Chromosome-centric Human Proteome Project
NEWSLETTER
No. 6
April 25, 2017

INSIDE

Editorial : What's NeXt After The neXt-50 Challenge?

Known Knowns, Known Unknowns, and Unknown Unknowns

Dr. Christopher Overall
Canada Research Chair in Protease Proteomics
and Systems Biology, University of British Columbia, Canada
Co-Chair, C-HPP Consortium: EC Member, HUPO

17,008... 17,008 PE1 proteins in NeXtprot now known!

A significant milestone is achieved, with more heralded. These 17,000 of the 20,159 human protein entries in NeXtprot are what the former U.S. Secretary of Defense, Donald Rumsfeld, might designate as "known knowns". This leaves 3,151 (PE2+3+4+5) "Known Unknowns" of which 2,579 are designated as Missing Proteins (PE2+3+4), but how many "Unknown Unknowns" lurk in the human proteome? Increasingly higher mass accuracy and faster instrumentation, deeper coverage proteomic analyses, and novel multi-omic approaches will render the shy known unknown missing proteins (MPs), and eventually even the very shy, increasingly better known. Both targeted MP searches and unbiased deep approaches, e.g. proteomic analyses of underexplored human cells and tissues, are ongoing by multiple Chromosome Groups aiming to each get to know better ~50 MPs. This is an integral goal of **the neXt-50 Challenge**—itself an interim goal in achieving an accurate first draft of the complete human proteome. Thus, **the neXt-50 Challenge** of the CHPP aims to introduce ~1000 found proteins—promoting known unknowns to known knowns, by HUPO-2018.

Achieving the goal of the HPP draws nearer… or does it?

As the human genome has taught us, nature has a talent for surprise in a multitude of unexpected ways to generate genome and proteome diversity, and ultimately human phenotypes with varying shades of disease susceptibility. Open reading frames (ORFs) are simple—we already know their cognate protein knowns and, by subtraction, the known unknowns in our proteome. Yet proteoforms constitute a huge number of protein variants, many of which have altered activity or even entirely new functions, and hence biological relevance, in homeostatic or pathological processes. Therefore, many proteoforms will form mechanistically informative biomarkers with higher sensitivity and selectivity for active vs. stable disease or remission states than mere constellations of tryptic peptides (Eckhard et al 2016). Clearly these proteoforms must be further explored and annotated to achieve deeper functional understanding of the human proteome. This endeavor should become more integral to the mission of the C-HPP. In addition, a relatively unexplored benefit in enriching for post-translational modified peptides is that these peptides will have different, yet often more favorable, ionization and fragmentation properties versus the parent tryptic peptide. This can be exploited to orthogonally know shy MPs that current analysis of regular tryptic peptides and MP searches tend to overlook as a resource for MPs (Eckhard et al 2016).

Semi-tryptic peptides generated by precision proteolysis exemplify this concept. For some protein substrates, cleavage by proteolytic processing (as opposed to degradation) removes peptide portions that may compromise ionization, fragmentation, or unambiguous mass spectrometric identification. Thus, apart for the intrinsic biological worth in identifying PTM-modified peptides of their parent proteoforms and characterizing the altered biological properties of the parent protein, such analyses will undoubtedly uncover recalcitrant MPs (Eckhard et al 2016). Terminomics by our TAILS approach, amongst others, reveals rich diversity in human proteolytic proteoforms. The Pandey Lab, amongst others, has an increasing collection of nonconventional protein translation start sites identified that promises to expand further the human proteome proteoform and protein repertoires. The TopFIND database annotates >330,000 protein termini in different species revealing the unexpected extent of this PTM.

But what of other unknown unknowns? Unexpected protein coding LncRNA and small open reading frame (smORF) protein sequences threaten to balloon protein numbers in the human proteome dramatically. Alan Saghatelian revealed at the C-HPP Scientific Workshop Session of the HPP at Taipei-2016 several hundred new small open reading frame (smORF) sequences that his lab identified by techniques such as RiboSeq, with a small number of their encoded proteins unambiguously identified and functionally characterized; many more are set for validation by targeted proteomics. The Roucou Lab tentatively claims that upwards of 24,000 smORFs are in fact present in humans. LncRNA and smORFs are very unexpected, unknown unknowns, potentially adding rich complexity, diversity and protein numbers to the human proteome upon validation. What other unexpected gene and transcript forms or whole new Gene/RNA classes lie unknown, waiting for recognition, identification, deciphering and characterization of their translated proteins or mere snippets of proteins?

Currently, the HPP goals center around the known knowns and the known unknowns. However, I know the unknown unknowns will be dramatic, breath-catching biological additions to our ever-expanding proteome—promising new keys for deciphering human phenotypes from what, not so long ago, was perhaps considered a limited human genome of just over 20,000 protein encoding genes and their protein products. So, not only

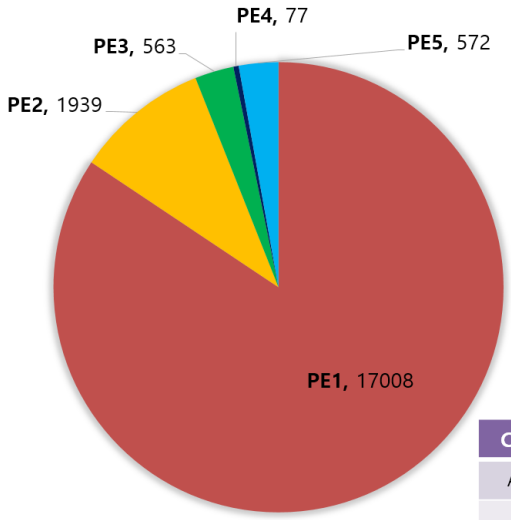
must we explore the less well known and unknown proteins, but even more, we must be open for very unexpected encounters with unknown unknowns.

The HPP goal is noble, achievable and significant. Let us hope the HPP goals and enthusiasm expands to encompass new discovery to find unknown unknowns and make them known. Whereas the known unknowns herald new biology in development, growth, homeostasis and disease—many potentially also bring to medicine new drug targets, but it is the unknown unknowns that I know will be as exciting or even more exciting. So Lets Go!

Reference

- Eckhard, U., Marino, G., Butler, G.S., and Overall, C.M. 2016. Positional Proteomics in the Era of the Human Proteome Project on the Doorstep of Precision Medicine. Biochimie 122, 110-118
- Eckhard, U., Marino, G., Abbey, S.R., Tharmarajah G., Matthew, I., and Overall, C.M. 2015. The Human Dental Pulp Proteome and N-terminome: Levering the Unexplored Potential of Semi-tryptic Peptides Enriched by TAILS to Identify Missing Proteins in the Human Proteome Project in Underexplored Tissues. Journal of Proteome Research 14, 3,568-3,582.

The neXtProt Status of Human Proteome



Missing Proteins (PE2+PE3+PE4):
☞ **2,579**

- ***PE1*** *: Evidence at protein level*
- ***PE2*** *: Evidence at transcript level*
- ***PE3*** *: Inferred from homology*
- ***PE4*** *: Predicted*
- ***PE5*** *: Uncertain*

Chr	PE1	PE2	PE3	PE4	PE5	Total
All	17,008	1,939	563	77	572	20,159
	84.37%	9.62%	2.79%	0.38%	2.84%	100%

Update: 2017-04-19 (neXtProt release : 2017-01-23)
https://www.nextprot.org/about/protein-existence

※ The official HUPO/HPP count of predicted protein-coding genes (not even known proteins), embracing neXtProt PE1-4, is now 17,008 + 2579 = **19,587**. HUPO HPP EC decided to exclude PE5, while leaving open the possibility that new evidence will justify making one or more of those "dubious or uncertain genes" a candidate for curation as PE1-4. See https://hupo.org/

C-HPP Leadership Update

C-HPP Executive Committee (EC) Composition

At Sun Moon Lake, C-HPP PIC members unanimously reelected three EC members: Young-Ki Paik (Chair), Peter Horvatovich (Secretary General), and Fernando Corrales (MAL). Their new terms began on January 1, 2017, and end on December 31, 2019. In March, C-HPP also elected Ping Xu (BPRC, China) and Gilberto B. Domont (Universidade Federal do Rio de Janeiro, Brazil) as new EC members. They will fill the positions previously occupied by Fuchu He (China) and Daniel Figeys (Canada) and will serve from March 1, 2017, to December 31, 2018. All seven members of the EC have now been appointed.

Young-Ki Paik

Chair

Lydie Lane

Co-Chair

Christopher Overall

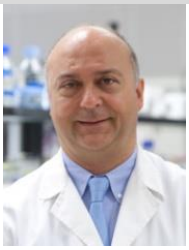
Co-Chair

Peter Horvatovich

Secretary General

Gilberto B. Domont

MAL



Fernando Corrales

MAL

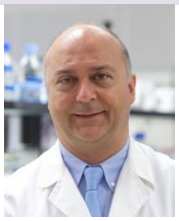
Ping Xu

MAL

Position	Name	Affiliation	Term Ends	Remarks
Chair	Young-Ki Paik (Asia Oceania)	Yonsei Univ., Seoul, Korea	Dec 31, 2019	PI, Chr 13
Co-Chairs	Lydie Lane (Europe)	SIB, Univ. of Geneva, Switzerland	Dec 31, 2017	PI, Chr 2
	Christopher Overall (America)	UBC, Vancouver, Canada.	Dec 31, 2018	PI, Chr 6
Secretary General	Peter Horvatovich (Europe)	Univ. Groningen, Groningen, Netherlands	Dec 31, 2019	PI, Chr 5 WiKi Manager
Members -at-Large	Gilberto B. Domont (America)	Federal University of Rio de Jeneiro, Brazil	Dec 31, 2018	PI, Chr 15
	Fernando Corrales (Europe)	CIMA, University of Navarra, Spain	Dec 31, 2019	PI, Chr 16
	Ping Xu (Asia Oceania)	Beijing Proteome Research Center (BPRC), China	Dec 31, 2018	PI, Chr 1

C-HPP Principal Investigators Council (PIC)

The current members of PIC are listed below. Note that PIs of Chr 3, 6, 19, 21 and X are newly joined the C-HPP in 2017. (see also next page for details).

Chr. 1	Chr. 2	Chr. 3	Chr. 4	Chr. 5	Chr. 6	Chr. 7
						
Ping Xu	**Lydie Lane**	**Takeshi Kawamura**	**Yu-Ju Chen**	**Peter Horvatovich**	**Christoph H. Borchers**	**Edouard Nice**
Chr. 8	**Chr. 9**	**Chr. 10**	**Chr. 11**	**Chr. 12**	**Chr. 13**	**Chr. 14**
						
Pengyuan Yang	**Je-Yoel Cho**	**Joshua Labaer**	**Jong Shin Yoo**	**Ravi Sirdeshmukh**	**Young-Ki Paik**	**Charles Pineau**
Chr. 15	**Chr. 16**	**Chr. 17**	**Chr. 18**	**Chr. 19**	**Chr. 20**	**Chr. 21**
						
Gilberto B. Domont	**Fernando J. Corrales**	**Gilbert S. Omenn**	**Alexander Archakov**	**Sergio Encarnación-Guevara**	**Siqi Liu**	**Albert Sickmann**
Chr. 22	**Chr. X**	**Chr. Y**	**Mitochondria**			
						
Akhilesh Pandey	**Yasushi Ishihama**	**Ghasem Hosseini Salekdeh**	**Andrea Urbani**			

A Brief Introduction of New Principal Investigators and Their Research Plans

In 2017, new five PIs recently joined us to take leadership of Chr 3 (Takeshi Kawamura, Japan), Chr 6 (Christoph Borchers, Canada), Chr 19 (Sergio Encarcion-Guevara, Mexico), Chr 21 (Albert Sickmann, Germany), and Chr X (Yasushi Ishihama, Japan). We warmly welcome them to our consortium. A brief introduction of their teams follows based on the information provided by each PI. As our project continues, unexpected challenges may await, but we believe there is a rewarding future for our C-HPP endeavor, which now becomes a project of national pride for each country. As stated in our 2016 JPR Editorials, we believe that by collaborating with our PIs and B/D-HPP colleagues, the pace of the C-HPP project will improve considerably. In this regard, we trust that our new PIs will bring new blood and resources to the consortium and to the proteomics community. Let's welcome them with great hope!

Chromosome 3

Takeshi Kawamura

- Associate Professor
- Proteomics Laboratory, Isotope Science Center, The University of Tokyo, Tokyo, Japan. 2-11-16 Yayoi, Bunkyo-Ku, Tokyo 113-0032 Japan.
- TEL/FAX: +81-3-5841-2872

Co-PI:

- Toshihide Nishimura (Department of Translational Medicine Informatics, St. Marianna University School of Medicine)
- Hiromasa Tojo (Department of Biophysics and Biochemistry, , Osaka University)

International Collaborative Team Member:

- György Marko-Varga and Siu Kwan Sze

Plans

PI has been changed from Prof. Nishimura to Kawamura in 2017 and Prof. Nishimura will continue to contribute the project as Co-PI. The takeover is ongoing. A short-term plan is searching missing proteins by continuing the previous proteomics for formalin-fixed, paraffin-embedded, and laser-microdissected lung adenocarcinoma cells of the early lepidic-types.

To search missing proteins, we have developed SQL-based in-house codes, which remove the effects of redundant peptides on FDR estimation. Non-unique proteins were replaced by the representative proteins with the codes using the following order of priority: 1, A protein containing a unique peptide with the highest protein evidence and the alphanumerically first entry name; 2. A reviewed protein with the highest protein evidence and the alphanumerically first entry name; 3, A un-reviewed proteins with the highest protein evidence and the longest amino acids.

We have re-analyzed our previously published data with the codes and found four proteins that have not been listed in the NextProt and PeptideAtlas. Our plan is to expand this method to other disease samples to find more missing proteins. The long-term plan is a new project that combines this method with our epiproteomics analysis to reveal relationships between chromosome-consisting protein histones and genes.

Chromosome 6

Christoph H. Borchers

- University of Victoria, Victoria, Canada
- McGill University, Montreal, Canada

Co-PI:

- Yassene Mohammed

Plans

Mining for human missing proteins in dissected mouse tissues

Mice are the most commonly used human model organism for biology and medicine. In addition to accessibility, around 90% of their genes are closely related to those in humans, making them a perfect tool for studying biology and diseases. In one project our goal is to develop mouse-specific MRM assays for 20 different tissues. Besides using these assays to answer various biological and disease-related questions, these assays also allows us to mine the proteome for missing human proteins. Especially because it is ethically inappropriate to collect and use human tissues for discovery analyses, using mouse tissue to detect homologs of the missing human proteins is a very promising approach.

In the first discovery phase, we performed analyses using high-resolution mass spectrometry on 40 different mouse tissues, with the goal of developing MRM assays using peptides that are shared between mouse and human. We have already identified 23 mouse proteins in various tissues corresponding to missing human proteins from chromosome 6 using this high resolution data. The majority of the proteins were found in only 2 out of the 40 different tissues (Table 1). Expression of these missing chromosome 6 proteins seems to be concentrated in bone tissue and marrow, as well as in ovaries and the nervous system.

While this high-resolution data provides the first indication of the presence of these proteins, it is our goal to verify these results with targeted MRM-based proteomics, using synthetic heavy labeled peptides which is the "gold standard" for the quantitation. We will also develop MRM assays for any predicted chromosome-6 proteins which could not be detected by high resolution MS, because these targeted assays are more sensitive than shotgun approaches, particularly if 2D LC-MRM/MS is use. For both groups of proteins, the MRM approach will provide the best information on tissue expression levels and in which human tissue the subsequent verification should be performed. Additionally, while our current results are related to Chromosome 6, this method is applicable to all missing human proteins as long as there is a mouse homolog.

Chromosome 19

Sergio Encarnación-Guevara

Co-PI: Jeovanis Gil Valdés, Julio Collado, Rafael Palacios Alberto Checa-Rojas.

Partners: Fernando J. Corrales. ProteoRed, Spain.

Plans

In the first stage the 236 MPs in the chromosome 19 will be analyzed in silico to determine the feasibility of their being unambiguously identified by MS experiments. All positive candidates will be selected to determine the cell lines and conditions of their expression at the transcription level. We will use chemical inhibition and siRNA silencing of epigenetic enzymes such as lysine deacetylases (KDACs) and histone and DNA methyltransferases to search for appropriate expression conditions of MPs. The next stage will involve large scale proteomics analysis to identify and validate the presence of the MPs using synthetic peptides. To go deeper into the characterization of these proteins, we will perform protein-protein interaction experiments and lysine acetylation analysis to search for potential targets of this PTM in the MPs. Some of these proteins will be selected to study their role in cancer cells. They will be transfected in cell lines with expression vectors and specific tags. The interacting proteins will be identified by MS/MS analysis.

Chromosome X

Yasushi Ishihama

- Department of Molecular & Cellular BioAnalysis Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshidashimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan
- Phone: +81-75-753-4555 (office), Fax: +81-75-753-4601,
  E-mail: yishiham@pharm.kyoto-u.ac.jp

Co-PI:

- Tadashi Yamamoto (Niigata University)
- Takeshi Tomonaga (National Institute of Biomedical Innovation, Health and Nutrition)

Plans

Currently we are working together with jPOST (Japan ProteOme STandard Repository/Database) team to develop a proteomic database for integration of proteome datasets generated from multiple projects and institutions. Researchers can now upload their proteome datasets to the jPOST repository through ProteomeXchange system, and the raw MS data is re-processed using the jPOST standard protocol to reduce both false positive and false negative hits. Finally, peer researchers can access to preset databases or can create customized queries including cHPP-focused missing protein databases.

Chromosome 21

Albert Sickmann

- Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V. Bunsen-Kirchhoff-Straße 11 44139 Dortmund
- TEL: +49 (0)231 1392-0, FAX: +49 (0)231 1392-200, E-Mail: info@isas.de
- Homepage: www.isas.de

Co-PI:

- Andreas Roos (The John Walton Muscular Dystrophy Research Centre, Newcastle University, United Kingdom)
- Robert Ahrends (Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V. Dortmund, Germany)

Plans

- Chromosome 21 and disease genes: Homocystinuria (CBS), Alzheimer's disease (APP), Leukaemia (AML1), Amyotrophic lateral sclerosis (SOD1), Autoimmune polyglandular disease (AIRE), Progressive myoclonus epilepsy (CSTB)
- Selected syndromes: Down Syndrome, Usher syndrome 1E, Knobloch Syndrome, Leucocyte adhesion deficiency, Betlehem Myopathy, Ullrich Myopathy
- Next steps:
  1. Detailed look at the data (Check original publication and Nextprot)
  2. Proteomics data for 18 genes available (PE2 to PE5)
  3. Definition of MRM / PRM assays for missing proteins
  4. Definition of sample material
  5. Stepwise validation of proteins
- Sample material which should be analyzed:
  1. Access to specimen from Down Syndrome patients
  2. Access to skin biopsy material
  3. Access to specimen from Mouse
     Homology to proteins from chr. 10, 16 and 17.
     Comparison with a dataset of about 12000
     proteins from 42 mouse tissues.
  4. Probing for missing proteins with SIL based MRM / PRM assays.
  5. Development of a Chromosome 21 assay (Whole chromosome assay)

Co-PI for Chr 22: Dr. Min-Sik Kim at Kyunghee Univ. Korea has been appointed as Co-PI for Chr 22 group with full support of Akhilesh Pandey (PI). We anticipate that Min-Sik will do some active roles in liaising between Chr 22 and the rest of Chr teams. He is the first author of 2014 Nature Paper (from Pandey Lab).

JPR HPP Special Issue (SI) in 2017

Given the joint efforts between the two teams, we now have a great opportunity to stimulate both C-HPP and B/D-HPP initiative groups by jointly publishing JPR SI in 2017. The contents in this joint SI will include more in-depth studies of missing proteins, biology, diseases, and technology developments. The deadline for manuscript submission is May 30, 2017.

Submission Procedure

Manuscripts must be submitted by May 30th, 2017 to be considered for this Special Issue. Manuscripts must be submitted electronically through the ACS Paragon Plus Environment online submission system. Specify in the authors' cover letter that the manuscript is intended for the HPP Special Issue.

Review and Publication Process

Initial editorial review will determine whether manuscripts are appropriate for the HPP Special Issue. Those falling outside of the scope of the Special Issue may be considered for a regular issue of the journal. Completion and inclusion of the HPP checklist (for more information: http://www.thehpp.org/guidelines/) and a full MS data submission to ProteomeXchange *must* be completed prior to initial submission and the PXD number provided in the abstract. Nonconforming papers will be returned without reviews. All relevant papers will go through full peer review. As papers are accepted they will go online and be available in time for HUPO-2017 in Dublin. Only papers accepted by September 1, 2017 will be published in this December 2017 HPP Special Issue. Papers requiring more time for revision will be considered for later regular issues.

HPP Data Guidelines

Papers must conform to both the *Journal of Proteome Research* mass spectrometry guidelines and the HPP guidelines v 2.1 (see Deutsch et al.) in order to be sent to review and for acceptance. Please check for any changes to the HPP guidelines before submission. All papers must analyze their data using the Human PeptideAtlas release 2017-01 and neXtProt release 2017-02, papers not doing so will be returned without review.

Journal of proteome research · ACS Publications Most Trusted. Most Cited. Most Read.

Human Proteome Project · C-HPP

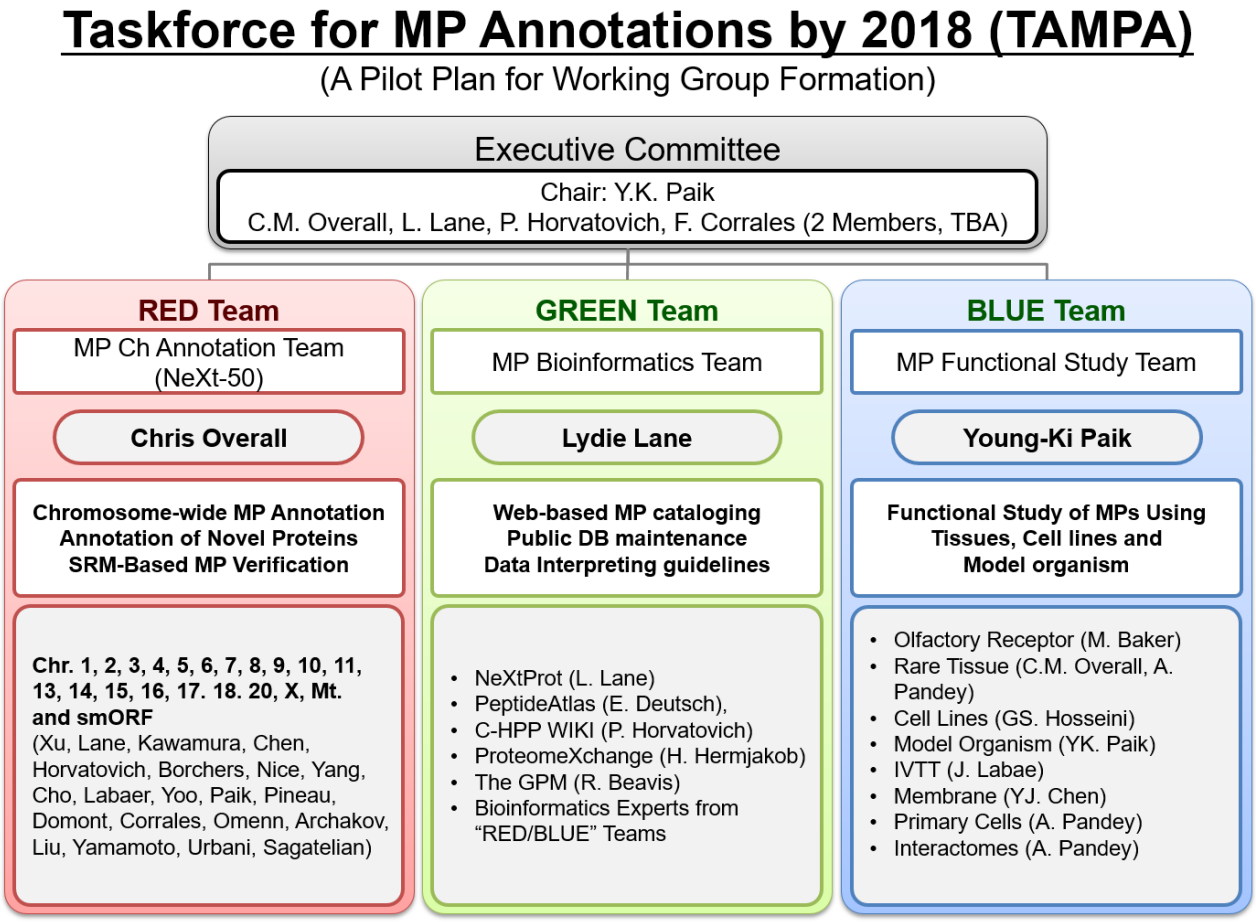
C-HPP Special Issue Editors

Associate Editor	Guest Editors			
				
Christopher M. Overall	**Young-Ki Paik**	**Gilbert S. Omenn**	**Eric Deutsch**	**Jennifer Van Eyk**

Working Group Formation of C-HPP

Background for reshaping working group: The HUPO and HPP leadership have suggested that annotation work on the missing proteins (MPs) should be expedited. Taking these constructive suggestions from the HUPO community into account, the C-HPP leadership has been working on the reorganization of the current C-HPP Consortium into more active working group modules in three teams. The current C-HPP consortium will continue to run, and the working groups are more focused on special missions as key players in the next Top MP50 campaign (e.g., MP search, annotation, and functional study).

- MP Annotation Team (Team Red): This team mainly comprises those involved in the MP50 campaign led by Chris Overall.
- MP Bioinformatics Team (Team Green): This team mainly comprises those involved in public DB maintenance or related bioinformatics services in each Red and Blue team. This team is led by Lydie Lane.
- MP Functional Study Team (Team Blue): To expedite the discovery of rarely expressed MPs, it was suggested that we organize several MP functional study groups that will use various cell lines, rate tissues (e.g., nasal epithelial for olfactory receptor), IVTT tech, membranes, and model organisms. This team is led by Young-Ki Paik.



Briefings in the 2016 C-HPP Workshops

Summary of C-HPP Cancer Cluster Meeting in Xiamen, China (May 20th – 21th)

Organized by Pengyuan Yang, China

Friday, May 20th

To explore how our cooperative works could best be activated and further developed, a Cluster group meeting was held during the C-HPP workshop in Xiamen. Because some cluster group members (e.g., IVTT, Membrane, and Reproduction etc.) were unable to attend and consequently organized their own meetings, the Cancer cluster meeting was the only one held in Xiamen.

The workshop started with a few good talks presented by Ed Nice (Monash U., Australia) on Common mAb resources for the Cancer cluster group, Heeyoun Hwang (KBSI, Ochang, Korea) on the "Search Pipeline of Single Amino Acid Variants using neXtProt Database," and other local speakers. A general discussion on the Cancer Cluster followed, led by Gil Omenn. In this session, topics were focused on (1) "Strategy for Cooperative Works on Mapping" and "Cataloging ASVs, SNP, Missing Proteins and nlc-RNAs that are relevant to Cancers: Open Discussion," (2) Joint project development of bioinformatics tools and DBs for cancer study (in collaboration with cancer groups in B/D-HPP), and (3) Resource sharing (Bio-Banks, Reagents, DBs, NGS analysis etc.).

Lydie Lane presents update on the neXtProt

Discussion session in Xiamen Workshop

Discussion Summary (Moderator: Gil Omenn)

- Strategy for Cooperative Works on Mapping and Cataloging ASVs, SNP, Missing Proteins, and nlc-RNAs that are Relevant to Cancers: Open Discussion
- Collaborations Issue: how to ensure more efficient collaboration between Chr teams on cancer research. Many suggestions and comments were made. The Chinese team (Chr 1, Ping Xu) noted that they have put some effort into proteomics analysis of testis samples and cross comparison between datasets obtained from different cancers (GC, HCC, CRC) and testis proteins using combined techniques (RNA Seq & MS). Gil Omenn suggested that we could

use the recent release of the Human Protein Atlas, which contains 999 testis proteins and 318 brain proteins, which would be quite useful for cross-checking missing proteins from testis tissues. A few groups (Chr 1, 11, 2, 14, Y) are apparently already engaged in testis analysis. Performance of protein analysis on the PE2 gene transcripts was encouraged. Gil Omenn added that cancer stem cells could be good targets for PE2 proteomic analysis due to their heterogeneous nature. Young-Ki Paik suggested that each C-HPP team needs to find teams that have strong technical know-how or resources (cell lines, clones etc.) and form a small collaborative team within the Cancer Cluster. Jong Shin Yoo stated that the Chinese (Chr 20) and Korean (Chr 11) teams have explored the possibility of sharing their specialty database (e.g., RNA Seq and MS profiling of SAAVs). They are also working on testis tissue and are looking for a good collaborator from Chinese teams on these subjects. Mark Baker commented that the teamwork between Chr 7 and 17 in the cancer proteomics work is an example of successful collaboration.

- Amplicon & lnc RNA Study: Gil Omenn emphasized the work of the Amplicon study on cancers by showing good examples of her2/neu case and ERBB1, 3 driver mutations for multiple cancers (breast, CRC, and stomach cancers) in the course of Chr 17 study. This work will be useful for precision medicine. Few teams are currently working on this topic, and it can be further discussed in a congress meeting. Several issues have emerged in regard to the functional aspects and detection of nlc RNAs and small proteins. These peptides or small proteins seem very important emerging targets for cancer biomarker studies.

Friday, May 20st

On May 21, following the general reports, a session presented by three co-chairs (YK Paik, L. Lane, and C. Overall) provided an update on the C-HPP consortium organization, neXtProt, and JPR SI. Peter Horvatovich provided an update on the C-HPP wiki site and website status. Invited talks were given by Gil Omenn (metrics) and Mark Baker (missing proteins-PE2-4 & olfactory receptors). C-HPP PIs gave presentations on their progress with mitochondria (Andrea Urbani and Paola Roncada), chromosome 21 (Daniel Figeys), chromosome 18 (Alexander Archakov and Victor Zgoda), chromosome 16 (Fernando Corrales), chromosome 7 (Ed Nice), chromosome 6 (Chris Overall), chromosome 5 (Peter Horvatovich), chromosomes 2 and 14 (Lydie Lane, with an introduction of Chr 14 group works led by Jerome Garin), chromosome 20 (Siqi Liu and colleagues), chromosome 8 (Pengyuan Yang), chromosome 1 (Ping Xu), chromosome 11 (Jong Shin Yoo), chromosome 9 (Je-Yuel Cho and Soo-Youn Lee), chromosome 12 (Ravi Sirdeshmukh), and chromosome 4 (Yu-Ju Chen).

Special talks were given by Jiashu Tang (Thermo Fisher Scientific Life Scientific Applications) on "Mass Spectrometry-based Proteomics in Chromosome Biology and Epigenetics" and by Dr. Wenhai Jin (Sr. Manager, Application Support, SCIEX) on "Combining Next Generation Proteomics and NGS through OneOmicsTM to Investigate CTB Cells for Studying Placental Abnormalities.".

Executive Committee Members of C-HPP and Gilbert S. Omenn, HPP Chair

C-HPP PIC members and their Co-workers in Xiamen

Summary of 15th C-HPP Workshop in Taipei and Sun Moon Lake, Taiwan (September 17th – 23th)

Organized by Yu-Ju Chen

Overview of HPP Activity: HPP's Annual General Investigators Meeting was held on Sunday morning (September 18). C-HPP and B/D-HPP investigators gathered to discuss HPP progress and future plans. In a brief overview session, the leaders of C-HPP (Young-Ki Paik) and B/D-HPP (Jennifer van Eyk) presented their yearly activity summary and directions. This overview was followed by presentations by chairs of the HUPO Technology Pillars: Eric Deutsch for Bioinformatics, Emma Lundberg for Antibodies, and Susan Weintraub for MS, who provided an update on the HPP guidelines for MS data handling and activity.

Young-Ki Paik Chairing the C-HPP PIC Meeting in Taipei

Invited Talks: The highlights of the morning session were lectures by Mike Snyder (Chair, SSAB of HPP, Stanford University, USA) and Alan Saghatelian (Salk Institute, San Diego, USA). These two talks informed us about conceptual advances in the scientific direction of HPP research and also provided new guidance on how we can more completely achieve our C-HPP goals and human proteome annotations. In his talk on "Genomics for HPP", Mike addressed the essential roles of RNA Seq for Proteogenomics-based HPP studies with respect to the identification of new isoforms (splice junction mapping, synthetic long reads). He also introduced the use of barcodes for sequence variants (e.g.,

Mike Snyder and Alan Saghatelian talk at the morning session

PacBio System). For his second topic, he focused on the full use of Global Project Resources such as GTEx Project Data, which aims to identify correct transcripts for each gene in tissues. He also discussed pQTL as a means to study the enhancer binding proteins (Tr Factors) in different tissues and genetic variation–driven protein abundance quantitation. He recommended the use of the same set of tissues across many people (vs. versa). In his talk on "Small ORF Analysis," Alan suggested the incorporation of microproteins (or small ORF encoding proteins, SEPs) into ongoing C-HPP activities. In practical terms, this would involve seeking ways to incorporate microproteins (or short ORF-encoding peptides; SEPs, <150 AA) and adopting general strategies for identifying SEPs. He provided some examples of biological studies of so-called "Nobody" proteins by showing their cellular localization, subcellular distribution, and functional aspect by using shRNA K.D. or rescuing its function. He also demonstrated how RNA Seq and proteomics validate actual small protein coding genes. Both talks greatly widened our endeavors on the remaining 2579 missing proteins (neXtProt 2-9-2017 release). Our thanks go to both speakers for their stimulating talks.

C-HPP Working Group Session I: After a coffee break, the C-HPP consortium held a discussion session on three topics: recapturing the concept of HPP cluster grouping (by Paik), the direction of MP50 challenge (by Overall), and progress in the annotation of missing proteins at neXtProt portal (by Lane). In addition, several C-HPP PIs, including Chr 3 (Toshihide Nishimura), 15 (Giberto Domont), 18

(Alexander Archakov), 22 (Min-Sick Kim), and Y (Ghasem Hosseini Salekdeh) showcased their progress on the missing protein hunt and related scientific work.

Christopher Overall Discussed 2017 JPR SI Plans

Joint HPP Workshop in Sun Moon Lake (I)

Joint Cluster Group meeting: The first Cancer Cluster Group Meeting involved more than 45 attendees. In this session, several guest speakers including Phil Robinson (ProCan), Jerry Lee (NCI Office), and Henry Rodriguez briefly introduced the "Moon Shot project" shared by Australia and the USA and its effect on international cancer proteomics research. Follow-up open discussions on the vision of cancer projects and related topics were held by those panels, especially Mark Baker, Hui Zhang (JHU), Chris Kinsinger (CPTAC), and Jacob Kagan (EDRN). The participants felt that the meeting was very successful. Two additional cluster group meetings (the IVTT Cluster Group led by Gyorgy Marko-Varga and the Reproductive disease cluster led by Charles Pineau and Ghasem Hosseini Salekdeh) were held separately. The next morning, the Membrane proteome (led by Daniel Figeys) and Neurodegenerative disease cluster groups held their own discussions. The latter meeting (led by Jong Shin Yoo and Alberto Urbani) was jointly held by HBPP (B/D-HPP) and the neuro-disease group (C-HPP), and the participants shared many views on resources, methods, and database construction.

Joint HPP Workshop in Sun Moon Lake (II)

Note: The term "cluster" was coined by Young-Ki Paik, Lydie Lane, and Gil Omenn when they met to discuss a strategy to activate HPP activity in Milano in 2015. It can be defined as an informal grouping to facilitate communications between the HPP working groups (i.e., B/D-HPP, C-HPP, Pillars, CPTAC, and related HPP groups) so that the HUPO membership and general public see it as a way of organizing sessions and workshops rather than as a parallel set of groups, as suggested by Chris Overall. After it was trialed at the 2015 Vancouver and 2016 Taipei HUPO congress meetings, C-HPP EC decided to operate the cluster team activity only as an internal study group within the C-HPP consortium (e.g., SAAV, lnc-RNA, etc.). This new attempt was designed to strengthen the chromosome-based mapping of proteins and functional study of missing proteins.

C-HPP PIC Members in Sun Moon Lake, Taiwan

Participants of Joint HPP Workshop in Sun Moon Lake

Summary of 16th C-HPP Workshop in Rio de Janeiro, Brazil (December 10th – 12nd)

Organized by Gilberto B. Domont, Brazil

The 16th C-HPP workshop was held in Rio de Janeiro, Brazil between 10 to 12 December 2016, in conjunction with the Centennial Anniversary of the Brazilian Academy of Sciences with message of "Translating the code of life into proteins and diseases".

In this workshop the following topics were addressed:

- Joint discussion and integration of the Chromosome-centric Human Proteome and the Human Biology Diseases Projects.
- Presentation of cutting-edge science on disease proteomics and mass spectrometry with impact on human diseases and targeted medicine.
- Assessment and evaluation of the first Human Proteome Map.
- Other goals of the C-HPP 1st phase not yet fully addressed are the pursuit of systematic ways for identification of proteoforms and complexes as well as cell & tissue localization. Need for distinct data deposition for human proteoforms & complexes.
- Definition of C-HPP phase 2: validation.
- Open discussion with scientists as well as lay citizens and media on the impact of the Human Proteome Project in daily life.

The meeting was opened Saturday December 10 by Pedro Vasconcelos representing the Brazilian Academy of Sciences, Marcelo Valle de Sousa, President of the Brazilian Society of Proteomics, Marcos Eberlin, president of the Brazilian Society of Mass Spectrometry and Mark Baker, HUPO President.

The keynote lecture on Zika virus epidemics: current situation and perspectives was presented by Pedro Vasconcelos from Instituto Evandro Chagas (Ministry of Health). This presentation provided a background and history on the Zika virus discovery, epidemic and current status of research and defense against Zika infection.

Sunday December 11st in the cutting edge section Catherine Costello (Boston University) presented the progress in the assignment of glycosylation post-translational modification in proteomics experiments. Manuel Fuentes (Universidad de Salamanca) and Fernando Corrales (Universidad de Navarra) provided update on application of in-vitro transcription and translation approach in identification of missing proteins, with promising identification of some missing proteins from collaborating chromosomes 5, 10, 15, 16 and 19, in cell lines, and in sperm. Peter Horvatovich (University of Groningen) gave a lecture on bioinformatics background of proteogenomics data integration of transcriptomics and proteomics for human lung tissue to reveal the molecular mechanism of COPD. Fabio Gozzo from the Universidade de Campinas gave a talk on evolution of cross-linking/MS: from dimers to large-scale quantitative interactome and to protein structure and the talk of Garry Corthals from University of Amsterdam focused on quality and accuracy of analysis for proteome-scale phosphorylation studies. Christoph Borchers from University of Victoria presented a talk on the development of targeted multiple-reaction-monitoring assays to determine more than 1000 proteins in mouse tissues and biofluids. Daniel Martins de Souza (Universidade de Campinas) and Peter Nilsson (Royal Institute of Technology) gave the summary on the status of Human Brain Proteome Project and Neuroproteomics HPP and provided insight into the use of proteomic profiling to reveal molecular mechanisms of psychiatric disorders. Melinda Rezelli, University of Lund talked on the use of proteomic profile to identify the proteome change during Alzheimer's disease onset and progression related to tau pathology. The workshop followed with a common section with the 1st Ibero-American Conference on Mass Spectrometry where the keynote speaker was Alexander Makarov, the inventor or Orbitrap mass analyser.

Monday 12nd December the workshop host Gilberto B. Domont (Universidade Federal do Rio de Janeiro) leaded a debate on the assessment and evaluation of the publication of the first HUPO Human Proteome Map. This was followed by discussion on the role of proteomics for personal and translational medicine by Gyorgy Marko-Varga and Johan Malm (University of Lund). The results of these discussions were summarized in the "The Rio Directive". This document recognizes the achievements of identification of 16,518 human proteins corresponding to 85% of the human proteome by the proteomics community and the human proteome project and made recommendation to restructure HPP as well as suggested closer work with clinicians and other experts working in human health. Rio Directives was submitted to the HUPO Executive Committee for discussion. Further details is available on the C-HPP Wiki at http://c-hpp.webhosting.rug.nl/tiki-index.php?page=16th+C-HPP+Workshop+in+Rio+de+Janeiro%2C+Brazil%2C+December+10-12%2C+2016.

Gilberto B. Domont Hosts the 16th C-HPP workshop in Rio de Janeiro, Brazil

Participants of 16th C-HPP Workshop in Rio de Janeiro, Brazil

Future C-HPP Workshops

- 17th C-HPP Symposium/Workshop, April 27-28th in 2017 (by Ghasem Hosseini Salekdeh)
- 18th C-HPP Workshop in Dublin during 16th HUPO Congress, 9/17-21, 2017
- 19th C-HPP Workshop in Santiago, Spain (Jointly Organized with EuPA & Fernando Corrales) Last week June, 2018
- 20th C-HPP Workshop in Orlando, USA (TBA), 2018
- 21st C-HPP Workshop in St. Malo, France, May13-14, 2019 (by Charles Pineau)

17th C-HPP Workshop in Tehran, Iran (April 27th-28th, 2017)

Registration

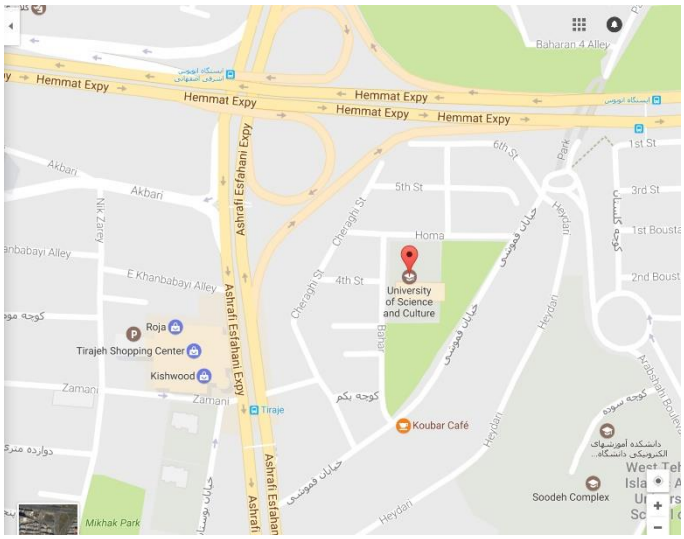
Resalat Highway, Banihashem St., Shaghayegh Alley, No: 9.

Tel/Fax: +98-21- 23562178

Email: info@royancongress.com; www.royaninstitute.org

Date & Venue

April 27-28, 2017; Conference Hall of University of Science & Culture, Tehran, Iran



Invited Speakers

Conference venue: University of Science & Culture, Bahar Ave, Shahid Ghomoushi Ave, Hemmat Highway, Ashrafi Esfehani Blvd, Tehran, Iran

Thursday, April 27th

08:00 – 08:50	Registration
08:50 – 09:00	Welcome Message: Prof. Hamid Gourabi (President of Royan Institute)
09:00 – 09:20	Plenary Lecture (Chair: Hossein Baharvand) Prof. Reza Malekzadeh (Deputy Minster for Research and Technology, Ministry of Health and Medical Education, Iran)
Invited Session I: Introduction and Overview: C-HPP Direction, Importance, Data Guidelines and Technology Resources (Chair: Young-Ki Paik)	
09:30 – 10:00	C-HPP: Past, Present and Future: Young-Ki Paik (C-HPP Chair, Korea)
10:00 – 10:20	JPR Special Issue and MP50 Campaign with HPP Data Guidelines: Chris Overall (C-HPP Co-Chair, Associate Editor for JPR)
10:20 – 10:40	Link between C-HPP and Biology/Disease HPP (Fernando Corrales, Spain)
10:40 – 11:10	Coffee Breaks
Invited Session II: Uncovering Missing Proteins by Using Various Biological and Clinical Samples (Chair: Chris Overall)	
11:10 – 11:30	Searching for missing proteins: Update on chromosome 16 activity (Fernando Corrales, Spain)

11:30 – 11:50	Looking for missing proteins in the testicular germ cell lineage: new insights into normal and pathological spermatogenesis (Charles Pineau, France)
11:50 – 12:10	Identification of Missing Proteins Encoded in Chromosome X (Tadashi Yamamoto, Japan)
12:10 – 12:30	Mitochondrial Proteome-Completion of Missing Protein Annotation (Andrea Urbani, Italy)
12:30 – 12:50	Updates of missing proteins and novel variant findings from chromosome 11 (Heeyoun Hwang, Korea)
12:50 – 13:10	SRM-based 'missing' proteins of UPS2 (sigma) and Chr18 expressed in human liver, HepG2 cells and in plasma (Alexander Archakov, Russia)
13:10 – 14:30	Lunch Break
Invited Session III: Functional Study of Missing Proteins with Disease Implication (Chair: Fernando Corrales)	
14:30 – 14:50	Functional Validation of the Previous Missing Proteins Involved in Human Reproductive Disease: Experimental Strategy and Pitfall (Young-Ki Paik, Korea)
14:50 – 15:10	Heart of Y Chromosome: Searching for Functions of Missing Proteins Involved in Disease (Ghasem Hosseini Salekdeh, Iran)
15:10 – 15:30	C-HPP Project (Chr 9): Proteogenomic Study in Human Lung Cancer (Je-Yoel Cho, Korea)
15:30 – 15:50	Progress in Chr 8 with Disease Implications (tentative) (Pengyuan Yang, China)
15:50 – 16:10	Update on the Food and Nutrition Initiative (Paola Roncada, Italy)
16:10 – 16:30	The hidden human proteomeL theprotein-coding LncRNA (Gong Zhang, China)
Invited Session IV: Invited Special Lectures (Chair: Fernando Corrales)	
17:00 – 17:30	Special Lecture 1: Antiviral Immunity Cellular Substrates revealed by TAILS Terminomics (Chris Overall, Canada)
17:30 – 18:00	General Discussion on All Pending Issues (Chairs: Young-Ki Paik, Chris Overall, Fernando Corrales)
19:00 – 21:00	Milad Tower (PIs and Co-PIs)

Friday, April, 28th

Young Investigator Invitation Session (Chair: Pengyuan Yang, Fudan Univ., China)	
08:30 – 08:45	Identifying 17000 human proteins in a single MS experiment using high throughput de novo identification aided by translatome sequencing (Dehua Li, China)
08:45 – 09:00	Human Y chromosome genes regulate Human Embryonic Stem Cell Differentiation to cardiac cell (Anna Meyfour, Iran)
09:00 – 09:15	Probable Physicochemical and Epigenetic Causes of Missing Proteins (Fan Zhong, China)
09:15 – 09:30	Human Y chromosome proteome project and male infertility (Mehdi Alikhani, Iran)
09:30 – 09:45	Bioinformatic tools of Chromosome 13 team for finding missing proteins (Seul-Ki Jeong, Korea)
09:45 – 10:00	Human Membrane proteomics: searching for missing protein (Faezeh Shekari, Iran)

10:00 – 10:30	Coffee Breaks
Working Group Formation: Networking and Interaction through the C-HPP Clusters (Co-Chairs: Young-Ki Paik and Chris Overall)	
10:30 – 12:00	Organization of Intra C-HPP Cluster Groups / Rare Tissues and Cell Lines Cluster Olfactory Receptor Cluster / Novel Protein Cluster PTM Cluster / IVTT Cluster / Infertility cluster / lnc RNA Cluster (Call-In: Gil Omenn)
12:00 – 13:30	Future Directions (Chair: Charles Pineau) Update on the 18th C-HPP Workshop in Dublin (Young-Ki Paik and Fernando Corrales) Update on the JPR Special Issue (Chris Overall, Co-Chair & JPR AE) Planning on the 19th C-HPP Workshop in Santiago (Fernando Corrales) Planning on the 20th C-HPP Workshop in St. Malo (Charles Pineau) Planning on the Collaboration with B/D-HPP Group (Fernando Corrales)
End of Symposium	

18th C-HPP Workshop in HUPO Congress in Dublin, Ireland (http://www.hupo2017.ie/)

Opening of registration	13th January 2017
Call for Abstracts	11th January 2017
Close of Abstracts	3rd May 2017
Notifications of Acceptance	31st May 2017
End of Early Registration	14th June 2017
End of Regular Registration	16th August 2017

HUPO
Dublin Ireland
17 - 21 September 2017
16th Human Proteome Organisation World Congress

DRAFT PROGRAMME

	SATURDAY 16th	SUNDAY 17th	THURSDAY 21st
	University College Dublin	University College Dublin	University College Dublin
8:30am-5:00pm	Congress Workshops on UCD Campus	Congress Workshops on UCD Campus	The HPP Post Congress workshop will take place in George Moore Auditorium in the O'Brien Centre for Sciences in UCD (University College Dublin) on Thursday 21st September

6 CONCURRENT THEMES

Cellular Proteomics	Drug & Biopharma-ceutical Proteomics	Systems, Bioinformatics & Omics Data Integration
Disease & Clincial Proteomics	Precision & Personalised Proteomics	HPP (B/D HPP and C-HPP)

	SUNDAY 17th	MONDAY 18th	TUESDAY 19th	WEDNESDAY 20th
	The Convention Centre Dublin	The Convention Centre Dublin	The Convention Centre Dublin	The Convention Centre Dublin
8:30-10:00am		Plenary Session	Plenary Session	Plenary Session
10:00-11:00am		Coffee Break Exhibition, Networking and Poster Viewing	Coffee Break Exhibition, Networking and Poster Viewing	Coffee Break Exhibition, Networking and Poster Viewing
11:00-12:40pm		6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations	6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations	6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations
12:40-2:10pm		Sponsor Lunchtime Workshops Exhibition, Networking and Poster Viewing	Sponsor Lunchtime Workshops Exhibition, Networking and Poster Viewing	LUNCH BREAK Exhibition, Networking and Poster Viewing
2:10-3:50pm		6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations	6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations	6 Concurrent Sessions comprising: Keynote Speakers, Invited Speakers and Selected Oral Presentations
3:50-4:50pm		Coffee Break Exhibition, Networking and Poster Viewing	Coffee Break Exhibition, Networking and Poster Viewing	Plenary Session
4:50-5:30pm	3:45-5:45pm HUPO Council Meeting	Plenary Session	Plenary Session	
5:30pm		Poster Reception	HUPO AGM (1 hour)	
6:00pm	HUPO Opening Ceremony			
7:00pm	6:20-7:00pm Opening Plenary	Optional Social Events Various Dublin Venues		
7:30pm	Opening Ceremony		HUPO Gala Evening The Guiness Storehouse	
8:00pm	Welcome Reception			

CHPP
Chromosome-centric Human Proteome Project

Newsletter Editor: Young-Ki Paik

Editorial Staff: Peter Horvatovich, Jin-Young Cho

HQ Office:
Yonsei Proteome Research Center, Yonsei University
50 Yonsei-ro Sudaemoon-gu, Seoul, 120-740, Korea
(www.c-hpp.org)

Chair: Young-Ki Paik
Co-chairs: Lydie Lane, Chris Overall