Meeting report

Genomics, proteomics and bioinformatics: all in the same boat

Robert B Russell

Address: EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany. E-mail: russell@embl-heidelberg.de

Published: 24 September 2002

Genome ***Biology*** 2002, **3(10)**:reports4034.1–4034.2

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2002/3/10/reports/4034

A report on the Genomics, Proteomics and Bioinformatics for Medicine (GPBM) 2002 meeting, St. Petersburg to Moscow, Russia, 22-30 June 2002.

A boat trip from Saint Petersburg to Moscow past many of the gems of Russian history sounds like a setting for a nineteenth century novel. But if you replace landed gentry on a river tour with molecular biologists and add nearly eighty scientific presentations, you come close to a picture of Genomics, Proteomics and Bioinformatics for Medicine. As the title suggests, the aim of the meeting was to discuss the current state of the art and how it impacts on medicine.

A great variety of medical applications of all techniques were evident throughout the meeting: genome-wide analysis of pathogenic organisms, searches for mutations implicated in diabetes, the proteomics of cancer, not to mention sessions devoted to cytochromes P450. Several talks showed that the P450 of electron-transport proteins and its functional partners have been subjected to virtually all forms of genomic, proteomic and bioinformatic investigations. Here, some highlights of the meeting are presented with a focus on advances in technology.

Genomics

Wilhelm Ansorge (EMBL, Heidelberg, Germany) discussed microarray data quality and the effect of several factors, such as oligonucleotide length and concentration, and improvements such as microelectric currents for hybridization. The sensitivity of the microarray technique has improved to the extent that clinical samples as small as 10 to 50 nanograms can be studied. Most provocatively, Ansorge said that his findings suggest that many of the early experiments with microarrays will need to be repeated. This theme of quality control would arise more than once over the course of the meeting.

Some problems with DNA chips are related to smearing of samples. Claudio Nicolini (University of Genova and Polo Nazionale Bioelectronica, Italy) discussed the DNAser technique for avoiding this problem. He demonstrated how the use of a heterogeneous surface, consisting of hydrophobic areas surrounding hydrophilic spots specifically activated for oligonucleotide immobilization, led to improvements in DNA-chip resolution.

Proteomics

Despite advances in technology for genome-wide analysis of nucleic acids, it is often the case that a better picture of the molecular biology of a cell comes only from the study of proteins. This fact was nicely illustrated by Laura Beretta (University of Michigan, Ann Arbor, USA), whose combined use of microarrays and proteomics suggested that dendritic cell maturation is controlled by events occurring after transcription and translation.

Denis Hochstrasser (Geneva University Hospital, Switzerland) gave an overview of his group's approach to studying components of blood plasma. They first analyze plasma with mass spectroscopy, and then use bioinformatics to identify the proteins. Novel proteins are then synthesized and tested *in vivo* for biological effects. Hochstrasser also discussed a new device called the molecular scanner, which couples tryptic digestion, SDS-polyacrylamide gel electrophoresis and mass spectrometry for real-time identification of sequences of proteins within samples.

Vadim Ivanov (Russian Academy of Sciences, Moscow, Russia) discussed 'peptidomics' as a follow-up to proteomics and genomics. Work in his laboratory has identified hundreds of new peptides derived from known proteins that are highly abundant in, for example, rat brain or spleen. He suggested that because of their high concentration, even weak binding to receptors might stimulate effects comparable to stronger-binding peptides, such as hormones, that occur at lower concentrations.

When one speaks of genomics or proteomics of eukaryotic systems, it is usually assumed that one is working with a model organism with a completed genome sequence, because without this information it is often not possible to identify proteins quickly. But what of the great majority of organisms for which no genome sequence is available? Andrej Shevchenko (Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany) discussed an approach that makes possible proteomic studies of organisms lacking sequenced genomes. His group combine *de novo* protein sequencing by mass spectrometry with new bioinformatic approaches in such a way that they can successfully identify homologs from organisms that have been sequenced, even when peptide sequences are error-prone.

Although knowledge of protein constituents and their abundance is key to understanding physiological and pathological processes, many more answers often come from an understanding of molecular or atomic details. Alexander Archakov (Russian Academy of Sciences, Moscow, Russia) displayed some of the first protein structures, albeit of low resolution, derived by using atomic-force microscopy (AFM). Using a variety of membrane proteins, he showed how AFM can be used to visualize single particles corresponding to binary and ternary protein complexes. When AFM is coupled with an optical biosensor, it is also possible to measure the kinetics of these interactions.

Large-scale means of determining the finer structural details of proteins were discussed by Hartmut Oschkinat (Institute of Molecular Pharmacology (FMP), Berlin, Germany) in his overview of the Berlin Protein Structure Factory (the first Structural Genomics project in Europe). He described the determining of quantitative structure-activity relationships (QSAR) by nuclear magnetic resonance (NMR) to detect chemicals that bind weakly to a protein drug target; two or more of these can subsequently be fused to produce a hybrid compound with a higher affinity for the target. NMR is unique in allowing studies of protein-ligand interactions dynamically in such atomic detail, an attribute further illustrated by Oschkinat's work on the selectivity of PDZ protein interaction domains for particular peptide ligands.

Bioinformatics

One of the fathers of protein annotation, Amos Bairoch (Swiss Institute of Bioinformatics, Geneva, Switzerland), discussed the current status of SWISS-PROT. What began as just one man with a mission and a few hundred sequences is now run by 75 to 80 people producing a database of 110,000 proteins (or over half a million if one includes the automatically annotated trEMBL set). Many staff are involved in annotation, with the ultimate aim of a complete description of protein function, which Bairoch aptly described as the "endgame of our life-sciences adventure". The process gets ever more complicated, necessitating more specialized databases, such as those specific to protein families or species (for example, the human proteomics initiative), in addition to features such as 'evidence tags' to permit users of the database to scrutinize the types and details of the information used for annotation.

The medical theme of the meeting made this a logical place to hear about the latest *in silico* methods of drug design. Eugene Shakhnovich (Harvard University, Cambridge, USA) presented a new approach that combines a knowledge-based atom-atom interaction potential derived from protein-ligand complexes with a computational combinatorial chemistry algorithm. The result is a set of possible lead compounds for a drug target of known structure. When the system was tested on the carbonic anhydrase II, a target for diseases such as glaucoma, Shakhnovich found an impressive correlation between predicted and experimentally determined free energies. A different *in silico* approach for drug discovery was later discussed by Vladimir Poroikov (Russian Academy of Sciences, Moscow, Russia). The prediction of activity spectra for substances (PASS) system correlates molecular features of compounds with biological activities and is thus able to predict prospective compounds for a biological activity of interest.

A mixed future full of complexity

Uwe Eichhoff (Bruker Biospin GmbH, Rheinstetten, Germany) discussed high-resolution NMR of body fluids to determine the concentrations of various metabolites. This 'metabolomics' approach appears quite powerful: it is apparently able to detect metabolic diseases based on urine samples, and intriguingly, has suggested other deviations from the norm that could correspond to metabolic disorders that have not yet been characterized.

As Denis Hochstrasser suggested at the outset of the meeting, the various subjects are united in that they all aim to study complexity. Many methods are needed to understand the complex behavior of biological systems, meaning that the future will involve a mix of genomics, proteomics, bioinformatics and many other new technologies and disciplines, some of which currently lack even a name. The methods, like the biological systems they study, often defy classification. Our ultimate endgame of understanding biological function means that we are all in the same boat and that we will always need to allow others on board.