Meeting report

Developmental biologists cast a net over sequenced genomes

Matthias Gerberding and Nipam H Patel

Address: Howard Hughes Medical Institute, University of Chicago, 5841 S. Maryland Ave., MC1028, Chicago, IL 60637, USA.

Correspondence: Nipam H Patel. E-mail: npatel@midway.uchicago.edu

Published: 24 September 2002

*Genome **Biology*** 2002, **3(10)**:reports4032.1–4032.3

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2002/3/10/reports/4032

© BioMed Central Ltd (Print ISSN 1465-6906; Online ISSN 1465-6914)

A report on the annual meeting of the Society of Developmental Biology, Madison, Wisconsin, USA, 21-24 July 2002.

The introduction of new concepts and of new techniques is equally important in advancing scientific research; ideally, they push each other forward. The talks at the recent Society for Developmental Biology (SDB) meeting gave ample evidence that technical progress continues to transform research in different areas of developmental biology. The last few years have seen remarkable progress in large-scale DNA sequencing, gene-expression analysis, and bioinformatics. The full genome sequence of several model organisms is now available, and new ones are being added continually. Another technical advance is the availability of microarrays and of serial analysis of gene expression (SAGE), and of methods to analyze the data they generate. Microarrays (and SAGE) measure levels of gene expression simultaneously for a large number of genes and, combined with the availability of whole-genome sequence data, have given us remarkable new tools for examining development. Several talks at the SDB meeting highlighted these advances and focused on new concepts of gene regulation and genetic networks that are beginning to emerge.

In the session on 'dealing with complexity', Stuart Kim (Stanford University, USA) described the kinds of new information that can be derived from a whole-genome approach to studying *Caenorhabditis elegans*. Microarray experiments can be used to analyze the expression and regulation of genes in different tissues and under different conditions. A number of different approaches have been used to isolate mRNA from specific tissues in cases where isolating the tissues by manual dissection is impractical. For example, in *Drosophila*, researchers have taken advantage of maternal-effect mutations to generate egg collections in which 100% of the embryos are mutant or have used green fluorescent protein (GFP) markers to allow automated sorting of specific embryos or cells. Recently, a technique has been developed called mRNA tagging, which allows mRNA to be labeled in living cells and subsequently to be isolated. A tissue-specific promoter drives the expression of a FLAG-tagged form of poly(A)-binding protein; all the mRNA in a desired population of cells is thus selectively tagged and can be isolated from the total mRNA of the animal. In this way, the Kim lab has been able to gather microarray data on specific expression in *C. elegans* tissues such as muscle cells, which are otherwise difficult to isolate. They have shown that the mRNAs of 1,364 genes are significantly enriched in nematode muscle. When mapped back to their chromosomal locations, the genes co-expressed in muscle were often found in clusters of 2-5 genes that cannot be explained as local gene-duplication events. This result suggests that the local chromosomal environment may play an important role in the regulation of muscle genes, perhaps as a consequence of chromatin opening.

The total amount of microarray data now available for model systems such as *C. elegans* has become staggering. Another innovation discussed in Kim's talk was the use of topographic mapping of array data to make it more accessible by presenting it in a useful visual format (Figure 1). This tool makes it far easier for the investigator to discover interesting links within the data and to examine the relationships between genes at many levels of resolution.

In the same session, Richard Young (Massachusetts Institute of Technology, Cambridge, USA) illustrated the power of a genome-wide analysis of the location of DNA-bound proteins on yeast chromosomes to aid understanding of genetic regulatory networks. In the initial experiments to assay genome-wide location, done in the Young lab two years ago, two DNA-binding proteins, Gal4 and Ste12, were co-immunoprecipitated with their respective DNA sequences

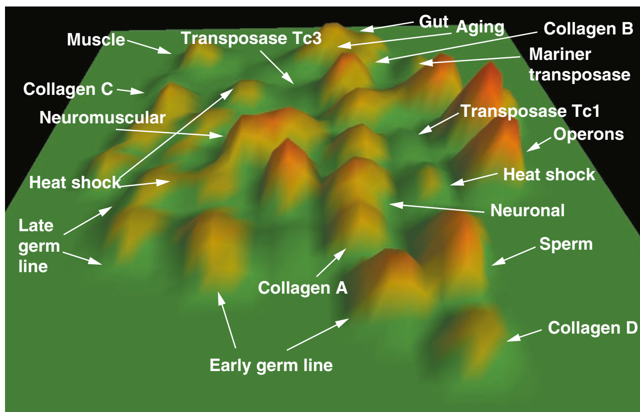

Figure 1
A 'topomap' of *Caenorhabditis elegans* gene expression. The co-regulation of genes is plotted in three dimensions; the z axis represents gene density and the position in the *x-y* plane is a relative measure of gene relatedness based on the microarray expression data. Selected classes of gene are enriched in specific 'mountains'. Image courtesy of Stuart Kim.

from cells grown under various conditions (such as different carbon sources and levels of pheromones). Using microarrays, it was then possible to determine which genes corresponded to the bound sequences. As a result, they were able to find new downstream targets of these two DNA-binding proteins that had not been detected by previous mutant screens. In subsequent experiments, this approach was expanded to the analysis of nine proteins that have roles in cell-cycle regulation. The fascinating picture emerging from this study is that transcriptional activators functioning during one stage of the cell cycle regulate transcriptional activators of the following stage. The data from these experiments are now being used to assemble models that depict large regulatory networks, and the approach will clearly be applicable to all systems. With knowledge of such networks, we can understand how different networks are inter-related and how different cellular functions are coordinated.

Also in this session, David Keys (University of California, Berkeley, USA) described the experimental design and early data from a screen for tissue-specific enhancers in the sea squirt *Ciona intestinalis*. Adult *C. intestinalis* could be easily mistaken for a plant or sponge, but as an embryo and larva *C. intestinalis* has a prominent notochord, a morphological structure also found in vertebrates, revealing its close relationship to this group of animals. The project is designed to allow a comprehensive analysis of the enhancer sequences of *C. intestinalis*. To locate the enhancers, genomic DNA of *C. intestinalis* is cut into pieces of about 1,500 base pairs (bp), cloned into a vector that allows potential enhancers to drive expression of the reporter gene *lacZ*, and introduced into embryos by electroporation. After just a few hours, the embryos can be assayed for *lacZ* expression and any tissue-specific patterns of expression can be recorded. The prediction is that the enhancers that drive similar patterns of expression will share sequence motifs that can be used to establish a '*cis*-regulatory code'. Because the genome of *C. intestinalis* has been sequenced, these putative enhancer sequences can also be associated with specific genes. The analysis is further improved by sequence comparisons with a closely related species, *Ciona savignyi*, whose genome has also been sequenced. These studies should also help us understand the evolution of enhancers, a central component of understanding evolution as a whole.

In the session on germ cells, Ruth Lehmann (New York University and Howard Hughes Medical Institute, USA) described a comprehensive approach to understanding the migration of germ cells in *Drosophila*. This particular avenue of research began several years ago with a mutagenesis screen that identified *Hmgcr* as a gene that mediates the attraction of germ cells towards the gonad. Somewhat surprisingly, *Hmgcr* encodes the HMG-CoA reductase enzyme, which is at the beginning of the well-studied metabolic pathway for cholesterol synthesis. Given that the entire *Drosophila* genome sequence is now available, Lehmann and colleagues were able to identify additional genes in the *Drosophila* genome that encode the other enzymes in the cholesterol biosynthesis pathway and to analyze their expression by *in situ* hybridization, to find out which parts of that pathway might also be necessary for germ cell migration. They found that the pathway in *Drosophila* lacks several of the terminal biosynthetic steps, explaining the long-known fact that insects cannot synthesize cholesterol. It will, of course, be of great interest to know exactly which product in the pathway is responsible for germ cell attraction.

In the same session, June Nasrallah (Cornell University, Ithaca, USA) presented an approach using a model plant, *Arabidopsis thaliana*, for the analysis of self-incompatibility, a process that prevents self-fertilization and ensures cross-fertilization in plants. Although self-incompatibility is absent from *A. thaliana*, it is present in a vast number of plant species, among them *Arabidopsis lyrata*, a species closely related to *A. thaliana*. Self-incompatibility in *A. lyrata* is mediated by stigma receptor kinases (SRKs) in the female stigma and their ligands, S-locus cysteine-rich protein (SCRs), in male pollen; the coding sequences of these two genes are among the most polymorphic known in eukaryotes. In *A. thaliana*, the SRK and SCR genes are truncated and nonfunctional, consistent with the absence of self-incompatibility. When a pair of the *SRK* and *SCR* genes from *A. lyrata* are transformed into *A. thaliana*, however, self-incompatibility is conferred on *A. thaliana*. These experiments demonstrate an elegant approach for cross-species studies and open up the analysis of self-incompatibility to the many genetic and genomic tools available for *A. thaliana*. They also present a fascinating case through which to explore how and why genes upstream in a genetic cascade can be lost during evolution while their downstream targets are maintained.

Modern developmental biology does not consist of just genome sequences and microarrays, nor are these even always the most exciting approaches. There is, however, no doubt that the availability of genomic tools is having a major impact on the field of developmental biology, and we can look forward to many exciting new breakthroughs as a result.