Meeting report

Signals and systems

Nevan J Krogan* and Timothy R Hughes†

Addresses: *Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143, USA. †Banting and Best Department of Medical Research, University of Toronto, Toronto, ON, M5S 3E1, Canada.

Correspondence: Timothy R Hughes. Email: t.hughes@utoronto.ca

---

A report on the 2006 Keystone Conference on Signaling Networks, Vancouver, Canada, 30 January-4 February 2006.

The boundaries between traditional notions of cellular signaling and more genomic and systematic approaches to biology are becoming increasingly blurred, and the recent Keystone conference on signaling networks reflected an expanded view of signaling. It is becoming increasingly accepted that genes, proteins, cells, and organisms function as components of larger systems, rather than independent activities contributing to a single defined outcome, and many presentations at the conference reflected this. If there was a single theme, it was the heavy reliance on technical approaches in functional genomics, proteomics, and computational biology, such that conceptual and technical discussions often dominated the resulting biology.

RNA interference screening

The impact of technology on the study of signaling networks was most evident in the widespread application of RNA interference (RNAi) screens. Screens to study signaling networks in multicellular organisms using RNAi technology are being performed with success, although several hurdles clearly remain. One of these is the apparently high rate of false positives. In his keynote lecture, Norbert Perrimon (Harvard Medical School, Boston, USA) predicted an uncomfortably high rate of false positives due to off-target effects, at least in *Drosophila*. Consistent with this, Phil Beachy (Johns Hopkins University School of Medicine, Baltimore, USA) described an ongoing RNAi screen in *Drosophila* looking for proteins involved in the Wingless signaling pathway. Following efforts to study a previously uncharacterized gene identified by this screen, it was noted

that there were 16 bases in the interfering RNA that were identical to *armadillo*, a known gene in the pathway, suggesting that it is an unanticipated off-target that could completely explain the phenotype conferred by the interfering RNA. Perrimon proposed that incorporating other types of data, such as protein-protein interaction information, can help decipher which hits are physiologically relevant. Rene Bernards (Netherlands Cancer Institute, Amsterdam, The Netherlands) commented that reporting lists of unconfirmed hits from genome-wide RNAi screens in humans might be detrimental to future work by creating distractions from more productive lines of research, and that scientists (and journals) perhaps need to be more mindful of this.

Despite these limitations, work presented on *Caenorhabditis elegans* showed unquestionably that useful biological information could be extracted from RNAi screens. Gary Ruvkun (Massachusetts General Hospital, Boston, USA) and colleagues have been involved in a variety of screens looking for factors involved in lifespan, molting, defects in fat storage, and the RNAi pathway itself. Ruvkun described the complete list of hits from the molting screen as a "rogues' gallery" of phenotypes and molecular functions. One could argue that this might be evidence not so much of a high false-positive rate as of the dependence of a system on many components and processes; a corollary of this is that perturbation of an individual gene can have a myriad of physiological effects. Andrew Fraser (Wellcome Trust Sanger Institute, Cambridge, UK) presented work that supports this view. His group has been performing synthetic interaction screens in *C. elegans* in which previously characterized worm mutants defective in the epidermal growth factor (EGF)-Ras-Raf-MAP kinase pathway are treated with libraries encoding double-stranded RNAs, in order to identify synthetic genetic interactions. A significant outcome of this work was that RNAi of genes in certain functional categories results in synthetic effects with a high proportion of all mutants tested,

suggesting that some cellular functions buffer or canalize many physiological traits, rather than perform a single physiological function.

David Sabatini (Massachusetts Institute of Technology, Cambridge, USA) presented a thorough overview of technical aspects of the development and use of genome-scale lentiviral RNAi libraries for human and mouse. A consortium of groups in the public and private sectors is contributing to the creation of the libraries, which are sold by both Sigma (St. Louis, USA) and Open Biosystems (Huntsville, USA) and are available as glycerol stocks, plasmids and viruses. Sabatini commented that while this effort is still incomplete, some functional categories (for example, transcription factors) are fairly comprehensively covered. The efficacy of the system has been validated by targeting tyrosine kinases, and Sabatini reported that 90% of these genes have at least one target that results in a significant knockdown of expression. Sabatini also discussed issues of high-content screening, such as image archiving (a screen can generate 1 Tb of images), image analysis, and defining a 'hit'. He described Cell Profiler [http://jura.wi.mit.edu/cellprofiler/], a free, open-source software package for analysis of thousands of cell images, which among other statistics can output cell counts, DNA content and mitotic index.

Sabatini gave his view on how to define true positives: as a rule of thumb, he suggested requiring two independent hairpins that work (the consortium library contains several for each gene), dose dependence, and complementation if possible. Finally, he noted that as antibodies are not available for the majority of human proteins, it is currently difficult to ascertain whether protein levels are impacted by RNAi. On a similar theme, Sumit Chanda (Genomic Institute of the Novartis Research Foundation, San Diego, USA) pointed out that 50% of work published in humans corresponds to less than 10% of the genome, with few (or no) publications for most known and predicted genes, which provides a justification for comprehensive screening efforts.

Chemical screening
Chemicals have long been used to perturb specific biological activities in experiments, and they have a particular appeal because of their potential application as drugs. Justin Lamb (Broad Institute, Cambridge, USA) described an ongoing effort to construct a transcriptional connectivity map for biomedical discovery, which relates microarray gene-expression profiles that are due to drug effects, disease, and gene perturbations. An initial goal is to find off-target uses for existing drugs; Lamb's group is currently analyzing 1,500 drugs approved by the US Food and Drug Administration (FDA), 500 non-drug bioactive compounds used as research tools, and the effects of the knockdown of 500 potential drug targets using lentivirally delivered short hairpin RNAs (shRNAs). A distinguishing feature of the connectivity map

is the use of the Kalmogorov-Smirnov statistic to score a relationship between two experiments, even in the absence of an overall correlation. Lamb presented evidence from these analyses to indicate that connections can be drawn between drugs and diseases even if the experiments are performed in different cell types, or possibly even different organisms, which is significant because large-scale RNAi screens are typically performed in well behaved cultured cells.

An alternative assay for chemical activities was presented by Soren Jensby Nielsen (BioImage, Soeborg, Denmark), who described screening compounds for their impact on protein translocation. He reported the effects of around 50 compounds on the localization of a dozen different green fluorescent protein (GFP) fusion proteins. A surprising number of effects were observed, and although effects appear to be protein-specific (that is, they do not correlate with growth inhibition or other general effects), Nielsen stated that this screening approach has not yet identified any primary targets, indicating that the observed mislocalizations are probably downstream effects of perturbing the primary target. Nevertheless, this is an intriguing approach to screening, especially as it requires only a relatively basic setup compared with high-content assays.

Protein-interaction networks
Giulio Superti-Furga (CeMM Center for Molecular Medicine, Vienna, Austria) and one of us (N.K.) both described efforts to purify protein complexes for the entire proteome of *Saccharomyces cerevisiae*. The two groups used a similar strategy involving the immunoprecipitation of affinity-tagged proteins followed by mass spectrometry. Although proteomic studies have been previously published in budding yeast, the current studies attempt to tag and purify approximately 10 times as many unique proteins than any previous study. Both studies have just been published and comparison of the datasets with each other and with others previously generated should allow the compilation of the most comprehensive protein-interaction map ever generated for an organism.

A next goal will be to define similar physical interaction maps in higher organisms. Marc Vidal (Dana-Farber Cancer Institute, Boston, USA) presented work on uncovering interacting proteins in human, focusing on two-hybrid analyses that are then confirmed by pulldowns. Like Perrimon, Vidal was enthusiastic about weighting the significance of interactions on the basis of multiple independent data types, most notably phenotypic profiles.

Pathways and networks
The term signaling pathway implies a cause-and-effect relationship among a series of steps. Two talks concerned the development of computational methods for the inference of cause and effect, or for determining the importance

of the nodes in a network, from quantitative genetic and phenotypic analysis. Eric Schadt (Rosetta Inpharmatics, Seattle, USA) pointed out that in genetic crosses information can only flow from genotype to phenotype. In contrast, gene-expression patterns can be either a cause or an effect of phenotypes. He gave a conceptual overview of the statistical approaches his group is using to distinguish whether gene-expression phenotypes are upstream or downstream of other phenotypes in large mouse crosses. In summary, one can ask which network structure is most consistent with the observed correlations, with quantitative trait loci used as an anchor. Alon Kaufman (Hebrew University, Jerusalem, Israel) presented algorithms for calculating the contribution of genes to a function that has been quantified via set of multiple mutants in factors that influence that function. These algorithms, known as the 'multi-perturbation shapely value analysis' and the 'functional influence network' algorithms, can recreate some features of known networks and should aid in predicting aspects of unknown pathways.

Analysis of the consequences of multiple perturbations has historically been a primary approach to dissecting genetic and signaling networks. On the other hand, these computational analyses require a collection of quantitative genetic data. Two talks described methods for generating such data on a large scale in budding yeast. Jef Boeke (Johns Hopkins University, Baltimore, USA) described a comprehensive genetic analysis of approximately 70 genes involved in DNA integrity. Analysis of the genetic patterns enables the establishment of 'congruence groups', functionally related clusters that share the same pattern of synthetic sick or lethal (SSL) interactions, but may not participate in synthetic sick or lethal relationships among themselves. One of us (N.K.) described the generation of an epistatic miniarray profile (a quantitative analysis of double-mutant growth phenotypes among a logically-selected subset of genes) of approximately 850 genes involved in most aspects of chromosome function. This analysis detected not only synthetic sick or lethal interactions, but also cases in which double deletions grow better than would be expected from the growth of each single mutant when crossed to most other mutants on the array. The genes involved in these latter interactions often displayed similar sets of genetic profiles and correlated well with physical interactions. Will RNAi enable the generation of such networks in higher organisms? Perhaps to some degree; however, in the question-and-answer period, Fraser suggested that his group had encountered a high false-negative rate in RNAi x RNAi 'crosses' in *C. elegans*, relative to RNAi x germline mutation, presumably because most RNAi knockdowns are incomplete.

Biological networks tend to have distinctive mathematical properties, most notably a scale-free degree distribution, the cause of which has been debated extensively. Ravi Iyengar (Mount Sinai Medical Center, New York, USA) and Albert-Laszlo Barabasi (University of Notre Dame, South Bend, USA) both presented arguments and modeling results supporting the notion that the origin of scale-free network structures is likely to be a consequence of selection pressure following gene duplication. Briefly, the idea is that newly duplicated genes have increased chances of survival if they develop a function that is tied to that of a gene that is already functionally coupled to many other genes, creating a 'rich get richer' scenario. The results of simulations of this paradigm appear to match experimentally derived network structures.

In contrast to all the network technology, a biological highlight was the description by Cori Bargmann (Rockefeller University, New York, USA) of her lab's demonstration that *C. elegans* can learn to smell which bacteria it should and should not eat, using a network of only a few hundred neurons. Overall, the conference successfully integrated a broad range of biological topics, including novel technologies and computational strategies that biologists can apply to obtain meaningful systems-level biological information on signaling networks.