REPORT

7th HUPO World Congress of Proteomics: Launching the Second Phase of the HUPO Plasma Proteome Project (PPP-2) 16–20 August 2008, Amsterdam, The Netherlands

Gilbert S. Omenn[1, 2, 3], Ruedi Aebersold[4, 5, 6] and Young-Ki Paik[7, 8]

[1] Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA
[2] Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA
[3] Center for Computational Medicine and Biology, University of Michigan, Ann Arbor, MI, USA
[4] Institute for Molecular Systems Biology, ETH Zurich, Zurich, Switzerland
[5] Faculty of Sciences, University of Zurich, Zurich, Switzerland
[6] Institute for Systems Biology, Seattle, WA, USA
[7] Yonsei Proteome Research Center, Yonsei University, Seoul, Korea
[8] Biomedical Proteome Research Center, Yonsei University, Seoul, Korea

The HUPO Plasma Proteome Project new phase, PPP-2, held its initial workshop on 17 August, 2008, at the 7th World Congress of Proteomics in Amsterdam. Technology platforms, data repositories, informatics, and engagement of research groups for the submission of major datasets were key topics. Plasma is expected to be the common pathway for biomarker development and application through collaboration and integration with other HUPO initiatives.

At the 7th World Congress of Proteomics in Amsterdam, 17th August 2008, the HUPO Plasma Proteome Project new phase, PPP-2, held its annual Workshop. The aims of the PPP-2 are (i) to stimulate submission of high-quality, large datasets of human plasma proteome findings with advanced technology platforms; (ii) to establish a robust, value-added informatics scheme involving EBI/PRIDE [1], University of Michigan/ProteomeCommons/Tranche, and Institute for Systems Biology/Peptide Atlas [2]; and (iii) to collaborate with other HUPO organ-based and disease-related initiatives to make plasma the common pathway for biomarker development and application.

After an introduction from Gil Omenn about the output from PPP-1 [3] and aims of PPP-2 [4], Ruedi Aebersold laid out a new strategy for plasma studies, emphasizing resolution and quantification deeper into the dynamic range of protein concentrations (<1 ng/ml), targeted analyses of high-interest proteins, and cooperative sharing of high-content data. He described selective or multiple reaction monitoring (SRM/MRM) of peptide ion transitions as a "mass spectrometry equivalent of multiplex ELISA" and noted the TIQAM MRM software and the PeptideAtlas from ISB as aids in designing these analyses.

The first datasets to be submitted under the PPP-2 ProteomExchange scheme developed by Lennart Martens and Eric Deutsch (www.hupo.org/PPP-2) were uploaded at PRIDE in time to be demonstrated real-time during the Workshop. This scheme provides the community the original highly annotated dataset and investigator's analyses at PRIDE [5]; distributed, secure file sharing by Tranche (www.tranche.proteomecommons.org); and automatic

Correspondence: Dr. Gilbert S. Omenn, Professor of Internal Medicine, Human Genetics and Public Health, Director, Center for Computational Medicine and Biology, University of Michigan, 2017 Palmer Commons, 100 Washtenaw Avenue, Ann Arbor, MI 48109-2218, USA
E-mail: gomenn@umich.edu
Fax: +1-734-615-6553

transfer to PeptideAtlas [6], where the raw spectra are uniformly re-analyzed with the TransProteomicPipeline to generate consolidated data on the plasma peptidome and proteome for comparative studies. Both PRIDE and PeptideAtlas are creating readily identified web pages for the PPP-2. The entire scheme is outlined in Figure 1.

Jon Jacobs of Richard Smith's group at Pacific NW Laboratory described three large datasets that were submitted: (i) 303 human N-glycoproteins, based on 63 runs on LCQ and LTQ instruments after treatment of a single normal EDTA-plasma specimen with hydrazide, trypsin, and PNGase [7]; (ii) the plasma proteome from trauma patients from the Inflammation and Host Response to Injury Program, identifying 1494 multi-peptide proteins and 662 N-glycoproteins with Mixed-12 partition and 120 runs on LTQ [8]; and (ii) a normal human EDTA-plasma proteome after Mixed-12 and Supermix depletion of 99 percent of protein mass, yielding 695 multi-peptide proteins [9]. The entire uploading process and release for public access at PRIDE was completed within a week.

Parag Mallick at the Spielberg Family Center for Applied Proteomics at Cedars Sinai Medical Center in Los Angeles, California, provided a terabyte of raw files, XML files, and related annotation of 32 experiments exploring reproducibility, sensitivity, fractionation, dynamic exclusion, selective inclusion, and gradient retention time warping. The sample is a pool from 20 healthy males, called "Bucket of Blood" (BOB). The PRIDE and Tranche steps performed smoothly. Over 5000 proteins were identified. These data had not yet been submitted for publication at the time of the Workshop. Mallick made several suggestions for enhancements of the quality assurance, metadata and software for parsing.

The third aim of the PPP-2 was addressed by leaders of other HUPO initiatives. Laura Beretta of the Fred Hutchinson Cancer Research Center, current chair of the Human Liver Proteome Project (HLPP), described the comparative analysis of liver and plasma proteomes and their changes with disease progression in two mouse models of hepatocellular carcinoma. She advocated future studies aimed at better understanding how proteome changes in diseased tissue translate into proteome changes in plasma and called for collaborative projects between HLPP and PPP. Numerous comparative studies will be desirable.

Jong Shih Yoo of the Korean Basic Science Institute described a label-free method for quantitation of pooled plasma proteins and N-glycosite peptides from normals and hepatocellular carcinoma (HCC) patients. Tadashi Yamomoto of Niigata University in Japan, chair of the Human Kidney-Urine Proteome Project, reported results available at www.hupo.org and at PRIDE for urine and for dissected glomeruli, with over 3000 proteins identified. The urine, of course, is a filtrate of plasma, altered by proteins released or secreted from the glomerular and tubular compartments of the kidney. On behalf of Weihong Yan, David Wong, and Joseph Loo at UCLA, Omenn briefly summarized results on the human salivary proteome, both whole saliva and separate parotid gland and submandibular/sublingual gland fractions [10]. Young-Ki Paik of Yonsei Proteome Research Center in Korea introduced BiomarkerDigger, which combines multiple databases with analytical functions which automate data analysis, searching, and metadata-gathering functions. BiomarkerDigger can highlight relationships between a given protein in a proteomic dataset and known biomarkers or biomarker candidates as illustrated for hepatocellular carcinoma. BiomarkerDigger was designed for use in the HUPO Plasma Proteome Project pilot phase and is being used routinely by the Paik lab for analyzing protein–protein interactions, protein networks, and functional expression information of the human plasma proteome. Finally, Omenn referred to Michigan Molecular Interactions (MiMI) (https://portal.ncibi.org/MiMI) [11]; and its Cytoscape plug-in [12] for convenient annotation of protein–protein interactions and visualization of networks.

Investigators are invited to submit major datasets to PPP-2, with raw spectra and full annotation. Notify Lennart Martens (lennart.martens@ebi.ac.uk) and Gil Omenn of your intentions. Readers are encouraged to nominate studies by others that the HUPO PPP-2 co-chairs should request. We hope many researchers will make use of these resources at PRIDE, Tranche, and PeptideAtlas to address their own biological and clinical questions.
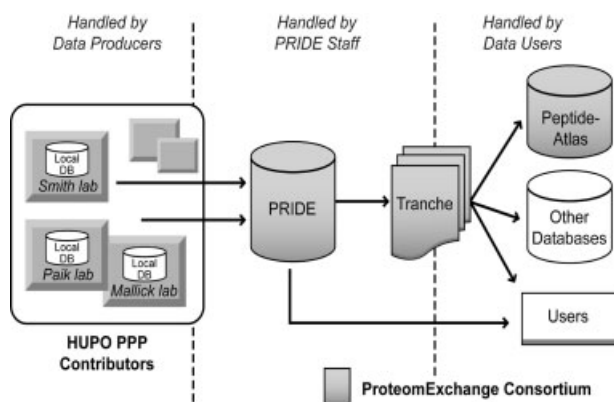

Figure 1. The ProteomExchange scheme shows the roles of investigators as data producers, the successive roles of PRIDE, Tranche, and PeptideAtlas, and the opportunity for researchers globally to utilize the original and the re-analyzed datasets from the Plasma Proteome Project (PPP-2) to explore and test their own queries. The links with other proteome studies will be particularly important.

References

[1] Jones, P., Cote, R. G., Martens, L., Quinn, A. F. *et al.*, PRIDE: a public repository of protein and peptide identifications for the proteomics community. *Nucleic Acids Res.* 2006, *34*, D659–D663.

[2] Deutsch, E. W., Eng, J. K., Zhang, H., King, N. L. *et al.*, Human Plasma PeptideAtlas. *Proteomics* 2005, *5*, 3497–3500.

[3] Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W. *et al.*, Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 2005, *5*, 3226–3245.

[4] Omenn, G. S., Aebersold, R., Paik, Y.-K., HUPO Plasma Proteome Project 2007 Workshop Report. *Mol. Cell. Proteomics* 2007, *6*, 2252–2253.

[5] Jones, P., Cote, R. G., Cho, S. Y., Klie, S. *et al.*, PRIDE: new developments and new datasets. *Nucleic Acids Res.* 2008, *36*, D878–D883.

[6] Deutsch, E. W., Lam, H., Aebersold, R., PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. *EMBO Reports* 2008, *9*, 429–434.

[7] Liu, T., Qian, W. J., Gritsenko, M. A., Camp, D. G., 2nd *et al.*, Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. *J. Proteome Res.* 2005, *4*, 2070–2080.

[8] Liu, T., Qian, W. J., Gritsenko, M. A., Xiao, W. *et al.*, High dynamic range characterization of the trauma patient plasma proteome. *Mol. Cell. Proteomics* 2006, *5*, 1899–1913.

[9] Qian, W. J., Kaleta, D. T., Petritis, B. O., Jiang, H. *et al.*, Enhanced detection of low abundant human plasma proteins using a tandem IgY12-supermix immunoaffinity separation strategy. *Mol. Cell. Proteomics* 2008, *7*, 1963–1973.

[10] Yan, W., Apweiler, R., Balgley, B.M., Boontheung, P. *et al.*, Systemic comparison of the human saliva and plasma proteomes. *Proteomics Clin. Appl.* 2009, *3*, DOI: 10.1002/prca.200800140.

[11] Jayapandian, M., Chapman, A., Tarcea, V. G., Yu, C. *et al.*, Michigan Molecular Interactions (MiMI): putting the jigsaw puzzle together. *Nucleic Acids Res.* 2007, *35*, D566–D571.

[12] Gao, J., Ade, A. S., Tarcea, V. G., Weymouth, T. E. *et al.*, Integrating and Annotating the Interactome using the MiMI plugin for Cytoscape. *Bioinformatics* 2008, DOI: 10.1093/bioinformatics/btn501.