restriction factors is an important component of virus-host interactions. However, the repertoire of restriction factors and mechanisms of viral restriction are not well known. Herpes simplex virus 1 (HSV-1) is used as a model virus to study restriction factors toward DNA viruses. HSV-1 encodes a viral E3 ubiquitin ligase, ICP0, which ubiquitinates and degrades antiviral proteins. We used Identification of Proteins on Nascent DNA (iPOND) to quantify the proteomes associated with HSV-1 genomes during infection. Comparing the iPOND proteomes for HSV-1 wildtype or ICP0-mutant virus, for which the ICP0 E3 ligase activity is abrogated, identifies host proteins that are recruited to viral genomes in the absence of ICP0-mediated ubiquitination and degradation. We hypothesize that the proteins enriched on ICP0-mutant viral genomes are host restriction factors that are counteracted by ICP0 during wildtype infection. A protein domain prediction analysis revealed numerous BTB/POZ and zinc finger domain proteins enriched on ICP0-mutant genomes. These proteins share a common potential function of transcription repression, suggesting that they may be restriction factors of HSV-1 that inhibit viral transcription. We used experimental approaches to confirm that the identified proteins are substrates of ICP0 and are characterizing their impact on viral transcription, genome replication, and progeny formation. Our innovative implementation of comparative proteomics identifies host proteins that are associated with viral genomes in the absence of viral-mediated ubiquitination. Our approach reveals a class of transcriptional repressor proteins that are likely novel restriction factors against DNA viruses and are targeted by viral countermeasures.

Contact: Joseph Dybas, dybasj@email.chop.edu

P15.04

Integrative Proteo-Transcriptomic Analysis Identifies the Immune Response in a Mouse Model of Congenital Zika Virus Microcephaly

Kimino Fujimura (1), Amanda Guise (1), Tojo Nakayama (2), Anais Meziani (1), Mukesh Kumar (1), Christoph Schlaffner (1), Long Cheng (1), Dylan Vaughan (2), Andrew Kodani (3), Simon Van Haren (4), Kenneth Parker (5), Ofer Levy (4), Ann Durbin (6), Irene Bosch (6), Lee Gehrke (6), Hanno Steen (5), Ganeshwaran Mochida (2), Judith Steen (1)

(1) F.M. Kirby Neurobiology Center, Boston Children's Hospital, Boston, MA, United States

(2) Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, United States

(3) Center for Pediatric Neurological Disease Research and Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN, United States

(4) Precision Vaccines Program, Division of Infectious Diseases, Boston Children's Hospital, Boston, MA, United States

(5) Department of Pathology, Boston Children's Hospital, Boston, MA, United States

(6) Department of Microbiology, Harvard Medical School, Boston, MA, United States

Congenital infection, occurring in utero or during delivery, is a significant cause of structural abnormalities of the brain in fetuses and neonates. The acute crisis of the 2015-2016 zika virus (ZIKV) epidemic in the Americas led to identifying a range of developmental abnormalities in the affected children's central nervous system, including microcephaly. To identify molecules and pathways involved in congenital ZIKV infection in a translationally relevant model, we conducted a tandem mass tag (TMT) mass spectrometry-based quantitative proteomics and paired RNA-sequencing analyses to study an immunocompetent congenital ZIKV infection mouse model. The ZIKV-infected embryos demonstrated smaller brain size and decreased cerebral cortical thickness as compared to the mock-infected embryos. Comprehensive developmental proteome coverage of the brain was achieved, resulting in the identification of >8,000 proteins (5% false discovery rate) and accurate quantitation of >6,800 proteins across subcellular compartments. Reflecting the active infection, we detected ZIKV polyprotein at >40% total sequence coverage. While we observed neurodevelopmental pathways such as NEUROD2-EOMES/TBR2 transcription factor cascades being dysregulated, the most robust response in the virus-infected brain was the induction of innate immunity. Our data showed the activation of immunoproteasome and MHC-I processing and presentation machinery associated with T cell and B cell infiltration and neuronal death during brain development. Together, the proteo-transcriptomic analysis describes the underlying pathology in the ZIKV-infected developing brain that informs fresh approaches to mitigate ZIKV-associated microcephaly. This combination of a congenital infection model and proteo-transcriptomic analysis provides insight into both the peripheral and tissue-specific immune mechanisms perturbed during infection, providing a resource and foundation for therapeutic interventions for many devastating congenital infection syndromes beyond ZIKV.

Contact: Kimino Fujimura, Kimino.Fujimura@childrens.harvard.edu

P15.05

Discovery of RSV-induced BRD4 Protein Interactions using native immunoprecipitation and PASEF mass spectrometry.

Morgan Mann (1), David Roberts (2), Ying Ge (2), Allan Brasier (1)

(1) UW-Madison, Department of Medicine, Madison, WI, USA

(2) UW-Madison, Department of Cell and Regenerative Biology, Department of Chemistry, Madison, WI, USA

Respiratory Syncytial Virus (RSV) is a ubiquitous airway virus that causes severe inflammation and airway pathology in children and the elderly. RSV infects the epithelial cells of the upper and lower respiratory tract, resulting in toll-like receptor and NF-kB mediated innate inflammation. In the process, the NF-kB subunit RelA/p65 translocates into the nucleus and interacts with the epigenetic scaffold BRD4. The resulting complex traffics to NF-kB early-intermediate genes to trigger the coordinated expression of pro-inflammatory chemokines and interferons to signal nearby immune cells and form an early antiviral defense. This program of gene expression is tightly regulated and stimulus specific. However, in severe infections, hyperstimulation and chronic inflammation can result in tissue damage and airway obstruction, often requiring hospitalization.

While the RelA/BRD4 complex is known to activate gene expression through transcriptional elongation, it remains unclear how the protein complex facilitates the high level of gene coordination exhibited during innate induction. One possible explanation is that BRD4 recruits additional proteins that impart gene-level specificity, such as transcription factors with specific DNA binding domains, or chromatin remodelers that can increase chromatin accessibility at a subset of gene promoters. To investigate this hypothesis, we utilized native immunoprecipitation to purify the BRD4 protein complex from RSV-treated human small airway epithelial cells, and analyzed the samples via shotgun proteomics. Samples were analyzed using a nanoElute LC system coupled to timsTOF Pro mass spectrometer (Bruker Daltronik) utilizing the PASEF acquisition mode. MaxQuant was used to process the data files and perform label-free quantitation.

At the conclusion of this study, we expect to observe a relative enrichment of transcription factors, such as AP1, RelA, and Smad3, in the virus-stimulated protein complex. Such interactors may modulate the expression of NF-kB early-intermediate inflammatory genes through a variety of mechanisms, and may serve as useful targets for clinical intervention.

Contact: Morgan Mann, mwmann@wisc.edu

P15.06

A rigorous evaluation of optimal peptide targets for MS-based clinical diagnostics of Coronavirus Disease 2019 (COVID-19).

Andrew Rajczewski (1), Subina Mehta (1), Dinh Duy An Nguyen (1), Björn Grüning (2), James Johnson (1), Thomas McGowan (1), Timothy Griffin (1), Pratik Jagtap (1)

(1) University of Minnesota, Minneapolis, Minnesota, USA

(2) University of Freiburg, Freiburg im Breisgau, Baden-Württemberg, Germany

The COVID-19 pandemic has wrought tremendous harm, with almost 65 million confirmed cases and approaching 1.5 million deaths since the disease was first observed in late 2019. To supplement the RNA-based diagnostic procedures that are currently used, direct detection of SARS-CoV-2 virus peptides with targeted proteomics is a promising alternative for virus detection in clinical and environmental samples. However, not all SARS-CoV-2 peptides would serve as adequate targets for targeted LC-MS analysis due to potentially poor ionization, similarity to host protein sequences, etc. Therefore, a rigorous evaluation to assess the most reliable peptide targets is an important step towards developing clinical tests. To this end, we developed bioinformatics workflows in the Galaxy platform to identify all SARS-CoV-2 peptides observable in a collection of publicly available SARS-CoV-2 datasets and then validate those peptides using the PepQuery search engine. PepQuery rigorously assesses putative MS/MS matches to SARS-CoV-2 peptides by testing the validity of these matches against peptide sequences from the human host or other coronaviruses. With these methodologies, we identified 639 peptides across three clinical samples and three in vitro samples that putatively belong to SARS-CoV-2, of which 87 were validated using our stringent bioinformatics approach that also assesses peptide-spectral match quality. Of these, 11 peptides were detected consistently across all six samples. Taxonomic analysis using MetaTryp 2.0 showed these peptides belonged to the Coronaviridae family, and subsequent sequence alignment with BLAST-P showed 8 of these peptides were distinct to SARS-CoV-2 over other coronaviruses. We contend that these peptides, which align to the nucleocapsid and membrane proteins of the viral particle, would serve as the optimal targets for mass spectrometry detection of SARS-CoV-2. In addition, our optimized workflows (available via https://covid19.galaxyproject.org/proteomics/) can be used to detect presence of the SARS-CoV-2 virus in clinical and environmental samples.

Contact: Andrew Rajczewski, rajcz001@umn.edu

P15.07

A quantitative proteomics studies delineate the mechanistic pathways associated with COVID-19 comorbidity with kidney disease

Alakesh Bera (1), Maura Watson (2), Madhan Subramanian (3), John Karaian (3), Michael Eklund (3), Robert Nee (2), Ofer Eidelman (3), Harvey Pollard (3), Rahul Jindal (4), Meera Srivastava (1)

(1) Department of Anatomy, Physiology and Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

(2) Department of Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

(3) Department of Anatomy, Physiology and Genetics, Bethesda, MD, USA

(4) USU-Walter Reed Department of Surgery, Bethesda, MD, USA

COVID-19 is a global issue, with over 62 million cases in 213 countries and territories on November 30th, 2020. Although this virus infects all groups, data indicates that the risk for severe disease and death is much higher in older population and having some underlying conditions. Data also suggest that more men than women are dying with COVID-19. Researchers found two host genes which are essential for severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2). ACE2 is a SARS-CoV-2 receptor, acting as a landing platform for the virus; whereas the serine protease, TMPRSS2, primes the virus for cell entry through cleavage of the viral spike protein. Several studies indicating that many organs including kidney and lung are targeted by this virus, which contributes significantly to COVID-19 morbidity. We hypothesized that the organ specific differential expression of ACE2 and TMPRSS2 and, induced level of cytokines expression are associated with many complications started with initial SARS-CoV-2 infection. The goal of this current study is to determine the underlying mechanism associated with COVID-19 severity and comorbidity related to kidney diseases. Preliminary data indicated that an increase level of serum TMPRSS2 present in male serum. Besides, tissue level RNA and protein expression data from different organs indicating that ACE2 and TMPRSS2 are high in intestine, stomach, liver, kidney and prostate. We are currently conducting a study to identify the biomarkers in the serum from different groups of chronic kidney disease (CKD) patients and kidney transplanted patients vs. healthy individuals. We analyzed the inflammatory markers in patients' serum samples by high throughput antibody arrays. Data indicating that AKT and SYK pathways probably play a critical role in induced cytokine release by SARS-CoV-2 infection. Further cellular and animal model experiments were currently pursuing in our lab to uncover the mechanism of CKD and COVID-19 related comorbidity.

Contact: Meera Srivastava, meera.srivastava@usuhs.edu

P15.08

Automated MHC-Associated Peptide Enrichment for Immunopeptidomics Analysis

Samuel Pollock (1), Shuai Wu (2), Jerry Han (2), Steve Murphy (2)

(1) Genentech, South San Francisco, California, USA

(2) Agilent Technologies, Santa Clara, California, USA

When there is a viral infection, cytotoxic T lymphocytes (CTLs) are activated after recognizing host cells presenting viral antigens via the MHC-I display system. These antigens are generated after viral proteins are degraded into peptides in the cytoplasm and transported to the endoplasmic reticulum (ER), where they are loaded into class I major histocompatibility complexes (MHC-I), which then migrate to the cell surface. These MHC-I complexes are of interest not only in viral defense, but also in cancer immunotherapy, as CTLs can also recognize peptides derived from endogenous, mutated proteins. Therefore, characterization of MHC-I presented peptides, or immunopeptidome, is of great interest in the design of vaccines and immunotherapies against cancer and infectious diseases.

Immunopeptidomics is generally considered more challenging in part due to the fact that most MHC-associated peptides are extremely low in abundance compared to other cellular peptides. One way to address this challenge is to immunoprecipitate the MHC-I complexes from large amounts of cell lysate using an anti-MHC-I antibody coupled to a solid phase. The MHC-associated peptides are then isolated from the complex and identified by mass spectrometry.

High-quality sample preparation is critical for obtaining good peptide signal and highly reproducible results. This workflow used the AssayMAP Bravo platform for automated immunoaffinity purification and peptide cleanup. In addition to high-quality sample preparation, a very sensitive and reproducible LC/MS system is required. For this study, an Infinity UHPLC Nanodapter converted standard LC flow to nanoflow. The nanoflow LC system was then coupled with a 6550 iFunnel Q-TOF LC/MS using data-dependent acquisition (DDA) for peptide analysis. The combination of high-quality sample preparation and a sensitive LC/MS system provide a reproducible and scalable solution to meet the challenges of studying MHC-I peptidomics.

Contact: Shuai Wu, shuai.wu@agilent.com

Quantitative Proteomics and Systems Biology

P16.01 Comprehensive Stoichiometric Characterization of the Thiol Redox Proteome

Tong Zhang, Nicholas Day, Matthew Gaffrey, Rui Zhao, Garett Grant, Tujin Shi, Thomas Fillmore, George Rodney, My Helms, Wei-Jun Qian

P16.02 Multiplexed Quantification of Insulin and C-peptide Without the Use of Antibodies

North Foulon, Elisha Goonatilleke, Michael MacCoss, Michelle Emrick, Andrew Hoofnagle

P16.03 Proteomic Evaluation of the Dynamic Protein Turnover in Human iPSC-derived Neurons

Ashley Frankenfield, Saadia Hasan, Michael Ward, Ling Hao

P16.04 MSstatsPTM: an R/Bioconductor software for detecting quantitative changes in post-translational modifications

Devon Kohler, Tsung-Heng Tsai, Ting Huang, Erik Verschueren, Trent Hinkle, Meena Choi, Olga Vitek

P16.05 Proteome-wide and Matrisome-specific Alterations during Human Pancreas Development and Maturation Revealed by 12-plex Isobaric DiLeu Tags

Zihui Li, Daniel Tremmel, Fengfei Ma, Qinying Yu, Min Ma, Daniel Delafield, Yatao Shi, Bin Wang, Samantha Mitchell, Austin Feeney, Vansh Jain, Sara Sackett, Jon Odorico, Lingjun Li

P16.06 Extending Capabilities of Real-Time Database Search on the Orbitrap Eclipse Tribrid Mass Spectrometer for Multiplexed Proteomics

Yang Liu, William D. Barshop, Amirmansoor Hakimi, Jesse D. Canterbury, Romain Huguet, Rosa Viner

P16.07 DIA Performance in Discovery and Quantitation Analysis on New Orbitrap Exploris Mass Spectrometers

Yang Liu, Khatereh Motamedchaboki, Joshua Nicklay, Sega Ndiaye, Aman Makaju, Amarjeet Flora, Ryan Bomgarden, Daniel Lopez-Ferrer

P16.08 Label-Free Proteomics Performance with New Orbitrap Exploris 480 mass spectrometer with Single-Cell Sensitivity

Daniel Lopez-Ferrer, Aaron S. Gajadhar, Aman Makaju, Aaron M. Robitaille, Julia Kraegenbring, Tabiwang N. Arrey, Michelle Dubuke, Emily Chen, Alexander Harder

P16.09 Robust Label Free Proteomics Performance with New Orbitrap Exploris 240 with Improved Separation

Daniel Lopez-Ferrer, Yang Liu, Aaron S. Gajadhar, Julia Kraegenbring, Tabiwang N. Arrey, Michelle Dubuke, Emily Chen, Alexander Harder

P16.10 Investigating bortezomib-induced protein degradation with pulseSILAC-DIA

Lindsay Pino, Josue Baeza, Richard Lauman, Birgit Schilling, Benjamin Garcia

P16.11 Comprehensive Proteomic Quantification of in vivo Bladder Stone Progression in a Cystinuric Mouse Model using Data-Independent Acquisitions

Jacob Rose, Tiffany Zee, Nathan Basisty, Pankaj Kapahi, Marshall Stoller, Birgit Schilling

P16.12 Draft Quantitative Proteomic Atlas of Human Body and Common Carcinomas

Liang Yue, Meng Luo, Wenhao Jiang, Ning Fan, Xiaolu Zhan, Fangfei Zhang, Fengchao Yu, Guoci Teo, Alexey Nesvizhskii, Ben Collins, Ruedi Aebersold, Fei Xu, Tong Liu, Yan Li, Tiannan Guo

P16.01

Comprehensive Stoichiometric Characterization of the Thiol Redox Proteome

Tong Zhang (1), Nicholas Day (1), Matthew Gaffrey (1), Rui Zhao (1), Garett Grant (2), Tujin Shi (1), Thomas Fillmore (1), George Rodney (3), My Helms (2), Wei-Jun Qian (1)

(1) Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, United States

(2) Department of Internal Medicine, University of Utah, Salt Lake City, UT, United States

(3) Department of Molecular Physiology & Biophysics, Baylor College of Medicine, Houston, TX, United States

Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are utilized as signaling molecules that form redox-dependent post-translational modifications of Cysteine thiol groups. Quantitative measurement of oxidative protein thiol modifications is paramount to understanding the state of the redox environment in a cell and helps to identify what Cysteine sites and proteins are susceptible to these modifications. Obtaining deep coverage of the thiol redox proteome continues to be a challenge as current methods can identify at most, several thousand Cysteine sites in a given tissue. In comparison, current phosphoproteomic approaches can recover more than 20,000 phosphorylated sites. Herein, we present a deep profiling and site-occupancy quantification of protein thiol oxidation (DOPO) strategy that integrates resin-assisted capture, multiplexed isobaric tagging, and 2D-LC-MS/MS to obtain an unprecedented coverage of more than 20,000 Cysteine sites for a tissue. We investigated the redox proteome in multiple tissue types, including mouse skeletal muscle that is challenging to profile due to presence of overabundant proteins. This approach enables site occupancy analysis of thiol oxidation, or stoichiometry, as well as profiling of transcription factors that are typically low in abundance. We present this deep redox profiling approach as a way to quantitatively evaluate redox-based post-translational modifications at the global and site-specific level.

Contact: Nicholas Day, nicholas.day@pnnl.gov

P16.02

Multiplexed Quantification of Insulin and C-peptide Without the Use of Antibodies

North Foulon (1), Elisha Goonatilleke (1), Michael MacCoss (2), Michelle Emrick (1), Andrew Hoofnagle (1)

(1) Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

(2) Department of Genome Sciences, University of Washington, Seattle, WA, USA

Background: Quantification of both C-peptide and insulin can be used to measure residual beta-cell function in diabetes, assess insulin resistance, and clinically determine the pathophysiological mechanisms of hypoglycemia. Previously developed LC-MS/MS methods to quantify serum concentrations of C-peptide and/or insulin have relied on intact C-peptide or the reduced B-chain, respectively, which ionize poorly. As a result, methods have leveraged immunoaffinity enrichment, which is plagued with interferences, or two-dimensional chromatography, which limits instrument throughput. We aimed to develop a workflow that uses proteolysis to enhance the sensitivity of traditional LC-MS/MS. **Methods:** Glu-C was selected as the protease, because C-peptide contains no trypsin cleavage sites and because it liberates peptides specific to insulin vs. proinsulin. After protein precipitation and solid phase extraction, lower molecular weight polypeptides were reduced, alkylated, and proteolyzed. Two peptide fragments were monitored for each protein (Acquity ULPC-Xevo TQ-S, Waters). **Results:** Glu-C was a reliable proteolytic enzyme with monotonic digestion kinetics. The assay was linear across clinically relevant concentrations for each analyte. The lower limits of quantification were similar to available clinical immunoassays and the precision of each analyte was appropriate for clinical care and clinical research. The C-peptide assay was fully validated for clinical care (according to CLSI C62-A guidelines) and was launched early in 2020 within the University of Washington Clinical Reference Laboratories. Long-term imprecision has been 10% (at 0.6 ng/mL). **Conclusion and future directions:** This is a robust method using Glu-C for the quantification of C-peptide and insulin in human serum. Without the need for an antibody to enrich the analytes, the method builds a foundation for further multiplexing of analytes relevant to the study of diabetes and obesity. The methods will be cross-validated within the Targeted Mass Spectrometry for Diabetes and Obesity Research (TaMaDOR) Consortium, which is coordinated by NIDDK and currently includes three institutions.

Contact: North Foulon, nfoulon@uw.edu

P16.03

Proteomic Evaluation of the Dynamic Protein Turnover in Human iPSC-derived Neurons

Ashley Frankenfield (1), Saadia Hasan (2), Michael Ward (2), Ling Hao (1)

(1) Department of Chemistry, The George Washington University, Washington, DC, USA

(2) NINDS/NIH, Bethesda, MD, USA

Neurodegeneration involves the progressive accumulation of misfolded and aggregated proteins within the brain which can be caused by dysfunctional protein quality control pathways. Cellular proteins are primarily degraded through the ubiquitin-proteasome system (UPS) or the autophagy-lysosome system. The UPS pathway tags proteins for degradation with ubiquitin, while the autophagy-lysosome pathway uses the acidic environment of the lysosome to digest and recycle macromolecules. While these pathways have been extensively studied, which proteins are degraded by a specific pathway are not clear, particularly for proteins in neurons. In this study, we coupled dynamic Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC) of human induced pluripotent stem cell (iPSC)-derived neurons with mass spectrometry (MS)-based proteomics to study protein turnover and determine the protein half-lives in iPSC-derived neurons. We inhibited the ubiquitin-proteasome system with epoxomicin treatment and the autophagy pathway with bafilomycin A1 treatment, respectively, to determine which neuronal proteins are mainly regulated by which degradation pathway. Dynamic SILAC was performed by culturing the iPSC-neurons with a “light” media. After neuronal maturation, the media was switched to a “heavy” (^{CN}) lysine-containing media. The heavy-to-light protein ratios were used to determine the protein turnover rates at both multiple time points and single time points enabling an accurate measurement of protein half-lives in neurons. Our study has shown that the median half-life for protein turnover within neurons is 4.8 days, which is significantly longer than mitotic cell types. The changes in protein turnover after drug treatment will enable us to identify neuronal proteins that are reliant upon a specific pathway for degradation and can help us better understand the mechanisms of the accumulation of misfolded proteins in neurodegenerative diseases.

Contact: Ashley Frankenfield, afrankenfield@gwu.edu

P16.04

MSstatsPTM: an R/Bioconductor software for detecting quantitative changes in post-translational modifications

Devon Kohler (1), Tsung-Heng Tsai (2), Ting Huang (1), Erik Verschueren (3), Trent Hinkle (4), Meena Choi (4), Olga Vitek (1)

(1) Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States

(2) Kent State University, Kent, OH, United States

(3) Galapagos, Mechelen, Antwerp, Belgium

(4) Genentech, South San Francisco, CA, United States

The scientific community widely utilizes mass spectrometry (MS)-based proteomics in biological and clinical research. One area of particular interest in MS-based proteomics is the impact of post translational modifications (PTMs), such as phosphorylation. Experiments targeting PTMs must consider additional complexities during evaluation over those solely targeting unmodified peptides. One complexity of particular importance is the impact of changes in the abundance of unmodified peptides. Changes in abundance of unmodified peptides can introduce bias when determining differentially abundant PTMs, masking crucial experimental conclusions. We argue that experiments targeting changes in PTM abundance require a workflow which is both reproducible and comprehensive in order to properlyassess the results.

To allow researchers to fully evaluate experiments related to post translational modifications, we developed a free and open-source R package MSstatsPTM (available in Bioconductor). MSstatsPTM addresses two major aspects of PTM experiments: modified and unmodified peptide summarization and visualization, and model-based statistical analysis, in particular testing for differential PTMs across experimental conditions. Additionally, the package supports experiments generated via label-free or tandem mass tag (TMT) labeling methods. MSstatsPTM takes as input indentified and quantified spectral peaks of modified and unmodified peptides as determined by the workflow used to generate the data. It then summarizes the feature intensities in a sub-plot method which involves imputation of missing values,and summarization of intensities using Tukey's median polish. Finally, the package models the summarized features using methods based on the family of linear mixed-effects models, while adjusting PTM measurements for changes in unmodified peptides. The function customizes the model to the specific experiment, taking into account the experiment design, and the acquisition strategy. We demonstrate the performance of MSstatsPTM through multiple case studies of recently published proteomics studies and simulated datasets.

Contact: Devon Kohler, kohler.d@northeastern.edu

P16.05

Proteome-wide and Matrisome-specific Alterations during Human Pancreas Development and Maturation Revealed by 12-plex Isobaric DiLeu Tags

Zihui Li (1), Daniel Tremmel (2), Fengfei Ma (3), Qinying Yu (3), Min Ma (3), Daniel Delafield (1), Yatao Shi (3), Bin Wang (3), Samantha Mitchell (2), Austin Feeney (2), Vansh Jain (2), Sara Sackett (2), Jon Odorico (2), Lingjun Li (4)

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

(2) Division of Transplantation, Department of Surgery, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

(3) School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA

(4) Department of Chemistry & School of Pharmacy, University of Wisconsin-Madison, Madison, WI, USA

The extracellular matrix (ECM) is a key component of the tissue microenvironment and plays an important role in maintaining optimal cell and tissue homeostasis through various biological mechanisms. However, how the proteome, including the ECM, dynamically changes over the lifespan has not been systematically studied in the human pancreas. In this study, we applied mass spectrometry-based quantitative proteomics using our custom developed N,N-dimethyl leucine (DiLeu) isobaric tags to delineate proteome-wide and ECM alterations at four age groups: fetal (18-20 weeks gestation), juvenile (5-16 years old), young adults (21-29 years old) and older adults (50-61 years old). We identified 3,523 proteins including 185 ECM proteins and quantified 117 of them. We detected previously unknown proteome and matrisome features during pancreas development and maturation. For example, we found that some previously reported tumor biomarkers (e.g. THBS1) displayed significant changes in protein expression levels at different developmental stages, which could be important knowledge for studies in biomarker discovery and clinical implementation. We also investigated ECM localization within islet or acinar compartments using immunofluorescent staining and revealed its dynamic changes with age. For instance, we have identified specific ECM proteins enriched in pancreatic islets (e.g. COL6A1, EMILIN1, FBN2, OGN) especially in post-natal stages compared to fetal, which provides new insights for the study of islet development, function and disease. In addition, we have made our dataset very accessible to a broad audience by developing an online searchable database. In summary, our study presents the most comprehensive proteomics analysis across the life cycle of human pancreas development to date and contributes to an improved understanding of the dynamic nature of the matrisome throughout pancreas and islet development. The high-quality normal baseline data will also benefit the field by promoting future investigations into mechanisms of various pancreatic diseases.

Contact: Zihui Li, zli597@wisc.edu

P16.06

Extending Capabilities of Real-Time Database Search on the Orbitrap Eclipse Tribrid Mass Spectrometer for Multiplexed Proteomics

Yang Liu (1), William D. Barshop (1), Amirmansoor Hakimi (1), Jesse D. Canterbury (1), Romain Huguet (1), Rosa Viner (1)

(1) Thermo Fisher Scientific, San Jose, CA, US

Introduction The Real-Time Search (RTS) feature implemented on the Orbitrap Eclipse grants the ability to generate peptide spectrum matches (PSMs) from MS2 spectra during acquisition via database search. Only upon confident identifications, the instrument may be triggered to proceed with SPS-MS3. In this work, we evaluated recent improvements to RTS-based TMT quantitation workflow including: 1) Close-out to prevent acquisition of excess MS3s for proteins which have already reached a sufficient depth of quantitative sampling, 2) online false discovery rate (FDR) assessments to predicate MS3 events on real-time PSM-level FDR estimates. Methods Thermo ScientificTM PierceTM TMT11plex Yeast Digest Standard or TMTpro 16 plex 2 proteome (human and E.coli) samples were analyzed using a Thermo ScientificTM Easy-nLCTM 1000 with 50 cm EASY-SprayTM columns coupled to an Orbitrap Eclipse Tribrid mass spectrometer. Data acquisition was carried out on the 3.4 ICSW. MS2 spectra were analyzed during RTS by the Comet search algorithm, and post-acquisition by Sequest HT in Thermo ScientificTM Proteome DiscovererTM Software 2.5. Results The RTS-MS3 workflow increased quantification accuracy compared to SPS MS3 methods, with the RTS “TMT mode” option (RTS-informed selection of TMT containing fragments for SPS isolation) outperforming the “trigger only” option. The proportion of PSMs with greater than 65% SPS window placements matching identified fragment ions increased by 40% on average when the RTS was applied. Enabling the Close-Out feature maintained quantitative accuracy while further boosting the improvement in quantified proteins. When the RTS FDR-filtering option was enabled, more MS3 spectra were collected compared to a heuristic score cutoff. Conclusions The new features in the RTS-MS3 workflow extend its capabilities for TMT quantitation. Improvements in numbers of quantified proteins and peptides were observed.

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

P16.07

DIA Performance in Discovery and Quantitation Analysis on New Orbitrap Exploris Mass Spectrometers

Yang Liu (1), Khatereh Motamedchaboki (1), Joshua Nicklay (1), Sega Ndiaye (1), Aman Makaju (1), Amarjeet Flora (1), Ryan Bomgarden (1), Daniel Lopez-Ferrer (1)

(1) Thermo Fisher Scientific, San Jose, CA, USA

Introduction Data-independent acquisition (DIA) mass spectrometry (MS) provides deep proteome analysis without the bias from peak intensity. In addition to a more comprehensive coverage of identification, DIA also shows accurate label-free quantification. In this work, we evaluated the DIA performance on both Thermo ScientificTM Orbitrap ExplorisTM 480 and 240 mass spectrometers for discovery and quantitative proteomics analysis. Methods Mixed Hela:E.coli digested peptides (1:2, 1:4, 1:8 ratio) were loaded on 25cm Aurora column (25cm x 75um ID, 1.6um C18) with an Thermo ScientificTM Easy-nLCTM 1200 system, separated by a 90min LC gradient before being injected to Thermo ScientificTM Orbitrap ExplorisTM 240 or 480 MS. Mixed Hela:Yeast peptides was analyzed using the same setup at different gradient length (15min, 30min, 90min). Chromatogram spectral library was built by matching the gas-phase fractionations (GPF) to the predicted spectral library1. Acquired DIA data were analyzed by SpectronautTM 14.0. Results Spectral libraries of 1000ng Hela:E.coli mixture were built on two Orbitrap ExplorisTM platforms and the dynamic ranges were demonstrated to be up to 7 orders of magnitude. By matching with each library, 9056 and 9472 proteins were identified on Orbitrap ExplorisTM 240 and 480, respectively. The Hela:E.coli spike-in proteome acquired on both systems were quantified by directDIATM. ExplorisTM 480 exhibited better quantification accuracy than ExplorisTM 240. Furthermore, we also proved that matching with the library built at a longer gradient and a higher sample load would improve the identification coverage. Therefore, once built at ultimate setup, the spectral library can be used for universal DIA discovery of the same sample. Conclusions Both Thermo ScientificTM Orbitrap ExplorisTM 240 and 480 mass spectrometers had excellent performance in DIA analysis in discovery and quantitation. References 1. Searle, B.C., Swearingen, K.E., Barnes, C.A. et al. Nat Commun 11, 1548 (2020).

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

P16.08

Label-Free Proteomics Performance with New Orbitrap Exploris 480 mass spectrometer with Single-Cell Sensitivity

Daniel Lopez-Ferrer (1), Aaron S. Gajadhar (1), Aman Makaju (1), Aaron M. Robitaille (1), Julia Kraegenbring (1), Tabiwang N. Arrey (1), Michelle Dubuke (1), Emily Chen (1), Alexander Harder (1)

(1) Thermo Fisher Scientific, San Jose, CA, USA

Introduction

LC-MS-based proteomics, a powerful technique for identification and quantification of peptides and proteins in complex samples. needs to provide robustness to analyze 1000s of samples without compromising on proteome coverage and quantitation performance. Here we demonstrate data reproducibility across different laboratories on the new Orbitrap Exploris 480 MS coupled to an High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Interface. The sensitivity of the Orbitrap Exploris 480 MS was evaluated in a data dependent label free method from just a single Hela cell to 5000 HeLa cells with great quantitation accuracy across a wide range using a HeLa:Yeast spike-in experiment.

Methods

Single HeLa cells isolated via fluorescence-activated cell sorting, processed on Nanodroplet Processing in One Pot for Trace Samples (nanoPOTS) platform and Pierce HeLa digest in a range of 0.2-1000ng were analyzed to evaluate instrument sensitivity with different throughputs (30,60, 90 and 120 min gradients) on either the UltiMate™ 3000 RSLCnano (single cells) or EASY-nLC 1200 coupled to an Orbitrap Exploris 480 MS with FAIMS Pro interface. Label free, quantitation performance was evaluated to demonstrate instrument sensitivity and methods reproducibility.

Preliminary data

The performance of this new benchtop mass spectrometer was evaluated in a data-dependent acquisition (DDA) for sample injection amounts of just a single Hela cell to 5000 HeLa cells (~1ug). This instrument sensitivity enables identification of ~7000 protein groups with 5.5 order of magnitude dynamic range from only a 200 ng of bulk HeLa digest and great replicate reproducibility and ~800 protein groups identification from a single Hela cell in 2hr gradient. The method performance and reproducibility were also evaluated across different instruments located in different laboratories around the world with great reproducibility in peptide and protein identification.

Novel aspect

Robust and sensitive Orbitrap Exploris 480 MS, providing throughput and sensitivity needed for label-free proteomics analysis.

Contact: Daniel Lopez-Ferrer, daniel.lopezferrer@thermofisher.com

P16.09

Robust Label Free Proteomics Performance with New Orbitrap Exploris 240 with Improved Separation

Daniel Lopez-Ferrer (1), Yang Liu (1), Aaron S. Gajadhar (1), Julia Kraegenbring (1), Tabiwang N. Arrey (1), Michelle Dubuke (1), Emily Chen (1), Alexander Harder (1)

(1) Thermo Fisher Scientific, San Jose, CA, USA

Introduction

LC-MS-based proteomics analysis is a powerful analytical tool for identification and quantification of thousands of proteins in complex biological samples, but a LC/MS systems need to be robust and easy to use for large scale proteomics analysis without compromising on performance. Here we present a label-free proteomics workflow on a new quadrupole-Orbitrap hybrid mass spectrometer coupled to a High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) interface with a robust analytical setup without compromising on performance. The performance of this new benchtop mass spectrometer was evaluated in a Data-Dependent Acquisition (DDA) across different laboratories.

Methods

LC-MS analysis was performed with an EASY-nLC™ 1200 system with 200 cm, C18 25cm Aurora column (25cm x 75um ID, 1.6um C18) coupled to a new quadrupole-Orbitrap MS with FAIMS Pro Interface. Different concentrations of Pierce HeLa digest are analyzed in DDA, top speed mode with multi CV FAIMS peptide fractionation. Instrument baseline performance across different labs was evaluated using a 2µm, 15 cm PepMap C18 columns at 200 ng HeLa digest in a 60min gradient, DDA method without FAIMS. Proteome Discoverer™ 2.4 software with an improved peptide identification workflow was used data analysis providing improved peptide and protein coverage with a 1% FDR rate.

Preliminary data

Label-free proteomics performance was evaluated in different labs and resulted in ~5000 protein groups with ~10% additional improvement to protein coverage with FAIMS. Untargeted label-free quantitation of peptides in different ratio and with mixed proteome showed great quantitation accuracy. In summary the new Orbitrap Exploris 240 mass spectrometer has shown great sensitivity and data reproducibility for variety of proteomics applications in a multidisciplinary proteomics laboratory with ease of use for different level of user's expertise.

Novel aspect

A novel easy-to-use and robust Orbitrap mass spectrometer with optimized LC-MS and data analysis workflow for high throughput proteomics analysis.

Contact: Daniel Lopez-Ferrer, daniel.lopezferrer@thermofisher.com

P16.10

Investigating bortezomib-induced protein degradation with pulseSILAC-DIA

Lindsay Pino (1), Josue Baeza (1), Richard Lauman (1), Birgit Schilling (2), Benjamin Garcia (1)

(1) University of Pennsylvania, Philadelphia, PA, USA

(2) Buck Institute for Research on Aging, Novato, CA, USA

Stable isotope labeling by amino acids in cell culture (SILAC) coupled to data-dependent acquisition (DDA) is a common approach in quantitative proteomics with the desirable benefit of reducing batch effects during sample processing and data acquisition. More recently, using data-independent acquisition (DIA/SWATH) to systematically measure peptides has gained popularity for its comprehensiveness, reproducibility, and accuracy of quantification. The complementary advantages of these two techniques logically suggests that combining them will lead to enhanced results, yet a benchmark comparison has not yet been published. Here, we explore SILAC-DIA and develop a pulsed SILAC-DIA-MS workflow for protein turnover experiments using free, open-source software. We determine empirically that using DIA achieves similar peptide detection numbers as DDA and further that DIA improves the quantitative accuracy of SILAC by an order of magnitude. We then apply pSILAC-DIA to determine protein turnover rates of cells treated with bortezomib, a 26S proteasome inhibitor FDA-approved for multiple myeloma and mantle cell lymphoma. With our current half-life model, we observe that SILAC-DIA produces more sensitive protein turnover models than SILAC-DDA. Of the 622 proteins tested for differential degradation in the DDA data, 52 were determined statistically significant (q-value < 0.05); of the 1373 proteins tested in the DIA data, 34 were determined statistically significant. Of the proteins determined differentially degraded by both acquisition methods, we find known ubiquitin-proteasome degraded proteins, such as heterogeneous nuclear ribonucleoprotein K (HNRNPK), eukaryotic translation initiation factor 3 subunit A (EIF3A), and eukaryotic initiation factor 4A-I (IF4A1/EIF4A-1), and a slower turnover for cathepsin D (CATD), a protein implicated in invasive breast cancer. As our approach is agnostic to the specific metabolic labeling scheme (e.g. 13C, 15N), the workflow easily generalizes to other labeling approaches. With improved quantification from DIA, we anticipate this workflow making pulsed SILAC-based experiments like protein turnover more sensitive.

Contact: Lindsay Pino, Lindsay.Pino@Pennmedicine.upenn.edu

P16.11

Comprehensive Proteomic Quantification of in vivo Bladder Stone Progression in a Cystinuric Mouse Model using Data-Independent Acquisitions

Jacob Rose (1), Tiffany Zee (1), Nathan Basisty (1), Pankaj Kapahi (1), Marshall Stoller (2), Birgit Schilling (1)

(1) Buck Institute for Research, Novato, Ca, US

(2) UCSF, San Francisco, Ca, US

Abstract Cystinuria is one of various disorders that cause bladder stone formation in humans and is most prevalent in children and adolescents as well as being more aggressive in males. Apart from limited disease management techniques to help solubilize the stones, there is no cure to this disorder. Additionally, recurrence, even after treatment, occurs frequently. There is a significant lack of knowledge pertaining to which factors, besides the build-up of cysteine, can be attributed to formation, expansion, and recurrence of these stones. In this study, we monitored the growth of bladder stones, guided by micro computed tomography (µCT), and profiled dynamic stone proteome changes in a cystinuria mouse model. Considering the stones' composition, we were very surprised to find a significant and dynamic organic component in the stones. Following their development in vivo, bladder stones were harvested and separated into four developmental stages (sand, small, medium and stone) based on their size and digested using a rigorous, 16 hour boiling method for stone solubilization. Data-dependent acquisitions (DDA) and data-independent acquisitions (DIA/SWATH) allowed for deep profiling of stone proteomics and provided over 1300 proteins among all stone sizes. We uncovered the proteomic signatures and pathways that show major changes as the stones grow. We see a remarkable increase in the fibrinogen side chains responsible for blood clots with increasing stone size. We also compared our mouse stones to the Human Kidney and Urine Proteome (HKUPP) and detected 643 overlapping proteins, suggesting high validity of the mouse disease model. We determined that stones initiate from a small nidus, growing outward, and revealed major enrichment in ribosomal proteins and factors related to coagulation and platelet degranulation, suggesting a major dysregulation in specific pathways that can lead to new therapies for this condition.

Contact: Jacob Rose, jrose@buckinstitute.org

P16.12

Draft Quantitative Proteomic Atlas of Human Body and Common Carcinomas

Liang Yue (1), Meng Luo (2), Wenhao Jiang (1), Ning Fan (2), Xiaolu Zhan (3), Fangfei Zhang (1), Fengchao Yu (4), Guoci Teo (4), Alexey Nesvizhskii (5), Ben Collins (6), Ruedi Aebersold (7), Fei Xu (2), Tong Liu (3), Yan Li (8), Tiannan Guo (1)

(1) 1, Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. 2, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China 3, Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China, Hangzhou, Zhejiang Province, China

(2) Department of Anatomy, College of Basic Medical Sciences, Dalian Medical University, No.9 West Section Lvshun South Road, Dalian116044, Liaoning Province, China, Dalian, Liaoning Province, China

(3) Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin 150081, Heilongjiang Province, China, Harbin, Heilongjiang Province, China

(4) Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA., Ann Arbor, Michigan, USA

(5) 1.Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA. 2. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann, Ann Arbor, Michigan, USA

(6) Queen's University Belfast, Belfast, Northern Ireland, United Kingdom, Belfast, Northern Ireland, United Kingdom

(7) 1, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland 2, Faculty of Science, University of Zurich, Zurich, Switzerland, Zurich, Switzerland

(8) Department of Anatomy and Physiology, College of Basic Medical Sciences, Shanghai Jiao Tong University, No.280 Chongqing South Road, Shanghai, 200025, China, Shanghai, China

Introduction: Numerous studies have investigated the proteins expressed in multiple normal and tumorous human tissues and carcinomas using qualitative or data-dependent acquisition (DDA)-based quantitative proteomics, and antibodies. However, systematic quantitative proteomic analysis using a coherent method remains to be accomplished. Methods: We collected 493 normal postmortem tissue samples from human body including 162 tissue types, 19 fetal tissues, and body fluids including blood, urine and saliva. We built a data-independent acquisition (DIA) spectral library using 71 pooled peptide samples from these tissues, fractioned them using high-pH HPLC, analyzed them by DDA-MS on timsTOF Pro, and built tissue-specific and comprehensive libraries using the Spectronaut and FragPipe. We further collected diaPASEF files of 493 normal samples and over 300 carcinomas/adjacent from 23 different organs and analyzed them using Spectronaut. Results: Altogether, we analyzed 752 DDA files from 62 types of human specimens and generated 48 tissue-specific sub-libraries, including ion mobility data, using FragPipe and Spectronaut. At 9857 proteins the fetal tissue library contains the highest number of proteins, whereas the tooth library at 629 proteins has the lowest number of entries. GO enrichment of tissue-specific proteins showed a high correlation to the specific tissue functions. For example, embryo morphogenesis was enriched in fetal tissues and salivary secretion in salivary gland, while the tissue-common proteins were mainly enriched in the metabolism of amino acid and cyclic compounds and exocytosis. We generated a library using FragPipe containing 13,328 proteins and 363,912 peptides, and searched over 700 DIA data sets from human samples acquired by diaPASEF using that library in Spectronaut. The characterization of the thus obtained quantitative proteomes from resulting from these analyses will be described in this presentation. Conclusion: We characterized over 10,000 proteins expressed in over 700 human samples from over 150 types of normal human tissues, body fluids, and common carcinomas.

Contact: Liang Yue, yueliang@westlake.edu.cn

Single-Cell Proteomics

P17.01

Rapid and deep profiling of bacterial proteome with low sample loading using diaPASEF

Guoting Qin (1), Rufeng Li (1), Mengfan Wang (1), Jennifer Copeland (2), Chengzhi Cai (1)

(1) Department of Chemistry, University of Houston, Houston, TX, USA

(2) VICI Valco Instruments Co. Inc., Houston, TX, USA

There is a growing interest in proteomic analysis using fewer than 10 ng of proteins. The demand for low sample loading is partly driven by the limited availability of many biological samples. For example, for research in catheter-associated urinary tract infection (CAUTI) using murine animal model, the amounts of microorganisms in biofilms grown on the inserted catheters are often limited to <100,000 cells (with <10 ng of total proteins),

especially at the early stage of the biofilm formation. It is challenging to achieve deep proteomic profiling of the bacteria in such samples. In this work, we show that this challenge can be overcome by the parallel accumulation – serial fragmentation (PASEF) method combined with data-independent acquisition (dia) performed on a trapped ion mobility mass spectrometry system (timsTOF Pro). Specifically, DIA library of the tryptic digest of E. coli 83972 proteins was prepared using a short (20 min) gradient with 8-24 fractions. With a total gradient time of less than 8 h, we obtained a spectral library covering up to 2100 protein groups. With 50 ng of tryptic digestion in 20 min single runs using a diaPASEF scheme, we could quantify 1972–2003 protein groups. Even decreasing the sample loading to 10 ng and 1 ng, we could still quantify 1394–1430 and 874–959 protein groups, respectively. Overall, we demonstrate that a deep proteome coverage for sub-10 ng of bacterial proteins at a speed of 200 samples per day can be achieved using diaPASEF. This method could be used to monitor the proteome of a small isolated bacterial population. Analysis of the proteome of extra cellular polymeric substances secreted by bacteria during biofilm formation on catheters using this method will be presented.

Contact: Chengzhi Cai, cai@uh.edu

P17.02

Increased Single-Cell Proteome Coverage Using NanoPOTS and Data Independent Acquisition

Richard Carson (1), Thy Truong (1), Yiran Liang (1), Ryan Kelly (1)

(1) Chemistry and Biochemistry, Brigham Young University, Provo, UT, USA

Single-cell proteomics holds great potential for the in-depth study of heterogeneous cell populations such as those found within tumor microenvironments, as well as expanding the data that can be gleaned from trace tissue samples such as needle biopsies. Nanoliter sample processing techniques including nanoPOTS enable the discernment of individual cellular proteomes by vastly decreasing sample losses to surfaces. Combining such sample processing with ultrasensitive LC-MS analyses, >1000 proteins have been detected from single mammalian cells. Here we report the pairing of nanoPOTS, nano-LC, and optimized data independent analysis (DIA) MS acquisition with the goal of dramatically increasing the depth of proteome coverage within a single cell. Our initial experiments using sub-nanogram amounts of commercially available cell lysate tryptic digests indicate that DIA can significantly outperform DDA (data dependent acquisition) methods, increasing the number of peptide identifications by ~50%, under otherwise identical conditions. Progress in the optimization of a DIA method for the proteomic analysis of single cells prepared using the nanoPOTS method is also presented. The ability to deeply and rapidly profile the proteomes of single cells is expected to advance the clinical study of disease conditions such as cancer in which understanding the heterogeneity of the abnormal cellular population can be crucial in making treatment decisions.

Contact: Richard Carson, richardc@chem.byu.edu

P17.04

Fully automated sample processing and analysis workflow for low-input label-free proteome profiling

Yiran Liang (1), Hayden Acor (1), Michaela A. McCown (2), Andikan J. Nwosu (1), Hannah Boekweg (2), Nathaniel B. Axtell (1), Thy Truong (1), Yongzheng Cong (1), Samuel H. Payne (2), Ryan T. Kelly (1)

(1) Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, USA

(2) Department of Biology, Brigham Young University, Provo, UT, USA

Downscaling sample processing volumes to the nanoliter scale using, e.g., nanoPOTS has greatly improved the sensitivity of low-input and single-cell proteomics. However, the required custom robotic and nanoLC-MS/MS systems have limited dissemination to other laboratories. Here we report a fully automated processing and analysis platform termed autoPOTS for low-input proteomics that uses only commercially available parts, which should be broadly enabling for low-input proteomics for many researchers and clinical applications. Cells and tissues were isolated and collected in a 384-well microplate with optical bottom using fluorescence-activated cell sorting and laser capture microdissection. One-pot sample processing was performed by a low-cost Opentrons liquid handling robot (OT-2) with a temperature control module in a total volume of ~6 µL. A ThermoFisher UltiMate autosampler was modified with a 10-port valve for sample desalting and injection. The workflow was then evaluated using in-house-prepared 30-µm-i.d. nanoLC columns and an Orbitrap Exploris 480 mass spectrometer. The evaporation rates at different temperatures were evaluated. OT-2 was programed to periodically dispense buffer or water to compensate for the limited evaporation through the well plate sealing mat during incubation. 1–500 HeLa cells were analyzed to evaluate the autoPOTS workflow. An average of ~200–3000 protein groups were identified, with peptide coverage reduced by a modest 24% for single cells and 12% for 150 cells relative to nanoPOTS. Mouse brain tissues were used to evaluate the compatibility of tissues, with an average of ~2000 proteins identified from a 200 × 200 × 20 µm^3 tissue voxels. We also applied autoPOTS to profile ~1200 proteins from ~130 B and ~130 T lymphocytes isolated from a human donor. General lymphocytes and specific B and T cell functions were characterized using the functional annotation, indicating the precision of this workflow for clinical samples.

Contact: Yiran Liang, lyrzju@gmail.com

P17.05

Proteome-wide Alkylation Enhances Peptide Detectability for Single Cell Proteomics

Xiuyuan Ma (1), Cory Matsumoto (1), Yu Gao (1)

(1) University of Illinois at Chicago, Chicago, IL, USA

Nowadays, the technologies to generate sets of single-cell genomic and single-cell transcriptomic data are reaching maturity. However, single cell proteomics analysis remains to be a challenging task. Unlike nucleotides that could be conveniently amplified by PCR, proteins and peptides cannot be amplified. Moreover, when only hundreds of picograms of proteins are present, the whole experiment suffers from significant sample loss during sample preparation and transfer, leaving only a small percentage of peptides eventually reaching the detector. Many researches are focused on minimizing the sample loss and using other substitutes to solve this problem and have indeed resulted in many great successes. Here we asked a simple question from another angle, instead of amplifying the sample, could we amplify the signal of peptides to achieve a lower detection limit? By alkylating lysine side chain and N-terminal primary amines using various functional groups ranging from 1-6 carbons we observed a significant improvement in data quality from minimum sample as low as a single mammalian cell. Proteome-wide alkylation enhances signal, improves the sequence coverage, and therefore increases the number of peptide identification when only minimum samples are injected. In addition, some of the alkylation also caused unexpected side-effects such as elution order swap. By a thorough investigation and optimization, we determined that a high-yield proteome-wide alkylation could be used to improve the detection of single cell proteomics without compromising the throughput.

Contact: Xiuyuan Ma, xiuyuanm@uic.edu

P17.06

Proteome-wide Alkylation Enhances Peptide Detectability for Single Cell Proteomics

Xiuyuan Ma (1), Cory Matsumoto (1), Yu Gao (1)

(1) University of Illinois at Chicago, Chicago, IL, USA

Nowadays, the technologies to generate sets of single-cell genomic and single-cell transcriptomic data are reaching maturity. However, single cell proteomics analysis remains to be a challenging task. Unlike nucleotides that could be conveniently amplified by PCR, proteins and peptides cannot be amplified. Moreover, when only hundreds of picograms of proteins are present, the whole experiment suffers from significant sample loss during sample preparation and transfer, leaving only a small percentage of peptides eventually reaching the detector. Many researchers are focused on minimizing the sample loss and using other substitutes to solve this problem and have indeed resulted in many great successes. Here we asked a simple question from another angle, instead of amplifying the sample, could we amplify the signal of peptides to achieve a lower detection limit? By alkylating lysine side chain and N-terminal primary amines using various functional groups ranging from 1-6 carbons we observed a significant improvement in data quality from minimum sample as low as a single mammalian cell. Proteome-wide alkylation enhances signal, improves the sequence coverage, and therefore increases the number of peptide identification when only minimum samples are injected. In addition, some of the alkylation also caused unexpected side-effects such as elution order swap. By a thorough investigation and optimization, we determined that a high-yield proteome-wide alkylation could be used to improve the detection of single cell proteomics without compromising the throughput.

Contact: Xiuyuan Ma, xiuyuanm@uic.edu

P17.07

Deep proteomic profiling of human spinal motor neurons using nanoPOTS reveals single-cell protein dynamics in ALS disease

Santosh Misal (1), Amanda Guise (2), Thy Truong (1), Yiran Liang (1), Samuel Payne (3), Edward Plowey (2), Ryan Kelly (1)

(1) Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, USA

(2) Biogen Inc., Cambridge, MA, USA

(3) Biology Department, Brigham Young University, Provo, UT, USA

Introduction: The progressive degeneration of motor neurons (MNs) in human spinal cord and brain can result in fatal or paralytic disorders including amyotrophic lateral sclerosis (ALS)[1]. To increase our understanding of ALS-associated protein dynamics and functions at the level of individual MNs, we collected in-depth single-cell protein expression measurements using nanoPOTS proteomic workflow from ALS and non-disease control (CTL) tissues[2].

Methods: Single MNs from 12-µm-thick fresh-frozen human spinal tissue sections from ALS and CTL donors were isolated by laser capture microdissection and processed in nanoliter droplets. Data were acquired on a Thermo Orbitrap Exploris 480 MS. Protein identification and quantitation were performed with Proteome Discoverer v2.5 and Perseus v1.6.7.0.

Results: To gain greater insight into ALS pathophysiology at single-cell level, we evaluated protein abundance changes in ALS and CTL donors (3 cases/disease group; 6 MNs/case) alongside boost sample (10 MNs) from each donor to facilitate feature identification and precursor matching. Application of stringent data filtering (1% FDR, ≥ 2 unique peptides/protein, ≥ 70% valid values either in ALS or CTL) allowed high confidence quantitative comparison of 929 proteins in ALS and CTL MNs. We identified 271 DE proteins (FDR 5%, FC ≥2) whose abundances were significantly changed (≥2-fold) in ALS relative to CTL MNs, including diminished abundance of numerous ribosomal proteins and translation factors. To identify functional relationships and pathways represented among DE proteins, we performed protein-protein interaction network and gene ontology enrichment analyses and report over-representation of proteins with roles in regulating cellular metabolic processes (n=194 proteins), subcellular localization (n=102), translation (n=89) and oxidative phosphorylation (9) in the DE protein population. Conclusion: Single-cell unbiased proteomic analysis identified multiple proteins and pathways dysregulated in the context of disease.

References: 1. Julien JP. Cell 2001 104 (4):581-591. 2. Zhu Y. et. al. Nat Commun 2018, 9 (1), 882.

Contact: Santosh Misal, samisal@chem.byu.edu

P17.08

Defining the carrier proteome limit for single cell proteomics

Tommy K. Cheung (1), Chien-Yun Lee (2), Florian P. Bayer (2), Atticus McCoy (1), Bernhard Kuster (2), Christopher M. Rose (1)

(1) Discovery Proteomics, Genentech Inc., South San Francisco, CA, United States

(2) Chair of Proteomics and Bioanalytics, Technical University of Munich, Freising, Germany

Single cell proteomics by mass spectrometry (SCoPE-MS) is a recently introduced method to quantify multiplexed single cell proteomes. While this technique has generated great excitement, the underlying technologies (isobaric labeling and mass spectrometry) comprise technical limitations with the potential to effect data quality and biological interpretation. These limitations are particularly relevant when a carrier proteome, a sample added at 25-500x single cell proteomes, is used to enable peptide identifications. Here, we perform controlled experiments with increasing carrier proteome amounts and evaluate quantitative accuracy as it relates to mass analyzer dynamic range, multiplexing level, and number of ions sampled.

Through the analysis of a sample containing four aliquots each of HeLa and K562 cells and various levels of carrier proteome we find that the relationship between the number of ions sampled and the quantitative accuracy of the measurement is dependent on the level of the carrier proteome. Low levels of the carrier proteome (e.g., 5x) display accurate quantification (≤20% CV) even when a small number of ions are sampled, while high carrier proteome levels (e.g., 400x) require many more ions to be sampled for accurate quantification. We also demonstrate that high levels of carrier proteome (e.g., ≥75x) may challenge the dynamic range of the mass spectrometer and limit the detection of single cell ions. Lastly, by diluting our samples to levels at or below the protein amount found in single cells, we demonstrate that common mass spectrometer settings tend to under sample ions for SCP and lead to inaccurate data when utilizing high levels of carrier proteome.

To enable rapid analyses of single cell proteomics data we introduce Single Cell Proteomics Companion (SCPCompanion) a program that enables rapid evaluation of single cell proteomics data and recommends instrument and data analysis parameters for improved data quality.

Contact: Christopher Rose, rose.christopher@gene.com

P17.09

Proteomic Profiling of a Metabolically Defined Mammary Stem Cell Enriched Population

Matthew Waas (1), Pirashaanthy Tharmapalan (1), Rama Khokha (1), Thomas Kislinger (1)

(1) Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

Breast cancer, a heterogeneous disease of multiple origins, is the most common cancer in women. Studies of the various subtypes of breast cancer have revealed similarities between the molecular profile of these subtypes to distinct populations of mammary gland epithelial cells. This evidence supports the hypothesis that the different epithelial cell populations serve as the cell-of-origin for the corresponding subtypes of cancer. Our hypothesis is that defining functional and molecular identities of distinct stem cell populations within the mammary gland epithelium will reveal potential markers and molecular vulnerabilities. Recently, we identified a metabolically defined subpopulation of mammary epithelial basal cells which harbor an increased capacity for colony formation representing a novel putative progenitor/stem population. We have adapted an innovative droplet-digestion sample preparation strategy relying on commercially available consumables and, in combination with systematic optimization of sample preparation, instrument acquisition, and data analysis methods, we reproducibly obtain >3,000 proteins on a Thermo Fusion mass spectrometer from 500 sorted cells. We have performed a repeatability study to characterize the observed variability inherent to sample preparation and data acquisition. Finally, we have applied this platform to profile the proteomic phenotype of the metabolically defined putative stem cells from individual mice revealing potential markers and drug targets. Future studies will compare the proteomic phenotype to stem cell populations sorted with other established markers and will investigate the functional capacity of this population. We anticipate the results of this research will harmonize our understanding of mammary stem cell biology and support breast cancer prevention, prognosis, and treatment efforts.

Contact: Matthew Waas, matt.waas@uhnresearch.ca

P17.10

Mass spectrometry-based single-cell proteomics enhanced by ion mobility separation

Jongmin Woo (1), Geremy C Clair (2), Song Feng (2), Sarah M Williams (1), Chia-Feng Tsai (2), Ronald J Moore (2), William B Chrisler (2), Richard D Smith (2), Ryan T Kelly (3), Ljiljana Pasa-Tolic (1), Charles Ansong (2), Ying Zhu (1)

(1) Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, United States

(2) Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, United States

(3) Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, United States

Unbiased single-cell proteomics (scProteomics) promises to advance our understanding of the cellular composition of complex biological systems. However, a major challenge for current methods is their ability to identify and provide accurate quantitative information for low abundance proteins. Herein, we describe an ion mobility-enhanced mass spectrometry acquisition method, TIFF (Transferring Identification based on FAIMS Filtering), designed to improve the sensitivity and accuracy of scProteomics. The TIFF method enabled unbiased proteome analysis to a depth of >1,700 proteins in single HeLa cells with

1,100 proteins consistently quantified, a significant improvement in overall performance. We applied the TIFF method to obtain temporal proteome profiles of 150 single murine macrophage cells during a lipopolysaccharide stimulation experiment and uncovered unanticipated temporal response trajectories. Further, we demonstrated, to our knowledge, the first application of scProteomics to classify cell populations of a human organ (the lung) without prior antibody labeling.

Contact: Jongmin Woo, jongmin.woo@pnnl.gov

Structural Proteomics, and Native MS

P18.01

Towards Lossless Structural Mass Spectrometry (LSMS): All Ion Unfolding Ion Mobility-Mass Spectrometry for Improved Structural Characterization of Glycoproteins

Gongyu Li (1), Lingjun Li (1)

(1) School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin, US

Structural analysis by native ion mobility-mass spectrometry provides a powerful means to understand protein interactions, stability, and other biophysical information. Commonly, this is assisted by using collision-induced unfolding (CIU) where increasing collisional energy is applied in trap cell, and through the calculated collisional cross section (CCS) of a protein from measurements of drift time in the drift cell. Traditionally for simplification of data analysis and for confident ion identification, precursor ion selection with quadrupole is performed prior to collisional activation. Due to the observed charge state distributions of proteins, this means only one charge state selected at a time, leading to a balance between total time cost across all observed charge states and structural analysis efficiency. Arbitrary charge selection (e.g. of most abundant one) brings an inherent bias in structural analysis towards other charge states, and/or inevitable loss of other charge-structure information. We herein propose the use of charge selection-free analytical workflow – All Ion Unfolding (AIU) – when appropriate, for protein activation with partial exemption from direct charge-structure correlation between those two phases. In summary, towards lossless structural mass spectrometry (LSMS), AIU, when combining with a new data integration method capable of reporting all charge structural information, offers an opportunity to maximally sample protein structural information with minimal time cost, where additional benefits include: 1) improved signal to noise ratios for unfolding fingerprints on same proteins, 2) higher tolerance to charge state shifts that normally affected by operating conditions and 3) better preserved (generally ~2% smaller median CCSs) nativelike confirmations during AIU operation without quadrupole selection. Notably, charge-separated structural information can still be extracted in an on-demand manner during the AIU operation regime.

Contact: Gongyu Li, gongyu.lims@gmail.com

P18.02

Effects of Charge versus Chemical Nature of Divalent Metal Ions on Alpha-Synuclein Structure

Rani Moons (1), Albert Konijnenberg (1), Anne-Marie Lambeir (1), Frank Sobott (2)

(1) University of Antwerp, Antwerpen, Belgium

(2) University of Leeds, Leeds, United Kingdom

The intrinsically disordered protein (IDP) α-synuclein has the propensity to self-aggregate and plays a major role in Parkinson's disease (PD). Various external factors are known to alternate the conformational space of α-synuclein monomers, a conformational transition which can influence the rest of the aggregation pathway and the toxicity of formed aggregate structures. We study the binding capacity of α-synuclein towards a variety of metal ions as well as how they alter the conformational behaviour of monomers. The charge state distribution, stoichiometry of metal ion binding, the relative specificity of binding, the effect of counterions and how all this can influence the monomer conformational space of α-synuclein, is analysed. It is interrogated if conformational effects of α-synuclein upon binding of metal ions are specifically related to the chemical identity of each metal ion, or if there is just a general effect when metal ions with a specific oxidation state bind the protein, which would indicate only a charge driven effect. Using nano electrospray ionisation (ion mobility) mass spectrometry (nESI(-IM)-MS) we compare the effect of binding of one to multiple mono- vs. di- and trivalent ions with increasing metal concentrations, including the role of counterions. Our data shows that the effect of metal ions on the α-synuclein monomer ensemble is linked to the chemical nature of the metal, its oxidation state, the counterion and most likely the mode and place of binding. This indicates that investigation of metal ions binding to an IDP, using nESI(-IM)-MS, reveals important factors that determine binding stoichiometry and conformational change.

Contact: Rani Moons, rani.moons@uantwerpen.be

Top-Down Proteomics, and Proteoform Biology

P19.01

Quantitative Top-down Thermal Proteome Profiling of E. coli lysate and Standard Proteins

Kellye Cupp-Sutton (1), Thomas Welborn (1), Ji Won Kang (1), Si Wu (1)

(1) University of Oklahoma, Norman, OK, United States

Thermal stability of proteins in their native or native-like environments has been shown to be indicative of protein functionality with respect to protein-protein and protein-ligand interactions as well as chemical environment. Thermal proteome profiling techniques have been developed that utilize bottom-up mass spectrometry-based proteomics for the high throughput evaluation of protein thermal stability. We have applied top-down proteomics methods and ultra-high pressure reverse phase liquid chromatography separation to thermal proteome profiling to better understand how proteoform variations such as posttranslational modifications and single nucleotide polymorphisms impact protein thermal stability and, subsequently, protein functionality. Furthermore, we have developed a high-throughput data analysis pipeline that utilizes existing software for proteoform identification and label-free quantitation, TopPIC and Biopharma Finder, respectively, in conjunction with an in-house software for data fitting and interpretation.

We have applied this top-down thermal proteome profiling platform to the analysis of the melting point shift of a standard protein with its known binding ligand (bovine carbonic anhydrase II and acetazolamide), the melting points of 2 proteoforms of the same standard protein (β lactoglobulin A and B), and E. coli cell lysate. We found that our platform was capable of measuring the shift in melting point upon acetazolamide binding with carbonic anhydrase II and measured a stability shift between β lactoglobulin A and B (proteoforms that differ by only 2 single nucleotide polymorphisms). Additionally, our analysis of the soluble E. coli proteome allowed for the determination of the melting points of 78 intact proteoforms. We also identified 3 proteoforms of the acyl carrier protein which demonstrated variable stability depending on the modification of the thiol group on the phosphopantetheine moiety on the serine 36 residue.

Overall, we have demonstrated a quantitative top-down thermal proteome profiling platform for the analysis of proteoform thermal stability in simple and complex protein mixtures.

Contact: Kellye Cupp-Sutton, kellyecuppsutton@ou.edu

P19.02

A novel tissue extraction technique combined with high-field asymmetric waveform ion mobility in top-down proteomics provides new insights into Alzheimer's disease

James Fulcher (1), Aman Makaju (2), Ronald Moore (1), Mowei Zhou (1), David Bennett (3), Phil De Jager (4), Vladislav Petyuk (1)

(1) Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, United States

(2) Life Sciences Mass Spectrometry Unit, Thermo Fisher Scientific, San Jose, California, United States

(3) Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, United States

(4) Department of Neurology, Center for Translational & Computational Neuroimmunology, Columbia University Medical Center, New York, New York, United States

Proteomic investigations of neurodegenerative diseases, such as Alzheimer's, have elicited many new discoveries over the past several decades. Thus far however, these investigations have been largely limited to bottom-up approaches that utilize enzymatic digestion which erase information on how different modifications and variants are distributed across intact proteins. Fortunately, top-down proteomics (TDP) can directly observe intact protein species as combinations of single-nucleotide polymorphisms, alternative exon splice variants, and post-translational modifications (collectively referred to as proteoforms). However, when analyzing complex samples TDP is often limited to observing proteoforms that are most abundant. Therefore offline chromatographic or electrophoretic fractionation techniques are required to reduce sample complexity, which in turn limits throughput. A faster alternative is through gas-phase fractionation such as ion mobility or high-field asymmetric waveform ion mobility spectrometry (FAIMS). Towards this end, we have utilized FAIMS with TDP to analyze cortex tissue samples from patients diagnosed with Alzheimer's disease and found that FAIMS is capable of more than doubling the average number of unique proteoforms observed in a single LC-MS/MS run (1,833 proteoforms with FAIMS compared to 754 without FAIMS with a 1% FDR). We also developed a new chaotropic-cosolvent tissue extraction technique, consisting of hexafluoroisopropanol/urea (HFIP/U), for isolating aggregation-prone proteoforms such as amyloid beta (Aβ). The combination of FAIMS-TDP with HFIP/U extraction allowed for identification of 2,637 proteoforms on average across 94 total cortex tissue samples at 1% FDR. This includes 40 non-redundant Aβ proteoforms identified (10 of which could be characterized across 75% of the samples analyzed) from as little as 20-30 mg wet weight of cortex tissue. We believe this robust methodology will offer new pathways for directly linking proteoforms to Alzheimer's disease processes.

Contact: James Fulcher, james.fulcher@pnnl.gov

P19.03

Optimization of Protein-Level Tandem Mass Tag (TMT) Labeling in Complex Sample with Top-Down Proteomics

Yanting Guo (1), Dahang Yu (1), Kellye A. Cupp-Sutton (1), Xiaowen Liu (2), Si Wu (1)

(1) Department of Chemistry and Biochemistry, University of Oklahoma, Norman, OK, USA

(2) School of Informatics and Computing, Indiana University-Purdue University Indianapolis, Indianapolis, IN, USA

Isobaric chemical tag labeling (e.g., iTRAQ and TMT) has been widely applied for the quantification of peptides and proteins in bottom-up proteomic studies. However, until recently, the quantification of intact proteins using isobaric chemical tag labeling methods has been limited. One issue that inhibits the application of isobaric chemical tags to intact protein labeling is under-labeling and over-labeling of proteins which complicates the analysis of complex protein samples. We will present the systematic evaluation of the labeling reaction parameters on the TMT labeling efficiency of an intact complex protein sample. Our results indicated that (1) under-labeling could happen on both the primary amine groups of the N-termini and lysine side chains. In the case of under-labeling, a higher TMT-to-protein mass ratio and higher initial protein concentration were advantageous for improved labeling efficiency. For example, the initial protein concentration of 1 µg/µL or higher reduced under-labeling when a TMT-to-protein mass ratio of 4:1 or higher was utilized. (2) Furthermore, over-labeling often occurred on the serine, threonine, and tyrosine residues; this over-labeling was often reversed under high pH conditions. We found that a high concentration of the quenching solution (e.g., 1.2% hydroxylamine) can facilitate the reversal of unwanted over-labeling. (3) We also evaluated other reaction conditions such as reaction buffer and reaction time, and our results suggested that these conditions may not significantly affect labeling status. Overall, our study provides practical guidance on performing efficient TMT labeling on complex intact protein samples which can be readily adopted in high-throughput quantitative top-down proteomics.

Contact: Yanting Guo, Yanting.Guo-1@ou.edu

P19.04

A novel charge-selection method enables targeted proteoform-resolved quantification of anti-SARS-CoV-2 neutralizing antibodies in human serum

Kerry Hassell (1), Yu Zhou (1), Debadeep Bhattacharyya (1)

(1) ThermoFisher Scientific, Somerset, NJ, USA

SARS-CoV-2 is the virus that causes COVID-19, the pandemic plaguing the world as of late. The research studies have revelead that anti-SARS-CoV-2 antibodies play an important role in treatment and prevention against and from this deadly virus. Fast, direct and accurate analysis of anti-SARS-CoV-2 neutralizing antibodies, usually with unknown sequences, is critical for understanding the mechanism with which virus-neutralizing antibodies combat viral propagation at the molecular level. Mass spectrometry (MS)-based approaches provide an selective, specific data used for identification and quantitation of therapeutic monoclonal antibodies (Mabs) in the clinical and pharmaceutical laboratory settings. However, use of MS based detection of Mabs with unknown sequences from blood faces technical limitations owing to sample complexity and required sensitivity for low abundance proteoforms. This study demonstrates development of a novel robust LC-Ms assay capitalizing high resolution and mass accuracy with Thermo Scientific™ Exploris™ 240 mass spectrometer that offers full scan windows for middle down subunit MS analysis. An anti-SARS-CoV-2 neutralizing antibodies mix was spiked in serum to test the performance. Enzymatic digestion at the hinge region, reduction and alkylation was done to produce the light, heavy and Fab domain subunits. This selective and reproducible LC-MS method detects down to low concentrations, while simultaneously detecting other proteoforms (glycoforms or oxidations) of antibodies. The high-resolution proteoform-resolved absolute quantification provide high quality insights about the proteoform-pathophysiology. The detection of these antibodies address the urgent needs of absolute mAbs quantification with unknown protein sequence. The experiment described provides high throughput and deploys the measurement of novel antibodies for rapid evolution of viral pathogens.

Contact: Kerry Hassell, kerry.hassell@thermofisher.com

Proteomics of Diseases and Precision Medicine

P20.01

Quantitative Proteomics Analysis of Alveolar Macrophages from Polymeric Ig Receptor-Deficient Mice Reveals Mechanistic Insights into COPD Pathogenesis

Abby Chiang (1), Chelsea Boo (1), Tara Kenny (2), John Kim (2), Antonio DiGiandomenico (2), Lisa Cazares (1), Sonja Hess (1)

(1) Dynamic Omics, Antibody Discovery and Protein Engineering, R&D AstraZeneca, Gaithersburg, MD, USA

(2) Microbial Sciences, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, MD, USA

Secretory IgA (sIgA) deficiency in the chronic obstructive pulmonary disease (COPD) lung is accompanied by reduced polymeric immunoglobulin receptor (pIgR) expression and is a known risk factor for exacerbations and progressive lung function decline. Interestingly, mice deficient in pIgR spontaneously develop a COPD-like lung phenotype as they age, which is characterized by neutrophilia, defective phagocytosis, microbial dysbiosis, and tissue destruction. The underlying mechanisms of disease in animals and how closely it mimics disease progression in COPD is not known. Herein, we performed quantitative proteomics on alveolar macrophages (AMs) to understand whether they play a role in progressive COPD lung disease. AMs were collected by bronchoalveolar lavage from 4- and 6-month old wild type or pIgR-deficient animals. Cells were lysed and processed based on the S-Trap digestion protocol in which proteins were alkylated and reduced in lysis buffer and digested with trypsin/LysC. The eluted peptides were analyzed with data-dependent acquisition (DDA) and data-independent acquisition (DIA) methods using LC-TIMS-QTOF (timsTOF-Pro, Bruker Daltonics). The raw data was analyzed using MaxQuant v1.6.17.0 and Spectronaut v14.0. Quantitative proteomic differences in the two groups of mice were identified using a label-free proteomics approach. Our preliminary data revealed differences in several canonical pathways and gene ontology categories enriched only in AMs derived from 6-month old pIgR-deficient mice. Notably, the most dysregulated pathways included phagocytosis, apoptosis, and inflammation, suggesting correlation with human COPD signatures. Overall, our results suggest the pIgR-deficient mouse could be a promising model to study the pathogenesis of COPD and the identified proteins from the enriched pathways could be further investigated as potential therapeutic targets.

Contact: Abby Chiang, abby.chiang@astrazeneca.com

P20.02

Comparative multiplexed interactomics of SARS-CoV-2 and homologous coronavirus nonstructural proteins

Jonathan Davies (1), Katherine Almasy (2), Eli McDonald (2), Lars Plate (3)

(1) Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

(2) Department of Chemistry, Vanderbilt University, Nashville, TN, USA

(3) Departments of Biological Sciences and Chemistry, Vanderbilt University, Nashville, TN, USA

Human coronaviruses (hCoVs) 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 these severe disease-causing hCoVs, each strain has distinctive virulence. A better understanding of the basic molecular mechanisms mediating changes in virulence is needed. Here, we profile the virus-host protein-protein interactions of three CoV nonstructural proteins (nsps) that are critical for virus replication. We use tandem mass tag-multiplexed quantitative proteomics to sensitively compare and contrast the interactome of nsp2, nsp3, nsp4 from three betacoronavirus strains: SARS-CoV-1, SARS-CoV-2, and hCoV-OC43 – an endemic strain associated with the common cold. In addition, we include nsp3 homologs from MERS-CoV and a common cold alphacoronavirus, hCoV-229E. This approach enabled us to identify both unique and shared host cell protein binding partners and further compare the enrichment of common interactions across homologs. Both nsp2 and nsp4 common interactors are strongly enriched for proteins localized at mitochondria-associated ER membranes suggesting a new functional role for modulating host processes, such as calcium homeostasis, at these organelle contact sites. The nsp3 constructs showed more variation from strain to strain, including interactions with nuclear import machinery for hCoV-229E and with rRNA processing for MERS-CoV. Our results shed light on the role these CoV proteins play in the infection cycle, as well as host factors that may mediate the divergent pathogenesis of common cold CoVs from SARS/MERS strains. Our mass spectrometry workflow enables rapid, robust comparisons of multiple bait proteins, which can be applied to additional viral proteins. Furthermore, the identified, common host-dependencies may present new targets for exploration by host-directed anti-viral therapeutics.

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

P20.03

Quantitative proteomics and phosphoproteomics of urinary extracellular vesicles define diagnostic and prognostic biosignatures for Parkinson's Disease

Marco Hadisurya (1), Li Li (2), Xiaofeng Wu (3), Roy N. Alcalay (4), Shalini Padmanabhan (5), Anton Iliuk (6), W. Andy Tao (7)

(1) Department of Biochemistry, Purdue University, West Lafayette, IN, USA

(2) Tymora Analytical Operations, West Lafayette, IN, USA

(3) Department of Chemistry, Purdue University, West Lafayette, IN, USA

(4) Department of Neurology, Columbia University Irving Medical Center, New York, NY, USA

(5) The Michael J. Fox Foundation for Parkinson's Research, New York City, NY, USA

(6) Department of Biochemistry, Purdue University; Tymora Analytical Operations, West Lafayette, IN, USA

(7) Department of Biochemistry, Purdue University; Tymora Analytical Operations; Department of Chemistry, Purdue University; Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University; Purdue Center for Cancer Research, Purdue University, West Lafayette, IN, USA

Compared to cancer, relatively lower genetic contribution to Parkinson's disease (PD) propels the search for protein biomarkers for early-detection of the disease. Utilizing 82 biological urine samples from the underlying populations - 21 healthy individuals (control), 13 healthy individuals with G2019S mutation in the LRRK2 gene (non-manifesting carrier/NMC), 28 PD individuals without G2019S mutation (idiopathic PD/iPD), and 20 PD individuals with G2019S mutation (LRRK2 PD), here we introduce a novel strategy to determine potential diagnostic and prognostic biomarkers for PD from urinary extracellular vesicles (EVs). After efficient isolation of urinary EVs through chemical affinity followed by mass spectrometric analyses of EV peptides and enriched phosphopeptides, we identified and quantified 4,480 unique proteins and 2,682 unique phosphoproteins. We established 4 panels of proteins and phosphoproteins as novel candidates for disease, genotype, risk, and progression biomarkers, which were further substantiated using ROC and in-depth network analysis. These findings demonstrate a general strategy of utilizing biofluid EV proteome/phosphoproteome as an outstanding and non-invasive source for a wide range of disease exploration.

Contact: Marco Hadisurya, mhadisur@purdue.edu

P20.04

A Robust modular DIA-MS workflow for characterization of three blood proteomes

Angela Mc Ardle (1), Aleksandra Binek (1), Christopher I. Murray (1), Vidya Venkatraman (1), Annie Moradian (2), Blandine Chazarin Orgel (1), Alejandro Rivas (1), Conor Phebus (1), Danica-mae Manalo (1), Simion Kreimer (1), Stephen Pennington (3), Koen Raedschelders (1), Justyna Fert Bober (1), Jennifer Van Eyk (1)

(1) Heart Institute, Los Angeles, CA, USA

(2) Precision Biomarker Lab, Los Angeles, CA, USA

(3) University College Dublin, Dublin, Leinster, Ireland

Many obstacles prohibit the translation of newly discovered protein markers into practical use. It is widely acknowledged that robust discovery workflows are essential to successful biomarker development pipelines and that a major challenge is balancing throughput, sensitivity, and reproducibility. We developed a flexible workflow that allows for high and mid throughput analysis and reliable quantification of proteins in plasma, depleted plasma and dried blood.

DIA-MS data acquisition was performed on an Exploris 480 (Thermo), coupled to either an Evosep One or an Ultimate 3000 chromatography system. Peptides were separated on 3 um x 15 cm C18 columns (Evosep and Phenomenex). A high select TOP14 depletion slurry (Thermo Scientific) was used to deplete plasma. Digestions were performed on an automated workstation (Beckman i7). Samples were denatured (30 % 2-2-2 Trifluoroethanol) and 40 mM dithiothreitol), alkylated (10 mM Iodoacetamide) and digested with trypsin (25:1). Five samples replicates were prepared to assess reproducibility on 3 separate days. Data was processed with OpenSwath search engine, mapDIA and plasma pilot.

Our high throughput workflow supported quantification of proteins with median intra-day Cvs ranging from 8-10 %, 12-15 % and 15-17 % while inter-day CVs were 18 %, 15 % and 20 % in plasma, depleted plasma and dried blood respectively. The mid throughput workflow supported quantification of proteins with median intra-day CVs of ranging from 10-13 %, 4-8 % and 10-13 % while inter-day CVs were 24%, 15 % and 18% in plasma, depleted plasma and dried blood respectively. The total number of proteins identified were 200, 683, 366 and 504, 473, 578 in plasma, depleted plasma and blood analyzed by the high and mid throughput platforms respectively.

We envision that implementation of our workflow will lead to more efficient and effective translation of candidate markers into clinical use through simplified and standardized discovery methods.

Contact: Angela Mc Ardle, angela.mcardle@cshs.org

P20.05

Serum Proteomics in Humans and Non-Human Primates

Sumaiya Nazli (1), Michael Olivier (2)

(1) Center of Precision Medicine, Wake Forest School of Medicine, Winston Salem, NC, United States

(2) Center for Precision Medicine, Wake Forest School of Medicine, Winston Salem, NC, United States

Protein biomarker signatures in plasma are often proposed as a less invasive means to assess proteomic changes in organs and tissues, which cannot be sampled over time in human patients. However, due to the complex composition of plasma, MS analysis is complicated by a small number of highly abundant proteins (14 proteins, including albumin and immunoglobulin, account for over 90% of total plasma protein), and the correlation between plasma protein changes and tissue changes often is unclear. Non-human primates (NHP) resemble humans in their physiology, genetics, and disease pathobiology. They are often used for cardiometabolic studies since tissue samples can be obtained under highly controlled conditions.

Here, we examined whether standard plasma proteomics approaches developed for human samples can be used to assess the plasma proteome of NHP.

We analyzed human NIST plasma my mass spectrometry after depleting the top 14 abundant proteins (A36369), followed by reduction, alkylation and tryptic digestion. Peptides were further fractionated using high pH RP peptide fractionation, and analyzed using a Fusion Lumos Orbitrap mass spectrometer. Based on peptide ion intensities, the depleted proteins still accounted for 13-30% of total proteins. Depletion was less efficient when NHP plasma was used, with depleted proteins still accounting for 42-64% of total proteins detected by MS. Overall, MS analysis identified an average of 73 proteins in unfractionated and 182 proteins in pH-fractionated samples with 20µg of starting protein, and 208 and 257 proteins in unfractionated and fractionated samples with 100µg starting protein. Similarly, plasma from NHP resulted in identification of 65 and 218 proteins in unfractionated samples of 20µg and 100µg, respectively. PH-fractionated samples containing 20µg of protein resulted in identification of 162 proteins, and 305 proteins were identified with 100µg, suggesting that the overall depletion and analysis protocol works effectively in NHP.

Contact: Sumaiya Nazli, snazli@wakehealth.edu

P20.06

Longitudinal Study of African Tuberculosis Patients

Samantha M. Powell (1), Lindsey Anderson (1), Leah G. Jarlsberg (2), Marina A. Gritsenko (1), Payam Nahid (2), Jon M. Jacobs (1)

(1) Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Lab, Richland, WA, USA

(2) School of Medicine, University of California San Francisco, San Francisco, CA, USA

Tuberculosis (TB) is one of the top 10 worldwide leading causes of death and the leading cause of death from a single infectious agent. TB is especially prevalent in Africa, which accounts for 84% of all TB-related deaths. In this study, serum samples were collected from 30 patients participating in an African region CDC-conducted Phase 2 TB clinical trial. Samples were collected at five time points across the course of one year encompassing pre- through post- treatment timepoints. A robust quantitative TMT-labeled approach was employed with proteome analysis via a sensitive LC-MS/MS platform. Following MS analysis and downstream protein level quantification, statistical analysis was performed to identify discriminatory protein/pathway signatures stratified based upon disease severity, i.e., cavitation status, for treatment, end of treatment, and post-treatment timepoints. Identification of protein signatures reflective of treatment response and post-treatment health resolution remained stratified based upon initial disease severity. The regulation of inflammatory and acute phase host responses, among other pathways, helped drive disease severity discrimination, even after >20 weeks post treatment. Identification of treatment efficacy markers that reflect resolution of infection in variable disease severity patients will help determine optimum regimens and possibly identify patients more susceptible to relapse. Additionally, understanding the patient post-treatment recovery period will provide information on post-treatment care needs and “return to normal” expectations.

Contact: Samantha M. Powell, samantha.powell@pnnl.gov

P20.07

Implementation Of Holistic Glycopeptide PASEF-DDA Data In Diagnostics

Hans JCT Wessels (1), Purva Kulkarni (1), Fokje Zijlstra (1), Merel Post (1), Melissa Bärenfänger (1), Pierre-Olivier Schmit (2), Alain J Van Gool (1), Dirk J Lefeber (1)

(1) Translational Metabolic Laboratory, Radboudumc, Nijmegen, Netherlands

(2) Bruker Daltonics, Wissembourg, France

Molecular diagnostics is on the verge of implementing high-throughput functional Omics data in routine clinical practice for high-precision personalized healthcare. Shotgun glycoproteomics in blood plasma offers unique possibilities for clinical diagnostics by providing site-specific glycosylation data for up to hundreds of proteins by a single measurement. The timsTOF Pro platform with parallel accumulation serial fragmentation (PASEF) provides the required data acquisition speed, sensitivity and instrument robustness for the highly anticipated application of holistic data in routine diagnostics at healthcare centers. Since both biomarker discovery and diagnostics can be performed using the same holistic data, the use of PASEF-DDA effectively eliminates the tedious process of developing and applying different methods for untargeted biomarker discovery and target biomarker measurement. However, implementation of holistic glycopeptide data in routine diagnostics poses considerable challenges with respect to data processing/analysis/management, clinical compliance and lucid reporting of test-results to clinicians. Our overall strategy to implement high dimensional holistic LC-IM-MS data in diagnostic workflows is to perform targeted data extraction for pre-defined glycopeptide biomarkers to enable efficient automated data processing and intelligent reporting by dedicated bioinformatic pipelines. In our contribution to the US HUPO 2021 conference we will discuss the diagnostic workflow implementation for diagnostics in congenital disorders of glycosylation (CDG) at the Translational Metabolic Laboratory of the Radboud university medical center in the Netherlands.

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

P20.08

Sirtuin 6 at the interface between gene expression and cellular metabolism during viral infection

Matthew D Tyl (1), Laura A Murray-Nerger (1), Yana V Miteva (2), Michelle A Kennedy (1), Joel D Federspiel (3), Ileana M Cristea (1)

(1) Department of Molecular Biology at Princeton University, Princeton, NJ, United States

(2) Columbia University, New York City, NY, United States

(3) Drug Safety Research Unit at Pfizer, Princeton, NJ, United States

Sirtuins are NAD+-dependent enzymes that function as “sensors” of the cellular environment by transmitting information through the addition or removal of posttranslational modifications (PTMs). Although initially designated as Class III histone deacetylases, recent work has demonstrated that sirtuins have a diverse range of enzymatic activities and substrates. One example is sirtuin 6 (SIRT6): initially characterized as a nuclear lysine deacetylase acting on histone 3 (H3), it has also been found to 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 cellular metabolism. Our lab previously demonstrated that SIRT6 restricts the replication of several viruses, including human cytomegalovirus (HCMV). However, the mechanism(s) underlying this antiviral effect remains unknown. Here, by generating CRISPR-mediated SIRT6 knockouts and SIRT6 overexpression, we verified the antiviral effect of SIRT6 against HCMV in primary human fibroblasts. To assess whether SIRT6 restricts viral gene expression, we developed a targeted mass spectrometry approach to quantify nearly one hundred HCMV proteins. SIRT6 overexpression resulted in a global reduction in viral protein abundances early in infection; however, viral proteins recovered by the intermediate stage of infection. In parallel, w e investigated SIRT6 protein interactions throughout infection using IP-MS to ascertain antiviral effectors of SIRT6. Following assessment of interaction specificity using the SAINT algorithm, we observed SIRT6 interactions with proteins involved in DNA damage repair, immune response, and transcriptional regulation. The interactions of SIRT6 with chromatin remodelers and transcription factors suggest a role in regulating

transcription, in particular at the early stages of infection. Altogether, our results lead us to a model in which SIRT6 exerts its antiviral effect by repressing metabolic genes, thereby interfering with virus-induced metabolic reprogramming.

Contact: Matthew Tyl, mtyl@princeton.edu

P20.09

Qualitative and quantitative proteomic and metaproteomic analyses of normal human urine sediment

XiaoLian Xiao (1), HaiDan Sun (1), Chen Shao (2), Wei Sun (1)

(1) Core Facility of Instrument, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, China

(2) State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing, China

Urine sediment is an integral component of urine and has been commonly utilized for disease diagnostics and health status monitoring. In this study, using 19 normal human urine sediments (5 adult males, 5 adult females, 5 child males, and 4 child females), we comprehensively analyzed the human proteome and metaproteome using a 2DLC/MS/MS approach. Next, interindividual variations in the human proteome and metaproteome were analyzed through the DIA approach. Finally, the physiological influencing factors (sex and age) of urine sediment were evaluated. As a result, a total of 2736 human proteins were identified in urine sediment, and their functions were primarily associated with the inflammatory response, energy metabolism, cell adhesion and oxidoreductive response. For the metaproteome, 65 genera were identified that were primarily involved in translation and carbohydrate metabolic processes. The median biological CV of the proteome/metaproteome of human urine sediment was 0.5/0.72, similar to the proteome of human urine supernatant. In addition, sex and age were observed to affect the proteome and metaproteome of human urine sediment. The results of this study serve to characterize human urine sediment and indicate that urine sediment might represent an alternative resource for disease research in addition to urine supernatant, but the influence of sex and age must be considered in the study design process. These results may facilitate the application of urine sediment to clinical research.

Contact: XiaoLian Xiao, 770877855@qq.com

AUTHOR INDEX

A

B

C

D

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z