Published in final edited form as:

Proteomics. 2009 November ; 9(21): 4830–4833. doi:10.1002/pmic.200990080.

The Human Liver Proteome Project (HLPP) Workshop August 2008, Amsterdam, The Netherlands

Laura Beretta

Fred Hutchinson Cancer Research Center, Seattle

Abstract

An HLPP workshop was held in conjunction with the 7th HUPO 2008 World Congress in Amsterdam, The Netherlands. The workshop was chaired by Laura Beretta (Fred Hutchinson Cancer Research Center (FHCRC)) and Xiaohong Qian (Beijing Proteome Research Center (BPRC)). The workshop was organized in three sessions: Session 1: The Human Proteome Project: the liver perspective, moderated by John Bergeron (McGill University); Session 2: The comparative analysis of the liver and plasma proteomes, moderated by Young-Ki Paik (Yonsei University) and Session 3: Opportunities for the study of liver diseases within the HLPP, moderated by Chantal Housset (INSERM, France).

Session 1: The Human Proteome Project: the liver perspective

The topics of discussion in this session included the following: What coverage of the human genome has been reached by the analysis of the liver proteome using mass spectrometry (MS)-based approaches and similarly using antibody (Ab)-based approaches?; How do these MS-generated and Ab-generated data integrate and complement each other?; What did we learn about localization, abundance and liver tissue specificity?; How to define pilot projects around the liver that would contribute to the Human Proteome Project?; Should a mouse proteome project be selected as a first milestone?.

Laura Beretta presented a first analysis of the human genome coverage obtained by extensive mass spectrometry analysis performed by her group, of two HLPP reference samples: the French Reference liver sample and the Huh7 hepatoma cell line. This analysis was performed in collaboration with Eric Deutsch at the Institute for Systems Biology, Seattle. All MS data obtained by the Beretta group at FHCRC on both HLPP liver reference samples were processed by Eric Deutsch in PeptideAtlas (www.peptideatlas.org). A total of 2 342 395 spectra were searched and 609 269 spectra were identified with a FDR of 5% (554 711 for FDR of 1%) corresponding to 50 611 unique peptides (44 710 for FDR of 1%). These peptides were mapped to 8039 distinct Ensembl gene IDs, 45% of them common to both the French liver reference sample and the hepatoma cell line. Coverage per chromosome was presented with a particularly high coverage of chromosomes 2, 14, 18 and 22 and an overall coverage of 37%. From the 2585 genes described in Protein Atlas (www.proteinatlas.org), 2353 were found expressed in liver tissue and/or in the HepG2 cell line. Integration of the MS-generated and Ab-generated expression data increased the human genome coverage to 42%. A similar analysis was performed on the mouse genome and liver proteome. Three mouse liver data sets were integrated within PeptideAtlas: a data set

generated by Matthias Mann (Max-Planck Institute for Biochemistry) using density gradient for protein separation, a data set generated by Andrew Emili (University of Toronto) using an organelle approach and a data set generated by Laura Beretta using 2D-HPLC followed by SDS PAGE for protein separation. Over 4 million spectra were searched and 980 824 spectra were identified with a FDR of 1% corresponding to 30 755 unique peptides. These peptides were mapped to 5716 distinct Ensembl gene IDs, 34% of them common to all three data sets, 65% common to at least 2 data sets and 35% unique to 1 data set. Each protein separation method provided a significant number of unique protein identifications and the overall coverage of the mouse genome reached 24%.

John Bergeron gave a presentation on the Liver cell map. He showed using Principal Component Analysis that the pattern of protein abundance measured directly by mass spectrometry distinguishes the RM and SM fractions and the Golgi. This analysis also allows for the evaluation of the degree of cross-contamination of the isolated subcellular fractions. A fraction that falls closer to another is likely contaminated with proteins from that fraction. Most importantly, this analysis demonstrated that the variability within different preparations of subcellular fractions is much smaller than the difference among the organelles. John Bergeron further described how proteotypic peptides can be used to characterize isoforms of important ER protein families. This approach was applied to 34 members of the cytochrome P450 family, identifying 2 novel isoforms of cytochrome P450 2c11, distinct by a single amino acid. A similar approach was applied to 11 glucuronosyltransferases and 9 carboxylesterases.

Bill Jordan (Victoria University of Wellington) presented the progress of the Asia Oceania HUPO Membrane Proteome Initiative (MPI). The goals of this initiative are to assess the ability of technologies to profile membrane proteomes and the extent to which membrane proteomes are comprehensively profiled. The approach is to isolate mouse liver microsomes, further carbonate washed to prepare the “MPI standard” enriched for TMD proteins and distributed in 2006/2007 to participating laboratories to apply 7 different protein and peptide separation techniques to the sample. Raw MS files were distributed to four bioinformatics groups. Transmembrane domain prediction was performed on the identified proteins. Twenty-nine cytochrome P450s were characterized.

Finally, Paul Boersema (Utrecht University) described how to apply stable isotope dimethyl labeling (labeling of primary amines on N-terminus and lysine residues) and protein quantitation to the study of protein complexes diversity between tissues. He applied this quantitative proteomics approach to the analysis of bovine liver and spleen 20S proteasome core, identifying compositions specific to the liver and to the spleen.

Session 2: Comparative analysis of the liver and plasma proteomes

The comparative analysis of the liver and plasma proteomes using 2 mouse models of liver cancer was presented by Laura Beretta. The proteome of liver tissue and corresponding plasma were obtained from healthy mice, from mice with liver fibrosis or steatohepatitis and from mice with hepatocellular carcinoma. An average of 20% of the proteins identified in liver tissue was also detected in the plasma. For these proteins, there was no correlation

between abundance in liver tissue and abundance in the plasma. In addition, changes in protein amount associated with liver disease progression in these two models did not translate into similar changes in plasma for a majority of these proteins.

Young-Ki Paik presented BiomarkerDigger, a proteome database recently developed by his group. The metadata-gathering function searches available proteome databases for protein–protein interactions (PPIs), Gene Ontology (GO), protein domain, Online Mendelian Inheritance in Man (OMIMTM), and tissue expression profile information and integrates it into protein data set profiles that are accessible by search functions in BiomarkerDigger. This database also facilitates cross-proteome comparisons by classifying proteins in the proteome on the basis of their annotation. BiomarkerDigger can therefore highlight relationships between a given protein in a proteomic data set and known biomarkers or biomarker candidates.

Session 3: Opportunities for the study of liver diseases within the HLPP

The topics covered in the final session dedicated to liver diseases were liver disease biomarkers with a specific focus on liver cancer, biobanks, mouse models and data repositories.

Chantal Housset presented an update on the biobanks established in France as well as the currently changing trends in the incidence of the different liver cancer forms. Hepatocellular carcinoma (HCC) remains one of the most common cancers worldwide and it is the third most common cause of cancer mortality. In France, a national biobank of HCC was initiated in 2001. Human biological samples are annotated, collected/stored/distributed under high quality-assurance procedures, according to international standards and ethical rules. This biobank is based on the OECD recommendations (Organisation for Economic Cooperation and Development) for Biological Resource Centres (BRC), which provides the guidelines for the certification of BRCs by national institutions. BRCs fulfill the “International declaration on human genetic data”, which was released by UNESCO in 2003. The biobank relies on a network of centers that share the same assurance-quality procedures for sample and data collection. Tumour and adjacent-non-tumour liver samples obtained in the course of surgical treatments are frozen and stored in each centre. Annotations related to demographics, clinical, biological and histological data are centralized within a remote unique database. The French Liver Cancer Bank is sponsored by the French national biomedical research agency INSERM, the National Institute of Cancer INCa, the Ministry of Health and the Ministry of Research. The cost estimate for eight collecting centres is 800 000€ The French Liver Cancer Bank is also a member of the European biobanks network initiative, sponsored by the European Commission. As of March 2008, the network has successfully collected over a thousand of liver samples from almost 900 patients associated with standardized clinical/laboratory and therapeutic information and informed consents. The samples have been used in basic, clinic and translational research, leading to 26 publications in the period 2004–2008. A major goal for the future is to strengthen connections with other national networks involved in similar programmes. Other goals include the collection of other forms of liver cancer, cholangiocarcinoma in particular. The increase in the incidence of cholangiocarcinomas is spectacular in France, as in many

western countries. Other biological samples such as liver samples in chronic liver diseases, liver cell types, serum and bile will also be included in the biobank.

Juan Falcon-Perez (CIC bioGUNE, Spain), collaborator of José Mato, co-chair of HLPP, gave an overview of mouse models of non-alcoholic fatty liver disease (NAFLD). This clinical term includes a broad spectrum of liver alterations ranging from a simple accumulation of triglycerides inside the hepatocyte to steatosis with different grades of inflammation and collagen deposition. This state may progress to more severe conditions such as cirrhosis and hepatocellular carcinoma. About eight million people in the European Union may develop non-alcoholic steatohepatitis (NASH). There are several mouse models that resemble the clinical manifestations of human NAFLD and NASH. They are either diet-induced or genetically induced, although a combination of both methods is often used to further resemble the human pathology. José Mato and colleagues have developed two mouse models of NAFLD/NASH by genetic modification of the methionine metabolism in the liver, a pathway central to the generation of S-adenosylmethionine, which is the main methyl donor in the methylation processes of proteins, lipids and DNA, and known for over seven decades to play an important role in NAFLD development. These mice are deficient in two liver-specific enzymes involved in the homeostasis of S-adenosylmethionine, MAT1A involved in the synthesis of this metabolite and the enzyme GNMT involved in elimination of the excess of this metabolite. These two mouse models represent powerful tools to identify biomarkers for the early diagnosis and prognosis of the disease. Collaborative studies to characterize their metabolome and proteome are ongoing. Juan Falcon-Perez is characterizing in particular the proteome of exosomes, a subcellular organelle particularly relevant to biomarker discovery studies.

Songfeng Wu (Beijing Proteome Research Center), collaborator of Fuchu He, co-chair of HLPP, introduced to the audience dbLEP, a database of liver proteome expression profiles that currently includes three data sets generated by participating laboratories in China: Chinese fetal liver with 2495 distinct proteins (898 high confidence proteins); French adult liver with 2570 high confidence proteins; and the Chinese adult liver with 6788 high confidence proteins. He also presented the progress made on the construction of the human liver protein linkage maps. Using yeast 2-hybrid, 3569 human liver protein-protein interactions (PPIs) were detected between 1300 baits and 1500 prey. The largest network includes 810 proteins and 1035 interactions. Validation is performed using co-immunoprecipitation and pull-down assays as well as PRINCESS: a PRotein Interaction Confidence Evaluation System with multiple data Sources. dbLIP, a database of human liver interacting proteins, contains 200 PPIs from 180 proteins expressed in the human fetal liver and 3405 PPIs from 2881 proteins expression in adult liver.

Finally, Laura Beretta brought to the attention of the workshop attendees three new initiatives focusing on liver cancer. The International Liver Cancer Association (ILCA) (<http://www.ilca-online.org>), founded in 2007, aims at advancing research in the pathogenesis, prevention and treatment of liver cancer. The International Liver Cancer Study (ILCS) also launched in 2007, is a network of clinicians, epidemiologists and molecular biologists coordinated by the International Agency for Research on Cancer (IARC) and its objectives are to improve understanding of the geographic, pathological, virological and

molecular diversity of the disease, as well as the identification of reliable biomarkers for risk assessment, early detection and monitoring of the disease. Finally, the International Cancer Biomarker Consortium (ICBC) (http://www.fhcrc.org/science/international_biomarker) launched by Dr. Lee Hartwell, Director of FHCRC, aims at establishing a collaborative network of research teams to both improve the process of biomarker discovery and to discover biomarkers for common types of cancer. Teams from the Fred Hutchinson Cancer Research Center, the Yonsei Proteome Research Center, the Korea Research Institute for Bioscience & Biotechnology, the Beijing Proteome Research Center and the University of Hong Kong are focusing their study on liver cancer biomarkers.

We acknowledge the important contribution of the HLPP co-chairs Fuchu He and José Mato who unfortunately were not able to attend the workshop but were represented by colleagues who presented their most updated progress.