

Biologists initiate plan to map human proteome

Project aims to characterize all human proteins.

Helen Pearson

Ambitious plans to catalogue and characterize all proteins in the human body — a Human Proteome Project — are being drawn up by a small group of researchers. But with a price tag of around US\$1 billion, some question whether the organizers can raise enough money or momentum for such an undertaking.

Researchers looked into the idea in the mid-1990s as the Human Genome Project was taking shape — the human proteome seemed a natural successor. However, a coordinated effort to index human proteins never emerged. One reason is that the scale and complexity of the problem proved daunting and nebulous. Protein-coding genes in the body can make tens of different versions of a protein, and each of these can be modified by the addition of chemical groups in countless different ways. All these proteins are being manufactured at differing levels, and at different moments in time, in the 200 or so types of human cell. “It was thought to be beyond comprehension,” says John Bergeron of McGill University in Montreal, Canada, former president of the Human Proteome Organisation (HUPO).

Now Bergeron and a group of leading proteomics researchers are putting together a proposal for a large-scale assault on the human proteome. It would reveal which proteins are present in each tissue, where in the cell each of those proteins is located and which other proteins each is interacting with.

(The human genetic sequence, by contrast, shows which regions code for proteins but not which are actually making them.)

Proponents say that this type of protein catalogue will be invaluable in revealing new drug targets or biomarkers to track the progression of disease. The cost “is absolute peanuts when you consider the importance of mapping the building blocks of life”, says Mathias Uhlen at the Royal Institute of Technology in Stockholm, Sweden, who is helping shape the new project.

Two preliminary workshops have been held to discuss the endeavour, most recently in Barbados in January this year. The group plans to consult with the wider proteomics community for the first time at HUPO's Amsterdam World Congress in August.

Those involved in the draft plan say that a human proteome project is now feasible partly because estimates of the number of protein-coding genes have shrunk. It was once thought that there might be around 50,000 or 100,000, but now, just 21,000 or so are thought to exist, making the scale of human proteomics more manageable. And the group plans to focus on only a single protein produced



Heart proteins seen separated in gel.

J. KING-HOLMES/SCIENCE PHOTO LIBRARY

from each gene, rather than its many forms. “We got rid of all this complexity,” Bergeron says. “We tried to craft a project that would be doable with easy-to-track milestones.”

The plan is to tackle this with three different experimental approaches. One would use mass spectrometry to identify proteins and their quantities in tissue samples; another would generate antibodies to each protein and use these to show its location in tissues and cells; and the third would systematically identify, for each protein, which others it interacts with in protein complexes. The project would also involve a massive bioinformatics effort to ensure that the data could be pooled and accessed, and the production of shared reagents.

Bergeron envisages the work being divided up between labs around the world. He says that the first stage of the project — which involves amassing existing mass spectrometry proteomic data — would take around six months, and that this would be followed by a pilot project lasting one to three years to do more comprehensive analysis of all the proteins manufactured by chromosome 21, the smallest human chromosome. The whole effort could take a decade.

“It will be a battle to convince people this is worth funding.”

“It’s a huge undertaking for HUPO and they’ve never done anything like it,” says Paul Tempst, an expert in proteomics at the Memorial Sloan-Kettering Cancer Center in New York. The organization has run several large-scale proteomics efforts, such as those to catalogue the proteome in human blood plasma, the liver and the brain. But results from the Human Plasma Proteome project and other proteomics efforts showed that different laboratories — and even the same lab — often identify very different sets of proteins from exactly the same sample. “It told us that there were lots of proteins in plasma and that if you do it in different labs without any effort at technology standardization you get different results,” says Tempst.

However, Bergeron and others say that improvements in mass spectrometry techniques have resolved many of the problems of reproducibility that have dogged the field, and that it is now possible to reliably identify a range of proteins if a sample is analysed many times. But an additional difficulty comes in trying to analyse samples containing proteins at wildly differing concentrations, with some present in only a few copies. The three-pronged approach is designed so that shortfalls in one technique can be made up by the others.

Steven Carr, director of proteomics at the Broad Institute in Cambridge, Massachusetts, says there is likely to be broad support for a large-scale proteomics effort, but much debate about how best to do it. Rather than analyse the proteome of one chromosome, he says it may be better to tackle the proteome of mitochondria or the cell membrane because it would reveal more about biology and diseases related to those structures. “It’s time to think about something in a systematic fashion — whether this is the project is a different question,” he says.

Coordination could also be a challenge. Compared with the human genome project, in which a handful of sequencing centres did the bulk of the work, a human proteome project would involve many more labs, with inevitable issues about data-sharing and competition. “It will be a battle to convince people this is worth funding,” says Michael Washburn, director of proteomics at the Stowers Institute for Medical Research in Kansas City, Missouri.

Those involved in the budding project say that they must now muster support from funding organizations such as the US National Institutes of Health and the European Commission — a clear challenge in today’s tight funding climate. “It will be a tough sale,” says Sudhir Srivastava, head of the Cancer Biomarkers Research Group at the US National Cancer Institute in Rockville, Maryland. Srivastava says that the project may have a better chance of success with a stronger focus on diseases such as cancer. “You need to show clinical utility of the approach before launching a mega-project,” he says.

ADVERTISEMENT



XX INTERNATIONAL CONGRESS OF GENETICS 2008
BERLIN, GERMANY,
JULY 12 – 17, 2008
Late Breaking Abstract Deadline
April 15, 2008

International Genetics Federation (IGF)

BERLIN 2008
XX INTERNATIONAL CONGRESS OF GENETICS

German Genetics Society (IGG)

In partnership with
nature genetics

See Editorial, [page 913](#).

Comments

Reader comments are usually moderated after posting. If you find something offensive or inappropriate, you can speed this process by clicking 'Report this comment' (or, if that doesn't work for you, email redesign@nature.com). For more controversial topics, we reserve the right to moderate before comments are published.

I do agree in respect to the points of challenges facing this project and I think the idea of having systemic steps approach will be ideal using mitochondrial or cell proteins or starting the project depending on diseases eg Breast cancer and others. Comparing human samples from different countries, tribes might show different results. So again the project can be based on HLA typing. Dr. Nazar M. Abdalla. PhD molecular biology.

Posted by: **Nazar Abdalla** | 24 Apr, 2008

The importance of systems level understanding is undisputed. Not only at the level of proteomics but also at epigenomic, lipidomic, metabolomic, glycomic levels, besides others. Although these projects are going to be challenging, costly and time intensive, they need to be undertaken because all these efforts may eventually lead to a comprehensive understanding of cellular function that will help develop drugs with specific therapeutic action and minimum side effects.

Posted by: **abhay sharma** | 26 Apr, 2008

I think this is the right time to map the human proteome. After having huge pile of data on gene expressions based on microarray and more recently with the no. of small RNA regulator MicroRNA no. going up, its very difficult to predict the regulatory output of an organ in a particular stimulus / disease / micro environment. Studying human proteome in a quantitative manner could solve most of the biological problems. However addition of protein protein interactions and integration of its biological output could add another dimension of to the human proteome. The complexity is undefined at present but I feel its worth going for human proteome. Pankaj Mishra, Ph.D. Biotechnology

Posted by: **Pankaj Mishra** | 29 Apr, 2008

Mapping the human proteome and applying bioinformatics tools for knowledge generation particularly for target identification and mechanism of drug action will have great importance in the context of the burgeoning interest in traditional polyherbal system of medicines as well as in the case of metal medicines that are often complexed with organics/phenolics/aromatics. This can open up a new frontier of knowledge with immense portents for new drug developments sureshkumar,head,pme,NIIST

Posted by: **suresh kumar** | 01 May, 2008

Add your own comment

You can be as critical or controversial as you like, but please don't get personal or offensive, and do keep it brief. Remember this is for feedback and discussion - not for publishing papers, press releases or advertisements, for example.

You need to be registered with Nature to leave a comment. Please log in or register as a new user. You will be re-directed back to this page.

[Log in / register](#)