Knocking sense into regulatory pathways

Guri Giaever & Corey Nislow

Simultaneous targeted perturbations illuminate the structure and function of regulatory networks.

The complex architectures of cellular signaling pathways are beginning to yield their secrets thanks to sophisticated combinations of experimental and computational tools. Two recent papers describe complementary new approaches to identifying the relationships between signaling components (kinases and phosphatases), the transcription factors they regulate and subsets of target genes. Both studies exploit the possibility of introducing multiple, simultaneous perturbations into model systems: Bakal et al.[1], in Science, study RNA interference of two genes simultaneously in Drosophila cells, and Capaldi et al.[2], in Nature Genetics, carry out double and triple genetic knockouts in yeast. Both groups use their perturbation data to order components in a signaling pathway by epistasis analysis. Yet each dissects a different level of the regulatory hierarchy—kinase cascades in Bakal et al.[1] and transcriptional circuits in Capaldi et al.[2]—illustrating the power and versatility of multiperturbation strategies.
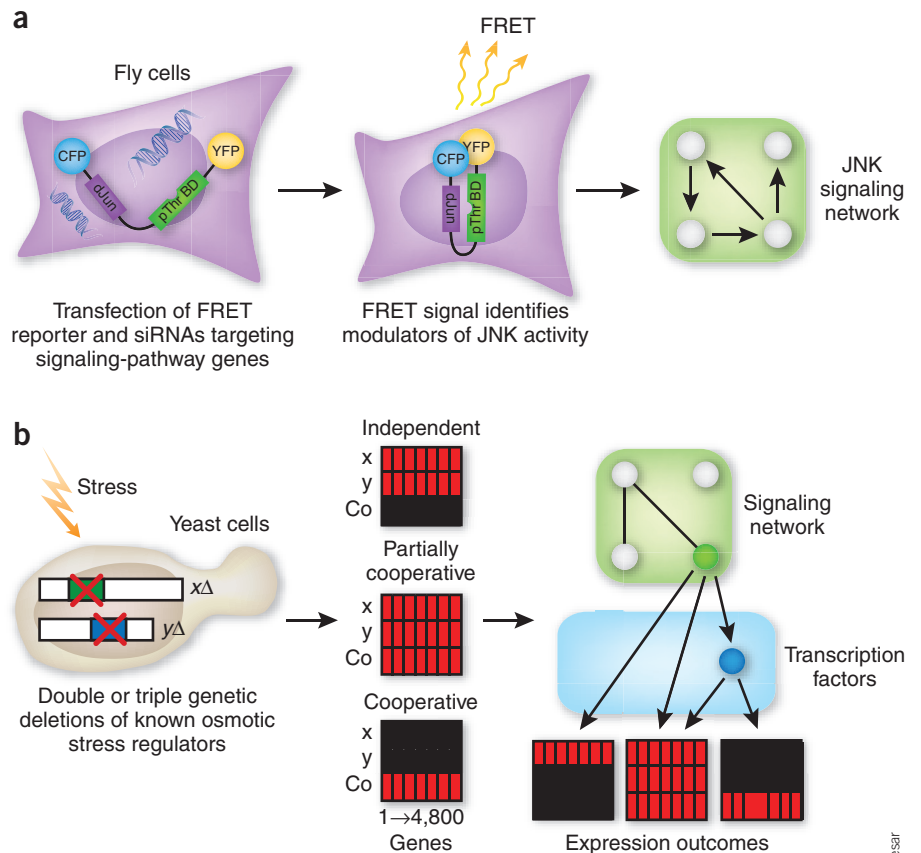
A grand challenge of genetics and systems biology in the next decade will be to integrate vast amounts of new data on gene and protein functional interactions into frameworks that define cellular pathways. Specifically, we will need to understand how cellular parts assemble into pathways, how multiple pathways are coordinated and how pathways are insulated and integrated to form a functioning cell[3]. Pathways do not generally act linearly, nor do they act in isolation. To dissect pathway architecture genetically, it will be necessary to perform epistatic analysis on data derived from experiments that employ multiple mutations or knockouts. This will be a challenging task, regardless of the phenotypic readout. Furthermore, the derived pathways might not reflect physical interactions between pathway components but rather a 'logical' architecture—that is, there may be more than one pathway that can predict the observed phenotypes.

Bakal et al.[1] used a Förster resonance energy transfer (FRET)-based reporter of phosphorylation of the JUN NH₂-terminal kinase (JNK)

to screen the effects of 1,565 double-stranded (ds) RNAs that target all known and predicted Drosophila kinases, phosphatases, regulatory subunits and adapters (**Fig. 1a**). Although 24 known and novel regulators of JNK phosphorylation were identified, several well-known regulators were not. To address this problem, the researchers screened the 1,565 dsRNAs against 12 cell cultures sensitized by simultaneous transfection with a second dsRNA that targeted

a canonical component of the JNK pathway. The resulting 17,724 (~1565 × 12) double perturbations identified 55 additional regulators of JNK. Next, they applied an integrative network algorithm that incorporates genetic and phosphoproteomics data to construct a JNK phosphorylation network, allowing them to propose a model of the architecture of JNK signaling.

This study is impressive in the number of double dsRNA combinations tested and in the

Guri Giaever and Corey Nislow are at the Terrence Donnelly Centre for Cellular and Biomedical Research, 160 College St., University of Toronto, Toronto M5S3E1, Ontario, Canada. e-mail: guri.giaever@utoronto.ca, corey.nislow@utoronto.ca

Figure 1 Multiple-perturbation experiments enable reconstruction of signaling pathways. (**a**) Strategy used by Bakal et al.[1] to identify the signaling network regulating JNK activity. Drosophila cells containing a dJUN-FRET sensor-reporter of JNK phosphorylation activity are transfected with one or two dsRNAs (left panel). FRET signals identify dsRNAs that modulate JNK activity (middle panel). A probabilistic computational framework is then applied to reconstruct the phosphorylation signaling network (right panel). (**b**) Strategy used by Capaldi et al.[2] to predict the transcriptional activation network of the Hog1 MAPK pathway. Yeast mutants with combinatorial deletions of genes known to be involved in the Hog1 pathway are generated (left panel). Gene expression profiles of the mutant strains are analyzed to derive the effects of individual deletions (labeled x and y) as well as the 'cooperative effect' (labeled 'Co') of two deletions together on all genes in the genome (middle panel). Genes are then clustered by whether x and y regulate them independently, partially cooperatively or cooperatively. The resulting modes of regulator interaction allow the order of the genes in the pathway to be inferred (right panel).

blending of experimental and computational approaches, but its scale also hints at the magnitude of the work that remains. Surely, additional double dsRNA screens beyond the 12 performed will reveal more JNK regulators. Moreover, there is room to optimize experimental protocols for dsRNA knockdown (e.g., in sequence selection and delivery) and for measuring the degree of silencing of each gene and off-target effects. Finally, the studies of Bakal et al.[1] were all carried out with unstimulated *Drosophila* cells. Given JNK's known roles in maintaining cell, tissue and organism fidelity in the face of cellular stress, additional experiments will be needed to determine if and how the architecture of the JNK network is affected by stress.

JNK is known to exert its influence through numerous transcription factors, which in turn regulate a diverse set of target genes. How might the influences of the JNK regulators extend into the downstream transcriptional program? Capaldi et al.[2], working with budding yeast, suggest a way of tackling this question. The authors focused on building a quantitative model of the Hog1 MAPK-dependent pathway, which regulates the osmotic stress response. They performed gene expression profiling experiments on single-, double- and triple-knockout mutants of *hog1*, *msn2/4*, *sko1*, *sok2* and *hot1*—all known components of the Hog1 pathway.

Their key insight was to computationally tease apart the effects of knocking out single genes from what they call a "cooperative component," which quantifies whether two genes function independently, cooperatively (epistasis) or partially cooperatively (**Fig. 1b**). Perhaps most importantly, their method computes the cooperative component for each gene in the genome, providing a fine-grained view of how pairs of regulators interact functionally over the entire genome. The resulting regulatory map shows, for the first time, how the Hog1 MAPK signal propagates through different combinations of transcription factors to regulate distinct subsets of genes. By applying their analytical method to expression profiles of salt-versus glucose-induced osmotic stress, Capaldi et al.[2] also suggest how different branches of the regulatory hierarchy are used in a context-dependent manner to respond to different types of osmotic stress.

In contrast to the Bakal study, Capaldi et al.[2] use complete knockouts of components of a well-understood pathway. These precise deletions have the advantage of being invariant from cell to cell and assay to assay, but have the disadvantage that they cannot be considered essential genes. An important feature of both studies is that they rely on phenotypic readouts other than fitness to define pathway architecture. Comparing the networks derived from multiple phenotypic measures should greatly expand the 'dynamic range' of network biology.

As these two studies show, the analysis of multiple mutants and multiply perturbed cells provides crucial information for reconstructing the 'wiring diagram' of the cell, but the impact goes further. Drugs, for example, are simply cellular perturbations that can be conditionally applied. This concept was recently explored by chemically perturbing multiple mutants to enrich for genetic interactions and to order the components in a DNA repair pathway[4,5]. In a clinical context, a patient taking a therapeutic agent represents a unique genetic background combined with a chemical perturbation. Thus, it is conceivable that the approaches discussed here might eventually help guide the analysis of complex perturbation experiments in therapeutic settings.

1. Bakal, C. *et al.* *Science* **322**, 453–456 (2008).
2. Capaldi, A.P. *et al.* *Nat. Genet.* **40**, 1300–1306 (2008).
3. Hartwell, L. *Nature* **387**, 855–857 (1997).
4. St. Onge, R.P. *et al.* *Nat. Genet.* **39**, 199–206 (2007).
5. Lehar, J., Stockwell, B.R., Giaever, G. & Nislow, C. *Nat. Chem. Biol.* **4**, 674–681 (2008).

Sequencing in real time

Michael L Metzker

DNA synthesis by single polymerase molecules has been visualized at the speed of catalysis, heralding a new sequencing technology of unparalleled throughput.

DNA sequencing methods generally work by halting the process of copying the template strand in one way or another, using dideoxynucleotides (in Sanger sequencing), reversible terminators or natural nucleotides[1]. Now, a report by Eid et al.[2] in *Science* shows that sequence information can be obtained by continuous monitoring of DNA synthesis itself. This strikingly different approach, which records the incorporation of fluorescently labeled nucleotides into single primer strands in real time, promises to increase the speed and read-length of DNA sequencing and to open new avenues in basic research on DNA polymerases and nucleotide analogs.

Most of the next-generation sequencing systems, such as those from Roche/454 (ref. 3), Illumina/Solexa[4] and Life Technologies/Agencourt Personal Genomics[5], are not single-molecule methods as they rely on DNA amplification. A single-molecule technique was recently reported by Helicos Biosciences[6], but its dependence on reversible terminators limits it to the analysis of short DNA fragments. DNA polymerases perform optimally with nucleotide concentrations in the low micromolar range, a requirement that presents a challenge to single-molecule detection methods, which typically use fluorophores at pico- to nanomolar concentrations.

Eid et al.[2], of Pacific Biosciences, solved this problem with the company's zero-mode waveguide array[7], a nanostructured device that reduces the observation volume to the zeptoliter range—an improvement of more than three orders of magnitude over confocal fluorescence microscopy. At this superresolution volume, an estimated 0.01–1 molecule enters the detection layer by diffusion, providing a very low background signal and a signal-to-noise ratio of ~25:1.

To enable parallel sequencing, the authors used a chip with thousands of nanoscale wells containing an immobilized DNA polymerase bound to a primed DNA template to be sequenced (**Fig. 1**). To allow uninterrupted monitoring of nucleotide incorporation, they labeled nucleotides with four distinguishable fluorescent dyes on the terminal phosphate group rather than on the base, creating nucleotide analogs that apparently do not interfere with DNA synthesis by φ29 DNA polymerase, a highly processive, strand-displacing polymerase.

The residence time of the phospholinked nucleotides in the polymerase active site is governed by the rate of catalysis and is on the millisecond time scale. The bound nucleotide generates a recorded fluorescent pulse as no other fluorescent molecules are present in the detection volume of the zero-mode waveguide. Formation of a phosphodiester bond releases the fluorophore, which quickly diffuses away, reducing fluorescence to background levels and generating a natural, unmodified DNA product (**Fig. 1**). Translocation of the template marks an

Michael L. Metzker is at the Human Genome Sequencing Center and the Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza N1409, Houston, Texas 77030, USA.
e-mail: mmetzker@bcm.edu