

Meeting report

## Systems biology of gene regulation fulfills its promise

Adrienne E McKee and Pamela A Silver

Address: Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA.

Correspondence: Pamela Silver. Email: [pamela\\_silver@hms.harvard.edu](mailto:pamela_silver@hms.harvard.edu)

Published: 22 May 2006

*Genome Biology* 2006, **7**:316 (doi:10.1186/gb-2006-7-5-316)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2006/7/5/316>

© 2006 BioMed Central Ltd

---

A report on the meeting 'Systems Biology: Global Regulation of Gene Expression' at Cold Spring Harbor, New York, USA, 23-26 March 2006.

---

A systems-level understanding of gene-regulation programs requires the synthesis of biological, computational, mathematical, and engineering approaches. One aim of a recent meeting on gene expression at the Cold Spring Harbor Laboratory was to promote this synthesis by bringing together experimentalists and computational biologists with a common interest in studying the organization and control of expression in complex biological systems. The presentations largely focused on identification and analysis of protein-DNA interactions, the discovery of *cis*-regulatory motifs, and the application of systems approaches to the study of post-transcriptional processes. Here we report on some of the remarkable experimental and computational advances in understanding gene regulation discussed at the meeting. A full list of abstracts is available at [[http://meetings.cshl.edu/meetings/abstracts/2006systems\\_absstat.html](http://meetings.cshl.edu/meetings/abstracts/2006systems_absstat.html)].

### Studying genome occupancy and protein-DNA interactions

Chromatin immunoprecipitation followed either by DNA hybridization to microarrays (ChIP-chip) or by DNA sequencing of paired end tags (ChIP-PET) are two widely used approaches to map protein interactions with the genome. One focus of the meeting was on the data obtained from these genome-occupancy studies. Several presentations showed that researchers are expanding the analysis of genome occupancy to tissue and developmental systems.

A strength of the ChIP-chip approach is its ability to define interaction sites for proteins with unknown targets. Peggy Farnham (University of California, Davis, USA) is exploiting

this to investigate the protein Suz12, a component of the Polycomb Group complex. Her group has not only isolated DNA targets of Suz12, but has also found that Suz12 can silence large regions of the genome in a cell-type-specific manner. The application of ChIP-PET to studying genome occupancy of key transcription factors in embryonic stem (ES) cells was the subject of a presentation by Huck-Hui Ng (Genome Institute of Singapore, Singapore). Ng and his colleagues have identified DNA targets of Oct4, Sox2, and Nanog. Among their findings is that promoter binding is not the rule; these factors are also present at intronic regions and target microRNAs (miRNAs).

Focusing on the genome occupancy of RNA polymerase II (PolII), Bing Ren (University of California, San Diego, USA) presented data from a large ChIP-chip study of five types of mouse tissue. He showed that PolII is largely located at two different genomic sites - promoters and putative enhancers. Data from his group also showed that most promoters are active in all the tissues analyzed and that these tissue-wide promoters, but not promoters that appear tissue-specific, are found near CpG islands. These findings have led to current investigations of chromatin signatures at human promoters with the goal of predicting promoters on the basis of histone modifications.

### Strategies and tools to investigate interaction and regulatory networks

Although an objective of most genome-location studies is to uncover the high-affinity interactions between proteins and DNA, these studies often generate many data on low-affinity interactions. Weak interactions often generate weak (but arguably significant) expression. Thus, much potentially useful information from ChIP-chip studies is largely ignored. Amos Tanay (Rockefeller University, New York, USA) discussed a computational strategy to garner information about low-affinity transcriptional interactions from ChIP-chip

datasets. This approach uses position-weight matrix regression to find new and previously characterized motifs in low-affinity interaction data. This algorithm has been used successfully with yeast datasets to find regulatory motifs.

An outstanding challenge in understanding gene regulation is to reliably identify the *cis*-regulatory element motifs that affect transcriptional and post-transcriptional processes. A further challenge is to integrate existing knowledge into approaches to the discovery of these motifs. Several speakers presented new computational efforts, web-based tools and wet-lab developments that focused on these challenges. Inherent in classical pattern discovery approaches is the problem of a high signal-to-noise ratio when searching for short, degenerate motifs in long spans of genomic sequence. Previously obtained information about the protein or sequence in question may help filter out the noise in the discovery of *cis*-regulatory motifs. This ability to use existing knowledge about sequence composition, phylogenetic footprinting, and factor binding to direct *cis*-regulatory element motif searches is the objective of a collection of software discussed by Wyeth Wasserman (University of British Columbia, Vancouver, Canada). The collection of online tools including PAZAR, a system for the collection and dissemination of regulatory sequence annotations, is available on the Wasserman lab's website [<http://www.cisreg.ca>].

Mathieu Blanchette (McGill University, Montreal, Canada) presented a large database of computationally predicted *cis*-regulatory element motifs (pCRMs) identified through a synthesis of human, mouse and rat transcription-factor binding sites [<http://genomequebec.mcgill.ca/PReMod>]. Known as PReMod (predicted regulatory modules), this dataset permits the evaluation of the distribution of the predicted motifs, thus providing an additional level of gene-regulatory information. For example, Blanchette showed that these motifs are enriched near 3' ends of genes and in regions far from genes. Scott Tenenbaum (University at Albany-SUNY, New York, USA) presented computational tools available at the Tenenbaum lab's bioinformatics tools website [<http://ribonomics.albany.edu>] for studying post-transcriptional regulatory elements. Among these is a collection of validated Training untranslated region (TUTR) datasets. These datasets comprise experimentally described RNA consensus sequences for use as blinded or non-blinded test sets.

### RNA networks and post-transcriptional programs

The regulation of gene expression extends well beyond transcription. Many groups presented data from studies designed to ask systems-level questions about tissue specification and mRNA localization. To appreciate how different tissues are established and how gene-expression networks act to specify similar yet distinct tissues, one needs to characterize the contributors that confer positional information.

John Rinn (Stanford University, USA) discussed the result of a large-scale positional expression study of more than 40 primary adult fibroblast cultures that map to the entire human body. Among the many intriguing findings he presented were data suggesting that expression signatures of hand and foot fibroblasts are more alike than those of hand and arm fibroblasts. In addition, he showed that specific Hox gene expression persists in adult tissues.

Large-scale, high-resolution investigations of the localization of gene expression in the *Drosophila* embryo were reported. Eric Lécuyer (University of Toronto, Canada) discussed findings from a genome-wide fluorescent *in situ* hybridization analysis of mRNA localization. Of the transcripts examined so far, more than 80% have an identifiable subcellular localization. In addition, Lécuyer's study has uncovered novel subcellular mRNA localization patterns. Soile Keränen (Lawrence Berkeley National Laboratory, Berkeley, USA) demonstrated a computational tool, Point-CloudXplore, which permits the analysis of morphology and gene expression at the cellular level. Data from confocal images of individual embryos stained for DNA and RNA or protein are converted into a data table. Hundreds of data tables can then be grouped to generate a virtual embryo onto which expression patterns of multiple gene products are resolved. These studies help to further our understanding of gene regulatory networks and provide a visual filter for directing additional gene-expression analyses.

Among the themes emerging from many of the RNA-centric studies is the power of posttranscriptional control in gene regulation. Lee Lim (Rosetta Inpharmatics, Seattle, USA) presented an elegant example of the application of a systems-level approach to a posttranscriptional process. He used microarrays to elucidate the effects of miRNAs on mRNA levels in HeLa cells, and his data show that miRNAs act as strong modulators of many different transcripts and have a broad effect on mRNA targets.

Gene-expression profiling affords investigators a view of steady-state mRNA levels. Unless experimental efforts are taken to broaden this scope, however, much of the resolution of mRNA expression dynamics is lost. To investigate the role of translational regulation in gene expression, Julia Bailey-Serres (University of California, Riverside, USA) compared the profiles of *Arabidopsis* mRNAs isolated from either polysomal or non-polysomal complexes. Her results point to a large discrepancy between populations of steady-state and actively translating mRNA, suggesting a direction for further investigation of the role of translational control in gene regulation. mRNAs in specific ribonucleoprotein (RNP) complexes were also the focus of Jack Keene's talk (Duke University, Durham, USA); he presented a study of mRNP populations following activation of Jurkat cells with mitogens. Specifically, messages encoding various RNA-binding proteins are differentially associated with HuR and PABP in

resting versus stimulated cells. The fact that Keene's group finds changes in the bound transcripts of RNA processing factors further highlights the importance of these proteins in gene-expression regulation.

### **The applications of systems-level data on gene regulation**

A significant ambition of the post-genomic era is to relate knowledge about gene regulation to outstanding questions of systems design and development. One of us (P.A.S.) presented recent synthetic biology efforts to use our present understanding of gene-expression programs to study other cellular processes. To this end, transcription-based logic is combined with protein localization and degradation to build cells that count mitotic divisions and ultimately act as a measure of cellular life span.

In a slightly different vein, Michael Levine (University of California, Berkeley, USA) reported that his group is using comparative genomics methods and knowledge of transcription networks to study how transcription factors drive organ development in the sea squirt *Ciona intestinalis*. His lab has developed a circuit diagram of the actions of transcription factors in the formation of the *Ciona* heart and has shown that perturbation of this 'heart network' can lead to the development of a multi-chambered heart rather than the normal simpler heart found in this organism.

It is now clear that the application of knowledge about gene regulation can further our understanding of the higher-level properties of complex systems. The meeting exemplified how the promise of systems-wide approaches is now being realized in full.