High-Throughput Proteomics Platform Demonstrated for Liver Disease Progression Biomarker Verification

Erin S. Baker[1], Kristin E. Burnum[1], Daniel J. Orton[1], Yehia M. Ibrahim[1], William F. Danielson III[1] Kevin L. Crowell[1], Jon M. Jacobs[1], Matthew E. Monroe[1], Gordon W. Slysz[1], Anuj R. Shah[1], Maria L. Luna[1], Marina A. Gritsenko[1], David Purdy[2], Nathan Susnow[2], Matthew M. Yeh[2], Renuka Bhattacharya[2], James D. Perkins[2], Robert L. Carithers[2], Iris W. Liou[2], Susan Strom[2], Anne M. Larson[4], Deborah L. Diamond[2], Brian J. McMahon[3], Brenna C. Simons[3], John D. Sandoval[1], Gordon A. Anderson[1], Michael G. Katze[2], Mikhail E. Belov[1], and Richard D. Smith[1]

[1]Pacific Northwest National Laboratory, Richland, WA; [2]University of Washington, Seattle, WA; [3]Alaska Native Tribal Health Consortium, Anchorage, AK; [4]University of Texas Southwestern Medical Center, Dallas, TX

Pacific Northwest NATIONAL LABORATORY

Overview

- A well characterized patient cohort was utilized to identify protein markers in serum that discriminate between patients with non-progressing and rapidly progressing liver fibrosis
- An initial discovery set of 8 pooled serum samples was analyzed with both high resolution MS and MS/MS to identify protein variations in different liver fibrosis conditions
- Verification of the significant protein markers was performed by analyzing 120 individual patients serum samples in triplicate using a LC-IMS-QTOF MS platform
- Overall platform performance and utility for initial large scale verification studies are summarized

Introduction

Mass spectrometry (MS)-based technologies are playing a growing role in the discovery of new candidate biomarkers for human and environmental health. However, in spite of significant advances, MS technologies remain challenged by biofluid samples such as blood plasma where protein concentrations span a dynamic range >10 orders of magnitude and inadequate measurement throughput hinders the ability to effectively account for biological diversity.

In response to this shortcoming, we are developing a new proteomics platform that combines LC, ion mobility separations (IMS) and MS. By spreading ions through a range of IMS drift times so they do not all arrive simultaneously at the detector, this new platform allows the detection of much lower concentration species and the ability to use shorter LC gradient times for higher-throughput analyses (**Figure 1**) and/or greater depth of proteome coverage. Since potential candidate biomarkers are expected at low concentrations and require the examination of many samples, LC-IMS-QTOF MS has enormous potential for improving biomarker discovery and verification.

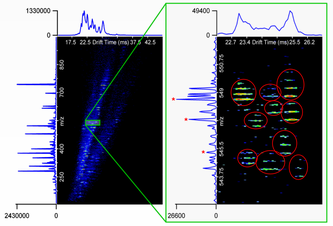


Figure 1. Selected LC elution time region from a complex sample of 20 proteins covering 10^8 dynamic range. Results show IMS-TOF MS can detect more proteins in the sample than FT MS. Ten peptides are easily detected in the zoomed region of the IMS spectra, but only 3 can be deisotoped and effectively identified by MS only (shown by *).

Methods

Experimental

- Initial discovery based sample set consisted of 60 post-liver transplant patient serum samples divided between 30 slow and rapid fibrotic progressors (progressors (P)) matched via donor age and liver cold ischemia time to 30 non-fibrotic patient controls (non-progressors (NP))
- Patient samples were pooled into 8 patient groups stratifying fibrosis progressing patients with matching controls and evaluated with a high resolution Exactive MS instrument and LTQ-MS/MS using a100-min LC gradient
- Potential liver disease progression diagnostic/therapeutic biomarkers were determined
- LC-IMS-QTOF MS with a 60-min reversed phase LC gradient was utilized for targeted verification of the determined protein biomarkers in 120 individual patients with varying fibrosis conditions
- Three technical replicates of each patient sample were blocked and randomized to reduce instrumental variation, resulting in 360 datasets

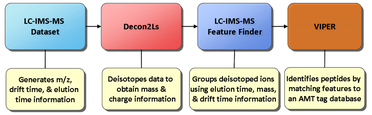


Figure 2. Data analysis flowchart

Informatics

- Data analysis flowchart is shown in **Figure 2**
- THRASH [1] algorithm implemented in Decon2LS [2] used to deisotope data
- Deisotoped ions are grouped into features utilizing drift time, monoisotopic mass and elution time information
- VIPER used to match features to sample information for identifications
- Heat maps for differentially expressed proteins were created in DAnTE using Rrollup [3] for significant peptides with p and q values less than 0.05

Results

Table1. The 29 proteins that discriminate fibrosis progression identified using an Orbitrap Exactive mass spectrometer

Protein	SP_Descriptions
THRB_HUMAN	Prothrombin
CO2_HUMAN	Complement C2
CO4B_HUMAN	Complement C4-B
TTHY_HUMAN	Transthyretin
CO4A_HUMAN	Complement C4-A
APOA4_HUMAN	Apolipoprotein A-IV
ITIH1_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H1
ANGT_HUMAN	Angiotensinogen
KLKB1_HUMAN	Plasma kallikrein
ITIH4_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H4
C1R_HUMAN	Complement C1r
HEMO_HUMAN	Hemopexin
ITIH2_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H2
A1BG_HUMAN	Alpha-1B-glycoprotein
CFAB_HUMAN	Complement factor B
ALS_HUMAN	Insulin-like growth factor-binding protein complex acid labile subunit
PROC_HUMAN	Vitamin K-dependent protein C
PLMN_HUMAN	Plasminogen
AFAM_HUMAN	Afamin
VWF_HUMAN	von Willebrand factor
FINC_HUMAN	Fibronectin
QSOX1_HUMAN	Sulfhydryl oxidase 1
A2MG_HUMAN	Alpha-2-macroglobulin
A1AT_HUMAN	Alpha-1-antitrypsin
DOPO_HUMAN	Dopamine beta-hydroxylase
CO7_HUMAN	Complement component C7
LAC6_HUMAN	Ig lambda-6 chain C region
IGKC_HUMAN	Ig kappa chain C region
LAC_HUMAN	Ig lambda chain C regions

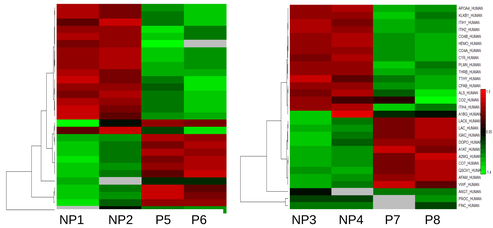


Figure 3. Pearson hierarchical clustering of normalized log2 MS intensity values of the 29 differentiating proteins showing consistency when comparing across matched non-progressor (NP) and progressor (P) patient pools

Discovery Phase

- High resolution Orbitrap Exactive MS instrument runs with a 100-min LC gradient were performed on the 8 pooled samples
- Informatics/statistical analysis identified 29 proteins that discriminate fibrosis progression (**Figure 3 and Table 1**)

Verification Phase

- 120 individual patient serum samples were analyzed in triplicate with the LC-IMS-QTOF MS platform and a 60-min gradient (**Figure 4**)
- The list of 29 discovery proteins was narrowed to 4 showing significant changes in the patients (**Figure 5**)

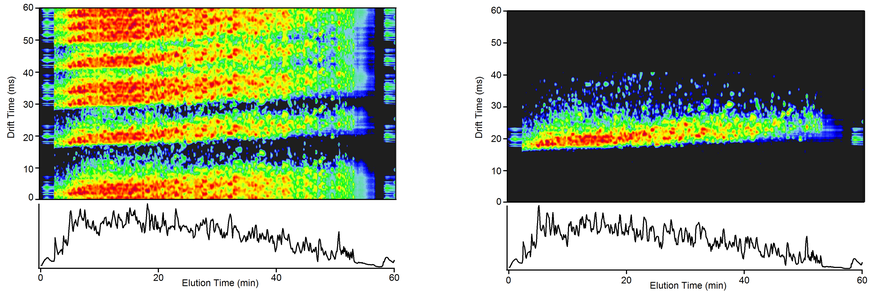


Figure 4. The 4-bit multiplexed 60-min LC-IMS-QTOF MS analysis of a typical LSDF human blood serum sample; encoded (left) and decoded (right) spectra

LC-IMS-QTOF MS Platform

High pressure automated LC system

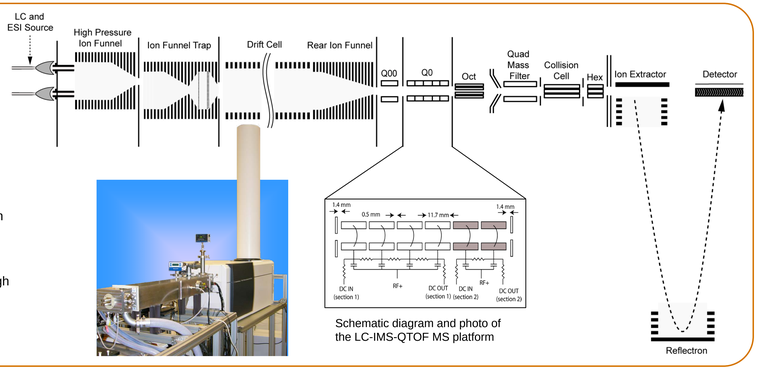
Four 30-cm long, 150 µm i.d. reversed-phase capillary columns packed in-house with porous 3 µm C18 bonded particles

Future ion source advances

- Near perfect efficiency nanoESI-MS SPIN interface
  J. Page, et al., *Anal. Chem.* 2008, *80*: 1800-1805
  I. Marginean, et al., *Anal. Chem.* 2010, *22*: 9344-9349.
- High sensitivity droplet-based microfluidics
  R. Kelly, et al., *Angew. Chem. Int. Ed.* 2009, *48*: 6832-6835

IMS-QTOF MS [4,5,6,7]

- High-pressure converging hourglass ion funnel focuses and traps ions prior to ion injection
- 1-meter IMS drift cell
- Second high-pressure ion funnel focuses the diffuse ion beam post-IMS separation
- High efficiency CID fragmentation performed in segmented quadrupole (Q0)
- Encoding was dynamically varied between signal averaging, 4-bit, 5-bit and 6-bit pseudo-random sequences; each 60-ms long IMS frame was acquired over 1 s
- Orthogonal Agilent QTOF MS provides high mass measurement accuracy after IMS separation



Schematic diagram and photo of the LC-IMS-QTOF MS platform

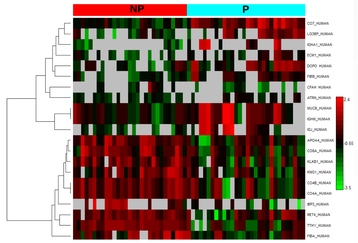


Figure 5. Clustering of normalized MS intensity values of significant proteins differentiating in individual patients in the LC-IMS-QTOF MS analyses of non-progressor (NP) versus progressor (P)

Conclusions

- **Application of an LC-IMS-QTOF MS platform allowed high throughput analysis in a large verification phase that will eventually include over 500 patients**
- **A preliminary set of 29 statistically significant peptides/proteins that discriminate fibrosis progression were identified from pooled samples in the discovery phase**
- **Analysis of 120 individual patient serum samples by LC-IMS-QTOF MS for the verification phase confirmed significant variation in 4 of the 29 statistically significant proteins**
- **Current results are also being correlated with complementary proteomic liver biopsy data and literature for mechanistic understanding of protein serum expression profiles**

Acknowledgements

This research was funded primarily though the Washington State Life Sciences Discovery fund (program ID# 2037574) and with support from the NIH National Center for Research Resources (RR 18522), the NIH National Cancer Institute (R21 CA12619-01), and the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory (PNNL). Work was performed in the EMSL, a U.S. Department of Energy/BER national scientific user facility located at Pacific Northwest National Laboratory in Richland, WA.

References

1. DM Horn, *et al.* Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules. *J Am Soc Mass Spectrom* 11:320-332 (2000).
2. N Jaitly, *et al.* Decon2LS: An Open-source Software Package for Automated Processing and Visualization of High Resolution Mass Spectrometry Data, *BMC Bioinformatics* 10:87 (2009).
3. AD Polpitiya, *et al.* DAnTE: A Statistical Tool for Quantitative Analysis of -omics Data, *Bioinformatics* 24:1556-1558 (2008).
4. ME Belov, *et al.* Dynamically Multiplexed Ion Mobility Time-of-Flight Mass Spectrometry. *Anal Chem* 80: 5873-5883 (2008).
5. K Tang, *et al.* High-Sensitivity Ion Mobility Spectrometry/Mass Spectrometry Using Electrodynamic Ion Funnel Interfaces. *Anal Chem* 77:3330-3339 (2005).
6. Y Ibrahim, *et al.* Improving Mass Spectrometer Sensitivity Using a High-Pressure Electrodynamic Ion Funnel Interface. *J Am Soc Mass Spectrom* 17:1299-1305 (2006).
7. BH Clowers, *et al.* Enhanced Ion Utilization Efficiency Using an Electrodynamic Ion Funnel Trap as an Injection Mechanism for Ion Mobility Spectrometry. *Anal Chem* 80:612-623 (2008).

CONTACT: Erin Shammel Baker, PhD
Biological Sciences Division, K8-98
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
e-mail: erin.baker@pnnl.gov