

Summary of United States Human Proteome Organisation (HUPO) Symposium Entitled “Standardized Clinical Proteomics Platforms”

Daniel C. Liebler‡, on behalf of Presenters

Studies by the National Cancer Institute (NCI)-funded Clinical Proteomic Technologies Assessment for Cancer (CPTAC)¹ program were highlighted in a United States Human Proteome Organisation (HUPO) symposium entitled “Standardized Clinical Proteomics Platforms” on February 23, 2009. The NCI CPTAC program includes the Broad Institute of MIT (Massachusetts Institute of Technology) and Harvard (with the Fred Hutchinson Cancer Research Center, Massachusetts General Hospital, the University of North Carolina at Chapel Hill, the University of Victoria, and the Plasma Proteome Institute), Memorial Sloan-Kettering Cancer Center (with the Skirball Institute at New York University), Purdue University (with Monarch Life Sciences, Indiana University, Indiana University-Purdue University Indianapolis, and the Hoosier Oncology Group), University of California, San Francisco (with the Buck Institute for Age Research, Lawrence Berkeley National Laboratory, the University of British Columbia, and the University of Texas M.D. Anderson Cancer Center), and Vanderbilt University School of Medicine (with the University of Texas M.D. Anderson Cancer Center, the University of Washington, and the University of Arizona). The CPTAC program was established to evaluate the performance of proteomics technology platforms, to develop system performance metrics, and to employ reference proteome standards to enable quantitative evaluation of analytical systems. A major goal of the program is evaluation of proteomics technologies in the context of a prototypical biomarker development pipeline. This begins with “unbiased discovery” of biomarker candidates through LC-MS/MS-based shotgun proteomics, followed by “verification” of biomarker candidates through specific, targeted assays employing LC-MRM-MS analyses.

Beginning in late 2006, the CPTAC program embarked on interlaboratory studies to evaluate proteomics platforms for unbiased discovery and targeted verification of biomarker candidates. Daniel Liebler (Vanderbilt University School of Medicine) described a series of studies by the Unbiased Dis-

covery Working Group, which employed a 20 human protein mixture and a complex proteome (*Saccharomyces cerevisiae*; yeast extract) spiked with a 48 human protein standard (Sigma UPS). The group performed a series of studies across six participating laboratories, in which a standard operating procedure (SOP) was developed to minimize variation due to analytical methods and thus enable quantitative comparisons of platform performance. Peptide detection was found to be highly variable, whereas protein detection was much more uniform across sites. The yeast reference proteome has previously been annotated by quantitative TAP-tag studies by O’Shea and colleagues (1), and this enabled quantitative comparisons of proteomics analyses between sites. Analysis of shotgun proteomic datasets enabled calculation of a “CN50” value, which denotes the yeast cellular protein copy number at which a 50% probability of detection is achieved by LC-MS/MS analysis. This parameter facilitates comparison of the depth of proteome coverage by different analysis platforms. In proteins spiked with Sigma UPS human proteins, concordance of detection efficiency across spike concentrations against the yeast background demonstrated the feasibility of multi-laboratory comparisons of protein biomarker discovery using a standardized platform.

Related work presented by Paul Rudnick (National Institute of Standards and Technology) described the development of 44 quality metrics describing the performance of multiple components of LC-MS/MS systems. These performance metrics enable documentation of system performance for proteomic analyses, as well as targeted troubleshooting, based on variances in specific metrics describing chromatography, ion source parameters, dynamic sampling, MS signal intensities, and peptide identifications. The metrics typically display variations less than 10% and thus can reveal even subtle differences in performance of system components. Analyses of data from interlaboratory studies conducted under a common SOP identified outlier data and provided clues to specific causes. These metrics could provide an unambiguous basis for distinguishing real proteomic differences from analytical platform instability by identifying outlier data using more criteria than simple peptide identifications.

Steven Hall (University of California, San Francisco) described interlaboratory studies by the Targeted Verification Working Group directed at evaluation of targeted LC-MRM-MS analyses of biomarker candidates in plasma. A

From the ‡Vanderbilt University School of Medicine, Nashville, Tennessee 37232

¹ The abbreviations used are: CPTAC, Clinical Proteomic Technologies Assessment for Cancer; LC-MS/MS, liquid chromatography-tandem mass spectrometry; LC-MRM-MS, liquid chromatography-multiple reaction monitoring mass spectrometry; MRM, multiple reaction monitoring; SOP, standard operating procedure; TAP, tandem affinity purification.

series of studies systematically evaluated sources of variation in LC-MRM-MS-based quantitation of signature peptides representing seven (7) proteins used as model biomarker candidates. Prototypic peptides from each were selected, and heavy isotopomers were used to construct stable isotope dilution MRM assays, which were performed according to an SOP. The purpose of the SOP was to minimize variation due to differences in laboratory protocol and enable estimation of variation due to instrument and analyst performance and process steps in the analytical workflow. Initial studies evaluated linearity and limit of detection and quantitation for the signature peptides. Subsequent studies employed digestion of the parent proteins in a single laboratory and finally digestion, sample preparation, and analysis in all of the participating laboratories. The data indicated the degree of variability attributed to both instrument analysis and process steps and suggested that variation in digestion efficiency for the target proteins in plasma is a major source of variability. The studies also identified other contributing factors to measurement variability and provided evidence that interlaboratory studies with standardized platforms can quantitatively estimate differences in concentrations of biomarker candidates in plasma.

Penny Drake (UCSF) described preliminary studies to evaluate targeted analysis of glycoproteins based on lectin affinity chromatography. The lectins *Sambucus nigra* (SNA) and *Aleu-*

ria aurantia (AAL), which bind sialic acid and fucose, respectively, were covalently coupled to POROS beads for high performance liquid chromatography (HPLC). Depleted plasma was trypsin-digested and separated into flow through, delayed and bound fractions by SNA or AAL HPLC. The fractions were treated with protein-N-glycanase F (PNGaseF) to remove N-linked glycans and then analyzed by LC-MS/MS. The experimental design included positive and negative controls that were used to monitor the specificity of lectin capture. Key features of this workflow include the positive identification of the captured and PNGaseF-treated glycopeptides from their deamidated Asn-XXX-Ser/Thr motifs, quality assessment using glycoprotein standards, high percentage recovery of deglycosylated peptides in lectin-captured fractions, reproducibility of deglycopeptide recovery among experimental replicates, and enrichment of low-abundance proteins through lectin chromatography. A proof-of-principle experiment using breast cancer cell line conditioned media in this workflow identified numerous cancer-relevant proteins. Future studies include biomarker discovery using plasma from individuals with breast cancer and healthy controls.

REFERENCES

1. Ghaemmighami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephore, N., O'Shea, E. K., and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737–741