

MEETING REPORT

Grand network convergence

Joel S Bader*

Abstract

A report of the Systems Biology: Networks meeting, Cold Spring Harbor, USA, 22-26 March 2011.

The success of the human genome project has provided a model for an analogous interactome project to map how proteins, genes, metabolites and other regulatory components interact to transform a biochemical soup into a living system. These maps promise to serve as a framework for models that predict how a biological system responds to a perturbation or an input, which is relevant to gene mutations and therapeutic treatment in human disease, and as a framework for designing new systems in synthetic biology.

Three major themes arose during the 2011 meeting: technological drivers and data generation, algorithmic advances, and convergence on biological applications with context-sensitive networks.

Technological drivers and data generation

Many recent biomedical advances are being driven by technological advances. Advances in DNA sequencing technology are paralleled by advances in network mapping technologies, although network mapping may be more complicated because the biochemical species (proteins, metabolites, RNA and small molecules) are diverse compared with genome sequencing (DNA only). Although knowledge of networks is far from complete, the numbers of unknown interactions are moving from 'unknown unknowns' to 'known unknowns'.

Physical interactions continue to be of great interest. Protein-protein binding interactions are being systematically mapped using mass spectrometry of protein complex components (Anne-Claude Gavin, EMBL, Heidelberg, Germany), and they are continuing to reveal interactions not anticipated by any existing data. In an advance that could revolutionize the yeast two-hybrid

system, next-generation sequencing is being incorporated as the back-end read-out (Pascal Braun, Dana Farber Cancer Institute, Harvard University, USA). Steady advances over the past several years have developed the two-hybrid system to the point that the false-positive rate is very low, with precision of high-throughput screens roughly equivalent to careful, small-scale studies. The significance of the next-generation sequencing application is that the coverage or true-positive rate, which in previous work has been low, could conceivably be increased to approach moderate to near full coverage of interactions amenable to two-hybrid assays. Enhanced yeast one-hybrid systems are also providing increased coverage of regulatory interactions between transcription factors and DNA (Marian Walhout, University of Massachusetts Medical School, USA).

New technologies are opening up the ability to probe unexplored types of interactions. Lipids and interactions between membrane-localized proteins have been difficult to study using traditional methods. New protein-lipid binding assays are becoming available for medium-scale applications (Gavin). Cell signaling networks are being mapped using membrane two-hybrid technologies (Igor Stagljar, University of Toronto, Canada).

Microarray technologies continue to be adapted to mapping biological interactions. Interactions between transcription factors and DNA using universal DNA probes have become highly reliable. Much like protein structure pipelines that have increasingly focused on discovering novel folds, protein-DNA binding assays are being focused on the transcription factors that are most likely to have novel binding motifs that cannot yet be predicted by homology (Timothy Hughes, University of Toronto, Canada). Microarrays of spotted proteins provide continuing opportunities for novel functional screens, such as mapping kinase-substrate interactions at the genome scale (Heng Zhu, Johns Hopkins University, USA).

A final theme of new technologies is a push to measuring interactions and activities in living systems. A recent single-cell mass cytometry technology allows simultaneous measurement of about 30 parameters about a cell, including surface and functional markers (Gary Nolan, Stanford University, USA). The resulting data provide a dynamic view of cell development and an indication of drug activity.

*Correspondence: Joel S Bader joel.bader@jhu.edu
Department of Biomedical Engineering and High-Throughput Biology Center,
Johns Hopkins University, Baltimore, MD 21218, USA

Algorithmic advances

A growing number of statistical methods use network data to link a biological input to an output (phenotype). In linear systems, given two out of three of input, system (transfer function) and output, we can reproduce what is missing. Being able to do the same for biological systems would have great utility in predicting disease risk, developing new therapeutics, and so on. Biology is more complicated because inputs and outputs are ill-specified, and knowledge of the system (network) is poor. Nevertheless, network data are sufficiently complete that they are proving useful in linking biological inputs and outputs.

An important new direction in algorithmic development is the integration of multiple data sources to provide a fuller picture of cellular activity. This is especially important in studying multiscale processes, such as animal development, in which protein-level interactions translate to patterns visible by eye. Imaging data are now being harnessed to improve the inference of developmental regulatory pathways, with predictions validated by mutant studies (Nicholas Luscombe, EMBL European Bioinformatics Institute, UK).

Even single-cell dynamic processes have been difficult to study because technologies that measure networks typically provide a static picture, requiring additional dynamic measurements to understand how network components change and networks reorganize over time. I described new methods for coupling interaction networks with transcription time series to provide a moving picture of network activity.

Data integration methods provide improved ability to predict disease outcomes (Kelvin Zhang, University of California, Los Angeles, USA, and Ontario Institute for Cancer Research, Canada), predict gene function (Quaid Morris, University of Toronto, Canada), and map regulatory networks (Sushmita Roy, Broad Institute, USA). Combining data from genetic interactions, physical interactions and protein sequence provides a more accurate picture of how networks have evolved (Amy Keating, Massachusetts Institute of Technology, USA, and Chad Myers, University of Minnesota, USA).

Network-based algorithms can assist in generating hypotheses about how a gene mutation leads to disease (Theresa Przytycka, National Institutes of Health, Bethesda, USA, and Patrick Aloy, Institute for Research in Biomedicine, Spain). Metabolic models have been developed that can link biochemical measurements, such as metabolite uptake, to growth rate (Zoltan Oltvai, University of Pittsburgh, USA).

Convergence on biological applications

Cells can have identical DNA but be very different because they express different genes. These different contexts imply the existence of different network

components (proteins in signal transduction or gene regulation, microRNAs, metabolites, lipids and small molecules) and different network states. Networks are usually not measured for a specific context, however, but rather through biochemical assays (protein-binding microarrays, and yeast one-hybrid and two-hybrid systems) or by superimposing data from many distinct conditions (chromatin immunoprecipitation with microarrays (ChIP-chip) and with sequencing (ChIP-seq), epistatic interactions, condition-specific pull-downs, or time series). Several groups presented work showing how network state can be inferred by integrating heterogeneous datasets and how differences in network state correspond to differences in phenotype.

Recent work has demonstrated that network contrasts, defined as differences in interaction patterns measured in different conditions, can be more informative of biological processes than interactions measured in individual states. Contrasts can be generated through several types of perturbations. Yeast genetic interaction screens have used small-molecule treatments to generate contrasts (Trey Ideker, University of California, San Diego, USA). Contrastive analysis in cell signaling using a combination of genetics and small molecules provides insight into pathways relevant to leukemia (Thomas Graeber, University of California, Los Angeles, USA). Host-pathogen networks can be probed as a function of the pathogen genotype; studies of human papilloma virus showed that high-risk and low-risk strains had interactions with different subsets of human host proteins (David Hill, Dana-Farber Cancer Institute, Harvard University, USA).

These observations may be valuable in human health, for example by identifying patient-specific differences in network state (Anna Goldenberg, University of Toronto, Canada). An intriguing possibility is that some of the heterogeneity in network state may not be genetic but rather purely stochastic (Suzanne Gaudet, Dana-Farber Cancer Institute, Harvard University, USA).

Beyond observing network activity is the challenge of shaping network state. Future drug treatments may involve perturbing a network to control the response to drug treatment (Michael Yaffe, Massachusetts Institute of Technology, USA), or using computational techniques to identify key components of disease pathways (Andrea Califano, Columbia University, USA). Finally, new synthetic biology technologies are providing an entirely new capability to rebuild biological systems from the DNA up, with exciting applications to human health and bioenergy (James Collins, Wyss Institute, Boston University, USA).

Networks are providing a clearer picture of the structure of biological systems. Individual datasets, focused on distinct types of interactions, are sufficiently complete to provide a coherent, though not yet seamless, framework of cellular behavior, from ligand-receptor interactions to

cell signaling to transcriptional response. Although these networks are often described as 'wiring diagrams', in reality the ability to predict the behavior of a biological system or to design a new system remains at the cusp of systems biology research. Transforming network maps to functional models is the crucial challenge for systems biology.

Published: 17 June 2011

doi:10.1186/gb-2010-12-6-306

Cite this article as: Bader JS: Grand network convergence. *Genome Biology* 2011, **12**:306.