National Cancer Institute
National Institutes of Health
U.S. Department of Health and Human Services

International Summit on Proteomics Data Release and Sharing Policy

August 14, 2008

Hotel Novotel
Amsterdam, The Netherlands

AGENDA

8:00 a.m. - 8:10 a.m. **Welcome and Review of Agenda**

- Meeting background and framework
- Defining the stakeholders (i.e., producers, users, funding agencies, others?)

Henry Rodriguez, Ph.D., M.B.A.
National Cancer Institute, NIH

Rolf Apweiler, M.Sc., Ph.D.
European Bioinformatics Institute

Session: History on Data Standards and Challenges

8:10 a.m. - 8:35 a.m. **Lessons Learned from Genomics**

- Current Bermuda Policies on pre-publication data access
- History of a successful policy adoption and implementation

Michael Dunn
The Wellcome Trust

8:35 a.m. - 9:00 a.m. **Challenges in Release and Sharing of Mass Spectrometry Data**

- MS data formats and analysis workflow
- Quantitative proteomics data and processing methods
- Diversity of software tools and data analysis methods
- Dependence of results on evolving data processing methods

Alexey I. Nesvizhskii, Ph.D.
University of Michigan

9:00 a.m. - 9:25 a.m. **Other Forms of Proteomics Data**

- Tissues arrays
- Antibodies
- Images

Fredrik Pontén, M.D., Ph.D.
Uppsala University

Session: Current Standards and Requirements

9:25 a.m. - 9:50 a.m. **Molecular and Cellular Proteomics (MCP Journal)**

- "Paris Guidelines" (MCP)
- Implementation and Success of the MCP Guidelines

Robert J. Chalkley
University of California, San Francisco

9:50 a.m. - 10:00 a.m. **Break**

10:00 a.m. - 10:25 a.m. **Emerging Standards in Proteomics and European Bioinformatics Institute (EBI) Policies on Data Sharing - HUPO-PSI**

- PRIDE (PRoteomics IDEntifications database)
- Other EBI guidelines

Rolf Apweiler, M.Sc., Ph.D.
European Bioinformatics Institute

10:25 a.m. - 10:50 a.m. **caBIG: a Federated Strategy for Proteomics Data**

- Interoperability
- Compliance standards

Liming Yang, Ph.D.
National Cancer Institute, NIH

10:50 a.m. – 11:15 a.m. **Repositories (including Formats)**

- Successful data sharing at National Center for Biotechnology Information (NCBI)
- Standards
- Curation
- Automation

Ron Edgar, Ph.D.
National Center for Biotechnology Information, NIH

11:15 a.m. - 11:55 a.m. **What's Missing and What Can Be Done To Improve Data Sharing?**

- Who wants to use the data?
- Processed results vs. raw data?
- How to integrate with existing data sharing efforts?
- What 'standards' are needed (formats, annotation, sequence, …)
- How to deal with evolving experiments and data analysis?
- How to minimize risks of pre-publication sharing?
- How to promote and what to avoid?
- Where data should be stored and made available?

Ronald Beavis, B.Sc., Ph.D.
University of British Columbia

Group Discussion

11:55 a.m. - 12:55 p.m. **Lunch**

Session: Breakout Groups

1:00 p.m. - 3:15 p.m. **Breakout Groups**

Group 1: Define metrics for data quality in the context of establishing a threshold as to what data is fit to be shared
Chair: Ronald Beavis, B.Sc., Ph.D.
University of British Columbia

Co-Chair: Stephen E. Stein, Ph.D.
National Institute of Standards and Technology

Group 2: Publishers – How should policies be designed and implemented?
Chair: Mike Snyder
Yale University

Co-Chair: Peter Hare
Nature Biotechnology

Group 3: Protein/Affinity Arrays – What data needs to be shared?
Chair: Albert J.R. Heck
Netherlands Proteomics Centre and Utrecht University

Co-Chair: Fredrik Pontén, M.D., Ph.D.
Uppsala University

3:15 p.m. - 4:00 p.m. **Break**
(chair and co-chairs write outcomes for presentation)

4:00 p.m. - 5:30 p.m. **Breakout Groups Report to Meeting Attendees**

5:30 p.m. - 5:45 p.m. **Closing Questions/Comments, Action Items**

5:45 p.m. **Meeting Adjournment**