3.1

Microscaled Cancer Proteogenomics, Progress and Barriers to Clinical Translation

Matthew J. Ellis[1]

[1]**Lester and Sue Smith Breast Center, Baylor College of Medicine**

Tumor analysis using next generation sequencing (NGS) of DNA is increasingly routine, however the impact in many disease areas has been modest, as therapeutic advances identified through single gene/mutation/drug therapeutic models often involve modest numbers of patients with therapeutic effects limited by clonal selection and resistance. Tumor immunotherapeutic approaches can be guided by NGS-based mutational load and mutational signature -based diagnostics but DNA sequencing does not probe the immune microenvironment of the tumor which is the determinant of the therapeutic outcome. Our Clinical Proteomic Tumor Analysis Consortium supported effort is addressing these issues using micro-scaled mass spectrometry-based proteomics. Tumor proteins are isolated from tumor core biopsies before DNA is prepared for NGS, effectively rescuing key analytes that are often discarded by genomics laboratories. We have applied our prototype pipeline to core biopsies accrued from patients receiving neoadjuvant chemotherapy for ErbB2 positive breast cancer. Our analysis approach, in combination with cloud-based proteogenomic analysis tools, promises a more complete evaluation of tumor biology, consequently a more accurate diagnosis of ErbB2 positive breast cancer. We are currently extending this approach to other breast cancer subsets and the overall strategic approach will be discussed.

3.2

Proteins to pathways: Functional insights from comprehensive proteomics and phosphoproteomics

Karin Rodland[1]

[1]**Pacific Northwest National Laboratory**

Changes in biological function represent a highly coordinated change in the flow of information within and between cells, some of which is manifested at the level of changes in mRNA and protein abundance - but those changes are just the tip of the informational iceberg. Changes in the level of specific post-translational modifications, particularly protein phosphorylation, can have significant impacts on cell function without necessarily impacting protein abundance. Deep quantitative phosphoproteomics provides a tool for mapping the flow of information within and between cells, potentially going beyond the canonical pathways that are well represented by antibody-based assays, and enabling the unbiased assessment of coordinated changes in interacting networks. This talk will demonstrate how comprehensive quantitative data from proteomic and phosphoproteomic studies of several cancer types (ovarian cancer, uterine cancer, and acute myeloid leukemia) can be used to identify interacting pathways and generate mechanistic hypotheses relevant to cancer prognosis and treatment.

3.3

Non MS-based techniques for proteomic profiling of large human cohorts: Trust . . . but verify

Robert Gerszten[1]

[1]Beth Israel Deaconess Medical Center, Harvard Medical School

Recent advances in non-mass spec based proteomic technologies have made high throughput profiling of low abundance proteins in large epidemiological cohorts feasible. We investigated whether aptamer based proteomic profiling could predict development of type 2 diabetes (T2DM) beyond known risk factors. We identified dozens of markers with highly significant associations with future T2DM across two large longitudinal cohorts (n = 2,839) followed for up to 15 years. We leveraged proteomic, metabolomic, genetic and clinical data from humans to nominate specific candidates to test for potential causal relationships in model systems. We will present studies of a top protein association with future diabetes risk on glucose homeostasis in vitro and in vivo. Further, loss-of-function variants associated with circulating levels of a novel biomarker were in turn associated with fasting glucose, hemoglobin A1c and HOMA-IR measurements – suggesting a causal role in disease pathogenesis. In addition to identifying novel disease markers and potential pathways in T2DM, we provide publically available data to be leveraged for new insights about gene function and disease pathogenesis in the context of human metabolism. The talk will also highlight applications to other cardiometabolic diseases as well as work flows established to verify the specificity of the observations.

3.4

The Proteogenomic Landscape of Lung Adenocarcinoma

Michael A. Gillette[1,2], Shankha Satpathy[1], Karsten Krug[1], Karl Clauser[1], Mohan Dhanasekaran[3], Suhas Vasaikar[4], Song Cao[5], Sara Savage[6], Francesca Petralia[7], Wen-wei Liang[5], Xiaoyu Song[7], Lili Blumenberg[8], Shaleigh Smith[8], Lijun Yao[5], Jiayi Ji[7], Lauren Tang[1], Melanie MacMullan[1], Shayan Avanessian[1], Harry Kane[1], David Heiman[1], Ramani Kothadia[1], Matthew Wyczalkowski[5], Mehdi Mesri[9], Ana Robles[9], Henry Rodriguez[9], Gil Omenn[3], Arul Chinnaiyan[3], Marcin Cieslik[3], Gaddy Getz[1,2], Alexey Nesvizhkii[3], David Fenyo[8], Bing Zhang[8], Kelly Ruggles[8], Ramaswamy Govindan[5], Pei Wang[7], Li Ding[5], D R. Mani[1], Steven Carr[1]

[1]Broad Institute Of MIT And Harvard, [2]Massachusetts General Hospital, [3]University of Michigan, [4]MD Anderson Cancer Center, [5]Washington University, St. Louis, [6]Baylor College of Medicine, [7]Mt. Sinai School of Medicine, [8]New York University, [9]National Cancer Institute

A persistent central deficiency in our knowledge of cancer concerns how genomic changes drive the proteome and phosphoproteome to execute phenotypic characteristics. Furthermore increasing evidence implicating epigenetic and post-translational changes in cancer biology reinforce the notion that molecular profiles based on nucleic acids are incomplete and are critically complemented by analyses of proteins and their post-translational modifications (PTMs). We present the first integrated proteogenomic study on a prospectively collected lung adenocarcinoma (LUAD) cohort, and provide new insights including on molecular taxonomy, novel mutations and fusions, protein and PTM associations with canonical driver mutations, metabolic dependencies, and the immune milieu. 102 LUAD tumor/normal adjacent tissue pairs and 9 additional LUAD tumors were collected under the auspices of the National Cancer Institute's Clinical Proteomics Tumor Analysis Consortium using rigorous standard protocols to minimize ischemic time and other pre-analytical variability. Approximately equal numbers of Eastern (China, Vietnam) and Western patients were enrolled and the population included ~ 40% never-smokers. Samples underwent comprehensive genomic and proteomic characterization, providing whole exome, whole genome, copy number, RNAseq, microRNA, long non-coding RNA, methylation, global proteome, phosphoproteome, and acetylome data. The distribution of top driver mutations paralleled that of large genomics studies; both novel structural variants in established driver genes and novel ALK fusion partners were defined. Association analyses highlighted important outliers seen only in the phospho data, including potential therapeutic targets such as SOS1 in KRAS mutant and Shp2 in EGFR mutant tumors. Novel KEAP1 mutants were described including one suggesting an alternative mechanism of NRF2 regulation. Multi-omics clustering revealed four distinct clusters, variably enriched for country of origin, gender, smoking and mutation status. An STK11-enriched cluster had a notably "cold" immune landscape; mechanistic hypotheses were nominated for this immune regulation. Kinase outlier analyses suggested novel therapeutic possibilities, while tumor-normal analyses defined candidate diagnostic biomarkers and helped illuminate carcinogenesis. While these and other analyses are intended to provide new insights into LUAD biology and facilitate testable therapeutic hypotheses, this diverse, densely characterized and closely annotated sample population provides a vast dataset that should be an important resource for the lung cancer and broader scientific communities.

3.5

The human iPS Cell Proteome in Health & Disease

Angus I. Lamond[1]

[1]University of Dundee

Deep mining of proteomes, using mass spectrometry (MS) based technology, can provide invaluable insights, at a systems level, into both physiological responses in healthy cells and mechanisms causing disease phenotypes. A further challenge concerns how to analyse and integrate these proteomic data with other parallel 'omics' and cell phenotypic data and how to manage the large resulting volumes of complex information. I will describe our progress in using quantitative proteomics for the large-scale analysis of human induced pluripotent stem cells (iPSCs), involving iPSC lines derived from both healthy donors and patient cohorts with specific inherited disorders. While previous studies have characterized iPS cell lines genetically and transcriptionally, little is known about the variability of the iPSC proteome. We have generated a deep proteome of human iPS cells and characterised the major determinants affecting proteome variation across human iPSC lines from healthy donors, thereby identifying key regulatory mechanisms affecting differential protein abundance. These data identified >700 human iPSC protein quantitative trait loci (pQTLs), for which we mapped trans regulatory effects and demonstrated an increased enrichment in disease-linked GWAS variants, compared with RNA-based eQTLs. Analysis of iPSC lines derived from donors with mutations causing the ciliopathy, Bardet Biedl Syndrome (BBS), show specific alterations in the expression of multiple proteins affecting pathways and functions relevant to all of the main signs/symptoms seen in BBS patients. Remarkably, these proteome changes in BBS lines, relative to lines from healthy donors, are already detected in undifferentiated iPS cells, even though the patient defects occur in differentiated tissues. To facilitate sharing with the community, we have incorporated all these data into our open access, searchable online EPD database (see; www.peptracker.com/epd).

3.6

Modulating nerve growth through RNA-protein interactions

Jeffery L. Twiss[1]

[1]University of South Carolina

Proteins synthesized in axons are used for regeneration of injured peripheral nerves and several lines of evidence indicate that axons in the brain and spinal cord can also translate mRNAs. RNA binding proteins (RBPs) regulate the transport, stability, and translation of subcellularly localized mRNAs, but only a few RBPs are known to localize into axons. RNA affinity mass spectrometry shows a complex population of RBPs interacting with the localization motifs of axonal Nrn1, Hmgb1, Gap43, and Actb mRNAs, but no proteins are shared between all 4 RNA motifs. A more systematic profile of the RBPs localizing into peripheral nerve axons from targeted mass spectrometry shows that axonal levels of hnRNPs binding to Nrn1 and Hmgb1 motifs increase after nerve injury and during regeneration. RNA-seq from axoplasm immunoprecipitates of these hnRNPs define 'axonal growth' RNA regulons compared to those for the RNA chaperone protein La/SSB. We had previously shown that KHSRP (also called FUBP2) and HuD proteins bind to Gap43 and Nrn1 mRNAs, with HuD stabilizing and KHSRP destabilizing these mRNAs. Both KHSRP and the related FUBP1 protein are increased in axons during regeneration. Proximity biotinylation studies show that KHSRP and FUBP2 proteins interact in axons. Interestingly, Fubp2 mRNA is also destabilized by KHSRP interaction, and FUBP2 protein levels increase in the absence of KHSRP. Deletion of the murine KHSRP gene alters axon growth in developing and injured neurons, indicating that KHSRP functions to attenuate axon growth, likely through neuronal mRNA destabilization. Together, these data indicate that axon growth potential is determined by RBP-mRNA interactions that regulate axonal mRNA dynamics.

[This work is supported by funding from the following: NIH (R01-NS041596 and R01-NS089633), South Carolina Spinal Cord Injury Research Fund, and the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation].

3.7

Mapping Axon Initial Segment Structure and Function by Multiplexed Proximity Biotinylation

Hamdan Hamdan[1], Brian C. Lim[1], Tomohiro Torii[1], Abhijeet Joshi[1], Matthias Konning[1], Cameron Smith[1], Donna J. Palmer[2], Philip Ng[2], Christophe Leterrier[3], Juan A. Oses-Prieto[4], Alma L. Burlingame[4], Matthew N. Rasband[1]

[1]Department of Neuroscience, Baylor College of Medicine, Houston, TX, [2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, [3]Aix-Marseille Univ, CNRS, INP, NeuroCyto, Marseille, France, [4]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA

Axon initial segments (AIS) generate action potentials and regulate the polarized distribution of proteins, lipids, and organelles in neurons. While the mechanisms of AIS Na+ and K+ channel clustering are understood, the molecular mechanisms that stabilize the AIS and control neuronal polarity remain obscure. To identify the AIS proteome, we used proximity biotinylation and mass spectrometry. We constructed AIS-targeted biotin ligase by generating fusion proteins of BirA* with NF186, Ndel1, and Trim46; these chimeras mapped the molecular organization of intracellular membrane, cytosolic, and microtubule compartments of the AIS, respectively. Separately, we live-labeled neurons using antibody-targeted proximity biotinylation to map the protein composition of the extracellular membrane of the AIS. Our experiments revealed a diverse set of biotinylated proteins not previously reported at the AIS. We show many are located at the AIS, interact with known AIS proteins, and their loss disrupts AIS structure and function. Our results provide new conceptual insights and an unprecedented resource for AIS molecular organization, the mechanisms of AIS stability, and polarized trafficking, among other functions, in neurons.