

Meeting report

Networks for all

Sebastian E Ahnert* and Sarah A Teichmann†

Addresses: *Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, UK. †MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK.

Correspondence: Sebastian E Ahnert. Email: sea31@hermes.cam.ac.uk

Published: 27 October 2008

Genome Biology 2008, **9**:324 (doi:10.1186/gb-2008-9-10-324)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2008/9/10/324>

© 2008 BioMed Central Ltd

A report on the Cold Spring Harbor Laboratory/Wellcome Trust conference on Network Biology, Hinxton, UK, 27-31 August 2008.

As molecular biology is driven by interactions between proteins, DNA and RNA, networks are a natural way to represent these systems. A recent network biology meeting in Hinxton was attended by scientists working on transcription networks and post-transcriptional gene regulatory networks, signaling networks, metabolic networks and contact networks in proteins and protein complexes. Here we discuss some highlights of the meeting, focusing on the newest research directions in the rapidly evolving field of network biology.

Transcriptional and post-transcriptional gene regulatory networks

Over the past decade, the role of microRNAs (miRNA) in genetic regulation has received much attention. Whereas some of the targets of miRNAs are now known, the mechanisms that regulate miRNA expression itself are very poorly understood. Marian Walhout (University of Massachusetts Medical School, Worcester, USA) addressed this important question by using *Caenorhabditis elegans* to construct the first genome-scale miRNA regulatory network that includes regulatory interactions of miRNA genes with transcription factors. In addition she showed that the presence of network motifs that contain both miRNA and transcription factors make it necessary to reconsider the relative network motif frequencies observed in transcriptional networks without miRNA, as the presence of miRNA nodes can increase the rate of information flow through the regulatory network. Eileen Furlong (EMBL, Heidelberg, Germany) presented work on the transcriptional network of mesoderm development in *Drosophila*. She is integrating chromatin immunoprecipitation and microarray (ChIP-chip)

time-course data with gene-expression profiles, including data from transcription factor mutants. This analysis revealed more complex combinatorial relationships than expected, including evidence for differential *cis*-regulatory module occupancy depending on different threshold concentrations at various stages of fly development.

Because of post-transcriptional effects, mRNA levels can be a poor indicator of transcription factor activity. Harman Bussemaker (Columbia University, New York, USA) described a way of detecting post-transcriptional modifications of transcription factor activity by using a statistical mechanics approach to predict expression levels from upstream regulatory sequence and by identifying chromosomal loci - activity quantitative trait loci (aQTL) - that affect transcription factor activity. More than a quarter of transcription factors appear to have at least one such aQTL, and in more than 90% of these cases the regulatory relationship would not be evident from mRNA expression experiments. This approach confirmed existing transcription factor regulations and also predicted a large number of novel interactions.

The fundamental question of whether transcriptional regulation is primarily determined by the genetic sequence itself or by its nuclear environment was addressed by Duncan Odom (CRUK Cambridge Research Institute, Cambridge, UK), who has studied hepatocytes from a strain of mice carrying a copy of human chromosome 21. The gene-expression program observed in these cells was almost entirely identical to that of human hepatocytes, leading to the conclusion that the primary responsibility for transcriptional regulation lies with the sequence, and that epigenetic effects are secondary.

Metabolic networks

In *Saccharomyces cerevisiae*, it has been established that 80% of genes can be knocked out without giving rise to a

phenotype. However, Guri Giaever (University of Toronto, Canada) showed that, in *S. cerevisiae*, 97% of genes exhibit a growth phenotype when perturbed by one of about 1,000 possible compounds and environmental stresses, suggesting that almost all genes are essential to growth in at least one particular condition.

Eytan Ruppin (Tel-Aviv University, Israel) introduced a computational approach for the development of tissue-specific metabolic models [<http://www.cs.tau.ac.il/~shlomito/tissue-net/>]. He has applied constraint-based modeling (CBM) to a combination of tissue-specific expression data and existing interaction data for metabolic networks. The CBM approach finds a network that is consistent with all input data, and reveals that as much as 18% of all human metabolic genes are involved in post-transcriptional regulation. Furthermore, the derived metabolic networks were shown to be highly tissue-specific.

From network to protein properties

In his keynote talk, Pawson used three-dimensional protein structures of Eph receptor tyrosine kinases to illustrate how allostery can occur within a single polypeptide chain. Using three-dimensional protein structures, he demonstrated an example of allostery within a single polypeptide chain through interactions between an SH2 and kinase domain. Wendell Lim (University of California, San Francisco, USA) presented a domain-based analysis of signaling, demonstrating that the choanoflagellate *Monosiga brevicollis* has SH2 and cadherin domains, previously thought to be limited to multicellular animals.

Anne-Claude Gavin (EMBL, Heidelberg, Germany) reported a new adaptation of affinity purification and mass spectrometry to study homomeric protein complexes isolated from a *Mycoplasma* species. In a first pass, this method identified a lower bound of 10% of such complexes, which consist of multiple molecules of the same protein. One of us (SAT) continued the theme, showing from a bioinformatics analysis of all proteins of known three-dimensional structure and from SwissProt annotations of *Escherichia coli* and human proteins that about two-thirds of proteins occur as homomers. She showed that homomers of dihedral symmetry have interfaces of different sizes, and that the larger interfaces are those conserved in evolution and in assembly intermediates. An example of this is the hexameric enzyme ATP sulphurylase, which assembles via a dimeric intermediate corresponding to the trimer of dimers predicted from the hierarchy of interface sizes evident from the three-dimensional structure.

Radek Szklarczyk (Radboud University, Nijmegen, Netherlands) traced various scenarios of how paralogous proteins interact with different partners. He and colleagues have found that paralogs often act as mutually exclusive,

condition-dependent subunits of different variants of the same complex, for example, RSC1/RSC2 of the RSC chromatin remodeling complex. Tanja Kortemme (University of California, San Francisco, USA) aims to re-engineer the interfaces between proteins to generate novel specificities or to abolish interactions of proteins with multiple interaction partners. One method she described for doing so was to map the individual residues involved in contacts between the different interaction partners of a protein, and to introduce mutations targeted towards residues specific to one interaction partner only.

In his presentation, Eli Eisenberg (Tel-Aviv University, Israel) highlighted the effect of relative protein concentration levels on the assembly of a protein complex. The concentration levels of a set of proteins forming a complex tend to be similar, and they also change in similar ways in response to environmental influences. Moreover, the fluctuations of concentration levels are found to be small for proteins in large complexes, or if the protein appears in multiple copies, and for the least abundant protein in the complex. Eisenberg reported that all these features can be shown to increase both the efficiency of protein assembly, as well as the robustness of the assembly process in the face of stochastic fluctuations.

Long Cai (California Institute of Technology, Pasadena, USA) described the behavior of the calcineurin-responsive zinc finger transcription factor Crz1 in *S. cerevisiae* in response to increasing calcium concentration. He showed that Crz1 is localized to the nucleus in bursts a couple of minutes in duration, and that the frequency of these bursts is proportional to calcium concentration. The consequence of this is that target promoters are activated according to the time the transcription factor spends in the nucleus.

Genetic networks

The work of Fabio Piano (New York University, USA) centers on the transition of oocytes to early embryos. Until now, most of the insights into this process have been gathered by studying its various components, such as fertilization, cell cycle, the establishment of cell polarity and cytokinesis, separately. Piano's aim is to describe these processes as functional modules of a larger interaction network by deriving a domain-based interactome network of proteins involved in *C. elegans* early embryogenesis. This network is more complete than previous networks of this kind, and reflects the modular organization of protein folding domains. This perspective can also be used to explain the robustness and evolvability of these functional units.

Trey Ideker (University of California, San Diego, USA) and colleagues are the developers of the widely used network processing and visualization software Cytoscape [<http://www.cytoscape.org/>], for which there now is a growing

community of independent plug-in contributors. Ideker described his team's efforts to integrate genetic and physical interactions into comprehensive regulatory networks. Of particular interest was his attempt to find an estimate of the number of times that each regulatory interaction would have to be sampled for a comprehensive network to emerge. This is analogous to the '5x' rule of DNA sequencing, which states that a genome needs to be sequenced at least five times to obtain a reliable dataset of the entire sequence. By assuming high false-negative rates and low false-positive rates, and by requiring that 95% of all interactions be found, with a false-discovery rate of less than 5%, Ideker arrived at factors of around 16x, but also showed that this figure can be reduced significantly under less simplistic assumptions. A realistic estimate is therefore likely to be on the same order of magnitude as the 5x rule for genome sequencing.

Eric Schadt (Rosetta Inpharmatics, Seattle, USA) showed that by comparing gene- expression patterns in different tissues, for example, adipose, liver, muscle and hypothalamus tissues in mice, genes that are co-expressed with genes in other tissues can be identified. Novel interaction networks that include these co-expressed genes in the different tissues can be derived that are independent of known genetic regulation within the tissues. These relationships between tissues also show how the subnetworks inside several different tissues influence each other. Schadt noted that this approach can be used to reveal interdependence relationships between treatments of different diseases; that is, treatment for one disease can exacerbate another, such as, for example, between obesity, diabetes and hypertension. Therefore, such diseases are likely to be the result of complex inter-tissue interactions in the first place.

This meeting demonstrated that the term 'Network Biology' encompasses a very broad range of topics and pervades many areas of current biological research. It is, therefore, likely that in future years, networks will be viewed more and more as the fabric that underlies much of biology, rather than as the subject of a distinct discipline called 'Network Biology'. The next meeting in this series will take place in Cold Spring Harbor in March 2009. Thereafter, annual meetings will be held in March, alternating between Hinxton and Cold Spring Harbor.