Meeting report

The post-genomic era for a select few

Paul Cliften

Address: Washington University Genome Sequencing Center, 4444 Forest Park Boulevard, St Louis, MO 63108, USA.
E-mail: pcliften@watson.wustl.edu

Published: 15 January 2004

Genome ***Biology*** 2004, **5:**308

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2004/5/2/308

A report of the 4th Colmar Scientific Symposium 'Biology in the Post-genomic Era', Colmar, France, 16-17 October 2003.

In genomics, as in life, the rich tend to get richer. Once a species has had its genome sequenced, it is more likely that close relatives will have theirs sequenced as well. This helps biologists collect a great deal of information about an organism, but it tends to exacerbate the imbalance of genomic resources. Only those studying these favored organisms can claim to have entered the 'post-genomic era'; the rest ask when this era will arrive.

Even those who work on the 'elite' organisms such as mouse and human are reluctant to concede that we have entered a post-genomic era. For instance, Jean Weissenbach (Genoscope, Evry, France) remarked that he would address biology in the 'genomic' rather than the 'post-genomic' era, pointing out that the human genome has yet to be completely annotated. Weissenbach described Exofish, a tool developed in his lab that identifies protein-coding DNA segments by comparing two genome sequences. Using Exofish, Weissenbach and his colleagues have compared the draft genome sequences of the pufferfishes *Takifugu* and *Tetraodon* with mammalian genome sequences. They identified additional exons within and adjacent to predicted genes in the mammalian genomes and thus were able to improve the annotation of the human and mouse genomes. As the sequences of many genomes are necessary to achieve a fuller understanding of mammalian genomes, Weissenbach made a compelling argument that completion of the human and mouse genome sequences does not necessarily signal the end of the genomic era.

Plant geneticists would certainly like many more complete genome sequences. Ralf Reski (Freiburg University, Germany) addressed the biology of the moss *Physcomitrella patens*. This primitive plant, which is a potential food source, has been engineered to express human proteins without giving them the plant-specific glycosylations that can make them allergenic in humans; this may lead to the use of such engineered plants to manufacture and deliver drugs. Furthermore, the haploid genome of *Physcomitrella* has a high rate of recombination, which enables gene-targeting approaches for functional genetic analysis that are not available in, for example, the model plant *Arabidopsis*. Thus, there are compelling practical and scientific reasons for obtaining the genomic sequence of *Physcomitrella*.

Although many bacterial genomes have been sequenced, and more are being completed almost daily, the answer to whether or not microbiology has entered into the post-genomic era depends largely on the organisms one wants to study and the questions one wants to ask. According to Jean-Michel Claverie (National Centre for Scientific Research (CNRS), Marseilles, France), there are plenty of microbial genomic sequences available for potentially identifying new targets for antibiotics. But, unfortunately, a large number of genomic sequences does not necessarily yield a large number of potential targets. On comparing the gene content of *Escherichia coli*, *Mycobacterium tuberculosis* and 36 other bacterial species, Claverie and colleagues identified only 71 'core genes' - genes that are present in all the species and that are therefore potential targets for broad-spectrum antibiotics. Of these 71 genes, 52 are linked to translation. From previous experience only about one in ten proteins is expected to be 'druggable' - that is, to be inhibited by a potential small-molecule drug. Claverie came to the disheartening conclusion that with only about seven druggable targets, there are few new broad-spectrum antibiotics on the horizon.

Bonnie Bassler (Princeton University, USA) works with a bacterium whose genome has not been sequenced and probably will not be sequenced in the near future. She studies quorum sensing in *Vibrio harveyi*, which communicates with members of the same and related species by emitting

light. Quorum sensing is the means by which bacteria assess population density. Despite the lack of comprehensive genomic sequence data in *V. harveyi*, Bassler and her colleagues have used comparative genomics to investigate the signaling systems that regulate light production in response to quorum-sensing signals. These systems consist of a sensor protein and an autoinducer, a small molecule that gives an indication of population density and that is recognized by the sensor protein. Bassler's group investigated *luxS*, a gene required for the synthesis of the autoinducer in the inter-species communication system in *V. harveyi*. They found homologs of *luxS* in more than 60 bacterial species, suggesting that inter-species communication is common in both Gram-positive and Gram-negative bacteria. They also identified the autoinducer from the discovery that the *Borrellia burgdoferi luxS* homolog is in an operon with two genes required for *S*-adenosyl methonine synthesis. This provided the key clue that led to discovery of a catalytic function for LuxS and identification of the autoinducer asfuranosyl borate diester.

Much of the meeting focused on the biology of *Saccharomyces cerevisiae*, an organism that has been in the vanguard of genomics and that is perhaps the only organism that has truly reached post-genomic status. Not only was it the first eukaryote to have its genome sequenced, its genome was nearly the first to be completed, coming in a close second to the tiny *Haemophilus influenzae* genome. André Goffeau (University of Louvain, La Neuve, Belgium), who initiated and husbanded the international yeast sequencing effort, observed how the availability of the yeast genome sequence, coupled with a community of labs poised to exploit the data, spurred the development of powerful tools for in-depth study of the yeast cell, such as genome-scale DNA microarrays, protein arrays, protein-interaction libraries and collections of gene deletions.

The power of these tools in yeast was demonstrated by several speakers. Jim Broach (Princeton University) described how his group has used DNA-microarray gene-expression profiling to determine the complete transcriptional output of signaling networks in response to glucose and other nutrients regulated by the small GTPase Ras and the phosphatidylinositol kinase Tor. He also showed that these kinds of experiment are sensitive enough to determine epistatic interactions between components of the signaling pathways.

Marc Vidal (Harvard Medical School, Boston, USA) presented an analysis of the comprehensive protein-protein interaction - or 'interactome' - maps of yeast and *Caenorhabditis elegans*. In the yeast interaction network, he found that 'hubs' (proteins that interact with many partners) are three times more likely than other proteins to be encoded by essential genes. He also distinguished two types of hub, which he called 'party' hubs and 'date' hubs. Party hubs interact simultaneously with many proteins, as determined by the similar patterns of expression of all the proteins involved. In contrast, date hubs interact with many different proteins at different times, as the proteins involved have very different patterns of expression. Removing the date hubs from the interaction network reveals sub-networks of functionally related proteins. Experiments like these are only possible, however, because the yeast genome has been extensively and accurately annotated. For other organisms, gene annotation is more difficult and time-consuming.

Comparative genome analysis is always helpful in the annotation of a genome, and several current comparative genome-sequencing projects are aimed at further characterization of *S. cerevisiae*. Claude Gaillardin (National Institute of Agronomy, Grignon, France) and Bernard Dujon (Pasteur Institute, Paris, France) described the sequencing of the genomes of four yeasts (*Candida glabrata*, *Kluyveromyces lactis*, *Debaryomyces hansennii* and *Yarrowia lipolytica*), representing major branches of the hemiascomycetes, and the comparison of these species to *S. cerevisiae*. These studies provide a broad overview of how *S. cerevisiae* relates to other yeasts in chromosomal number, repeat content (tRNAs, rDNA repeats and retrotransposons) and gene content.

Mark Johnston (Washington University, St Louis, USA) described comparative sequencing of five species within the *Saccharomyces* genus (*S. bayanus*, *S. castellii*, *S. kluyveri*, *S. kudriavzevii* and *S. mikatae*) with the goal of identifying conserved intergenic sequences that are likely to regulate transcription. Comparison of these closely related species has significantly improved annotation of *S. cerevisiae*. This type of study represents a new level of comparative genome annotation aimed at identifying all functionally conserved sequences, not just those for protein-coding genes.

The genomic era will persist for many organisms long after some have entered the post-genomic era. In the meantime, we should be sensitive to those who do not have a genome sequence to analyze.