

Comprehensive, Quantitative, Intact Proteoform Measurements of Patient-Derived Breast Tumor Xenografts Using an Improved Top-Down Proteomics Pipeline

Tao Liu¹, Paul Piehowski¹, Samuel Payne¹, Sangtae Kim¹, Jungkap Park¹, Christopher Wilkins¹, Carrie Nicora¹, Yufeng Shen¹, Rui Zhao¹, Anil Shukla¹, Ronald Moore¹, Sherri Davies², Shunqiang Li², Reid Townsend², Matthew Ellis³, Emily Boja⁴, Henry Rodriguez⁴, Karin Rodland¹, and Richard D. Smith¹

¹Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA; ²Department of Medicine, Washington University, St. Louis, MO; ³Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX; ⁴National Cancer Institute, National Institutes of Health, Bethesda, MD



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Overview

- Many biologically active proteoforms cannot be identified or distinguished by conventional bottom-up approaches since the information is lost with trypsin digestion.
- Integrating data from bottom-up and top-down proteomics measurements with available genomic data enables characterization of protein variants, including coding polymorphisms, processing products, and post-translational modifications.
- There is an urgent need to be able to make comprehensive and reproducible measurements of proteoforms in biological or clinical samples.
- The available patient-derived xenografts (PDX) for two breast cancer subtypes (luminal B and basal)¹ provide an ideal system for this evaluation.

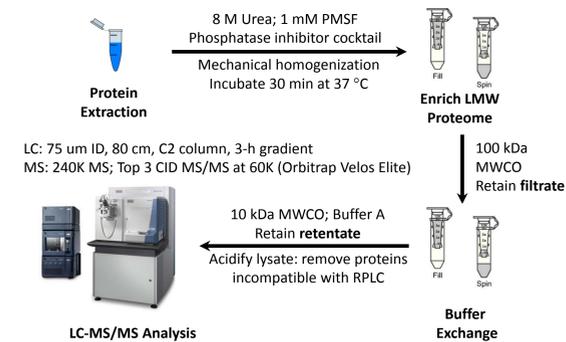
Introduction

- The widely used bottom-up proteomics approaches in many cases cannot identify or distinguish many different processed and/or modified versions of the gene products whose coverage is important for understanding the development and treatment of cancers.
- Bottom-up proteomics approaches are inherently limited by the "peptide-to-protein" inference problem, affecting both accuracy and precision of the quantification.
- We have recently developed a significantly improved top-down proteomics pipeline that allows for comprehensive quantitative intact proteoform measurements.
- Application of this pipeline to the comparative analysis of PDX models of basal and luminal B human breast cancer provided much improved performance in the ability to obtain both comprehensive proteome coverage, and precise intact proteoform quantification, resulting in novel biological insights otherwise unseen in bottom-up analysis approaches.

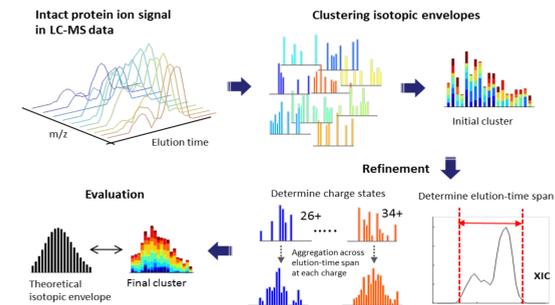
Methods

Sample preparation

Washington human-in-mouse (WHIM) PDX samples representing luminal B (P6 and P33) and basal (P5 and P32) breast cancer subtypes

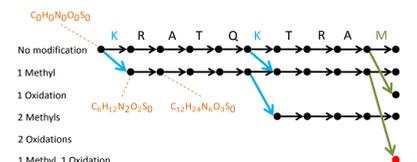


LC-MS feature finding in ProMx



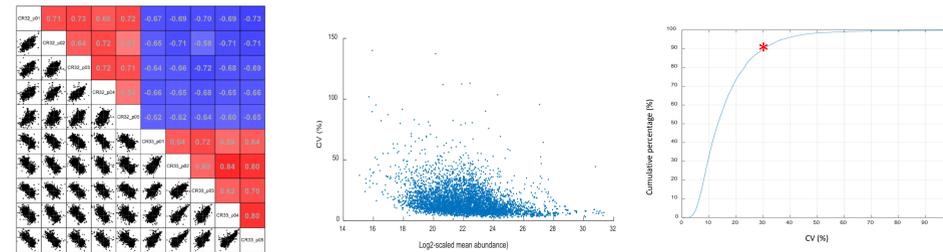
- 1) clustering isotopomer envelopes across adjacent time and charge state; 2) the initial cluster is refined to accurately determine its elution-time span and range of charge states; 3) calculates the likelihood that the final cluster is an LC-MS feature.

Proteoform ID by sequence graphs



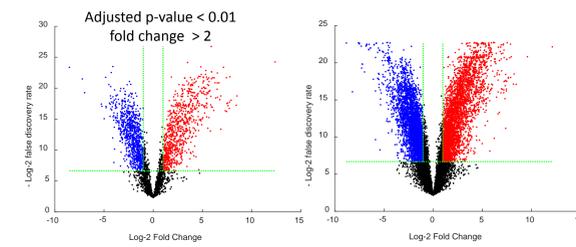
Results

Comprehensive and reproducible detection of proteoforms



P32: basal (WHIM2)
P33: luminal B (WHIM16)

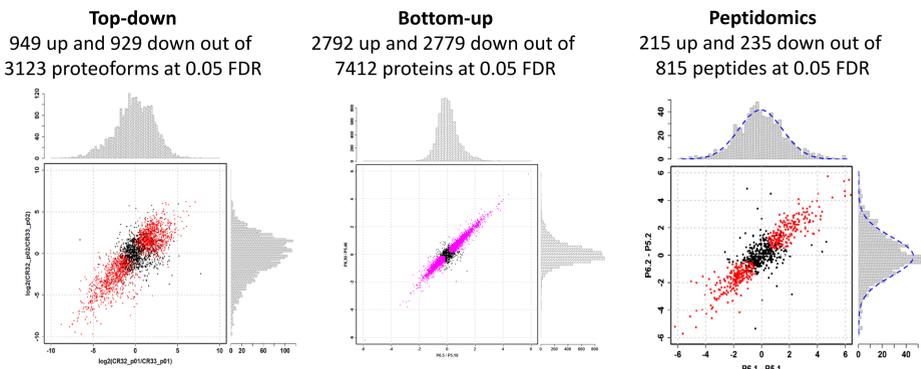
Precise quantification of the proteoforms



A total of 1780 differentially expressed, out of 3123 proteoforms

A total of 7392 differentially expressed, out of 14412 clustered features

Comparison of top-down, bottom-up, and peptidomics results



Top-down vs. bottom-up: Enriched pathways

Upregulated proteins

REACTOME_INFLUENZA_VIRAL_RNA_TRANSCRIPTION_AND_REPLICATION
REACTOME_PEPTIDE_CHAIN_ELONGATION
REACTOME_VIRAL_MRNA_TRANSCRIPTION
KEGG_TRANSLATION
REACTOME_REGULATION_OF_GENE_EXPRESSION_IN_BETA_CELLS
KEGG_RIBOSOME
REACTOME_GTP_HYDROLYSIS_AND_JOINING_OF_THE_60S_RIBOSOMAL_SUBUNIT
REACTOME_METABOLISM_OF_PROTEINS
REACTOME_FORMATION_OF_A_POOL_OF_FREE_40S_SUBUNITS
STRUCTURAL_CONSTITUENT_OF_RIBOSOME
REACTOME_INSULIN_SYNTHESIS_AND_SECRETION
STRUCTURAL_MOLECULE_ACTIVITY
REACTOME_PACKAGING_OF_TELOMERE_ENDS
REACTOME_RNA_POLYMERASE_I_PROMOTER_OPENING
REACTOME_RNA_POLYMERASE_I_PROMOTER_CLEARANCE
REACTOME_TELOMERE_MAINTENANCE
REACTOME_TRANSLATION_INITIATION_COMPLEX_FORMATION
REACTOME_PROCESSING_OF_CAPPED_INTRON_CONTAINING_PIE_MRNA
REACTOME_RNA_POLYMERASE_I_III_AND_MITOCHONDRIAL_TRANSCRIPTION
REACTOME_TRANSCRIPTION
KEGG_SPLICOSOME
KEGG_OXIDATIVE_PHOSPHORYLATION
REACTOME_INTEGRATION_OF_EXPRESSION_METABOLISM
REACTOME_PREFOLDIN_MEDIATED_TRANSFER_OF_SUBSTRATE_TO_CCT_TRIC
REACTOME_ELECTRON_TRANSPORT_CHAIN
REACTOME_MRNA_SPLICING_MINOR_PATHWAY
NUCLEOSIDE_NUCLEOTIDE_AND_NUCLEIC_ACID_METABOLIC_PROCESS

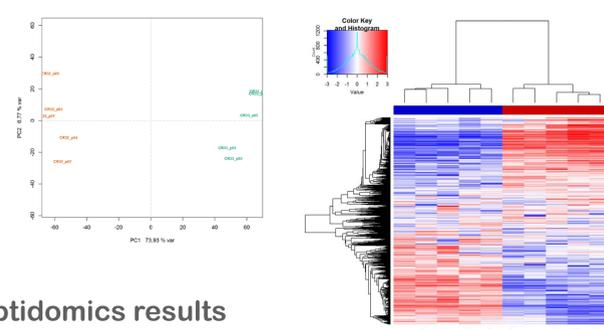
Downregulated proteins

HSIAO_HOUSEKEEPING_GENES
REACTOME_DIABETES_PATHWAYS
KEGG_OXIDATIVE_PHOSPHORYLATION
KEGG_TRANSCRIPTION
REACTOME_FORMATION_OF_ATP_BY_CHEMOSMOTIC_COUPLING
PROTON_TRANSPORTING_TWO_SECTOR_ATPASE_COMPLEX
BIOCARTA_PROTEASOME_PATHWAY
REACTOME_GLYCOSE_TRANSPORT
REACTOME_REGULATION_OF_INSULIN_SECRETION
REACTOME_INTEGRATION_OF_ENERGY_METABOLISM
MITOCHONDRIAL_INNER_MEMBRANE
MITOCHONDRIAL_ENVELOPE

Fisher exact test
adjusted p value < 0.05

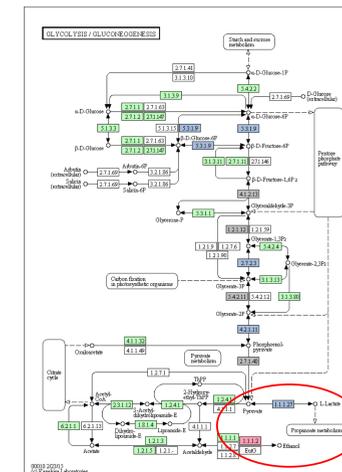
Many proteoforms of histones and 60S ribosomal proteins showed upregulation in the luminal subtype

Separation of the two BrCa subtypes



Example: Affected pathways

Warburg Effect



Green: KEGG default
Blue: up in luminal
Red: up in basal
Grey: mixed
Blank: not in human

Conclusions

- Our top-down proteomics pipeline provided significant improvement in proteome coverage, precision of the quantification, and the ability to identify statistically significant changes on proteoform abundances, in comparison to another recent report that analyzed proteoforms in the same PDX samples².
- Although still relatively limited in coverage compared to the conventional bottom-up approach, top-down analysis provides a complementary view of the proteome (e.g., truncated and/or post-translationally modified proteoforms, variants, additional protein identifications).
- The top-down, bottom-up, and peptidomics analysis of the same PDX samples showed the same degree of changes in protein/peptide abundances between the two subtypes.
- The top-down measurements provided unique information on up-regulation of many proteoforms of histones and 60S ribosomal proteins as well as key pathway level changes (e.g., glycolysis) in the luminal subtype. This holds great promise in providing new insights into cancer biology.
- Fractionation and new technologies (e.g., SLIM IMS) will greatly improve proteome coverage.

Acknowledgements

This project was funded by NIH grant U24-CA-160019 from the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). Samples were analyzed using capabilities developed under this grant and NIGMS Biomedical Technology Research Resource P41GM103493, and performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) Office of Biological and Environmental Research (OBER) national scientific user facility on the Pacific Northwest National Laboratory (PNNL) campus. PNNL is a multiprogram national laboratory operated by Battelle for the DOE under contract DE-AC05-76RL01830.

References

- Li et al., Cell Rep. 2013 Sep 26;4(6):1116-30
- Ntai et al., Mol Cell Proteomics. 2016 Jan;15(1):45-56

