SYSTEMS BIOLOGY

# Networking an organism

**Researchers have built predictive probabilistic networks for a halophilic archaeon and for *Caenorhabditis elegans*.**

As any teenager will tell you, who you spend time with determines to a large extent what you spend it doing. In this regard, genes and teens may not be so different.

Genes that are associated with each other, in terms of co-regulation, co-expression, physical linkage between their protein products or other associations, are more likely to have similar functions than unassociated genes. Moreover, understanding the functional coupling between genes in an organism should help to describe its biology (**Fig. 1**).

In two separate reports, scientists have built networks that describe the unicellular archaeon *Halobacterium salinarum* (Bonneau *et al.*, 2007), and the multicellular nematode *C. elegans* (Lee *et al.*, 2008). "We're trying to develop pragmatic approaches to go from changes in genotype to changes in phenotype," explains Edward Marcotte at the University of Texas in Austin, senior author on the paper describing the *C. elegans* model, which was done in collaboration with the group of Andrew Fraser at the Sanger Institute. Whereas similar models have been built for simpler organisms, Marcotte and colleagues have now shown that the approach can be applied to a complex multicellular animal such as the worm.

To build the network, postdoctoral researcher Insuk Lee integrated data from many previous studies, including microarray data, physical and genetic interaction data, and associations mined from the literature. Overall, the researchers integrated more than 20 million experimental observations to generate the network, which, simply stated, is a set of relationships that defines the probability that genes function together.

The best test of a network is whether or not it is predictive. "We were testing the notion," says Marcotte, "that if we had a large functional network, then associated genes in that
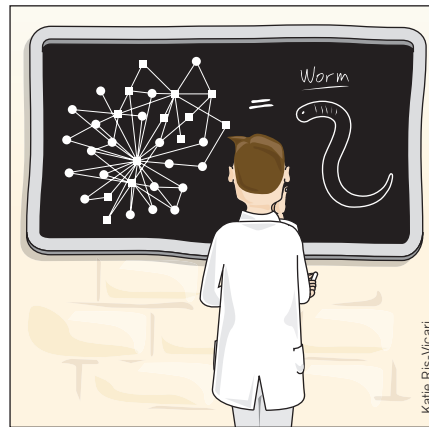


**Figure 1** | Can networks describe organisms?

network would be involved in similar traits, so if we knocked them down we'd get similar effects. You could think of it as guilt by association." Indeed, the researchers showed that, given a known exemplar set of genes involved in a particular trait, such as lifespan, the network could identify other genes with similar functions based on the strength of their associations with the exemplar genes. Absolute phenotypic prediction, however, was limited to predicting whether or not a given gene is essential.

"One of the most exciting things here is that a single network can be predictive for different cell and tissue types in a multicellular organism," says Marcotte. "It was not clear in advance that there would be that kind of specificity encoded in the network." An obvious application of this approach therefore is the identification of candidate genes in human genetic disease. As Marcotte points out, many human diseases do have some associated exemplar genes, and more are likely to be uncovered in the future. These could be used in a network-based approach to identify other candidate disease genes.

What distinguishes the work of Richard Bonneau at New York University and Nitin Baliga and colleagues at the Institute for Systems Biology from other network analyses

is that the archaeon that they chose to model, *H. salinarum*, is relatively unstudied. As there was very little information other than gene ontologies available, the researchers resorted to a data-driven approach. They subjected the archaea to environmental and genetic perturbations, used microarrays to monitor changes in gene expression over time, identified clusters of co-regulated genes, and used these clusters as the basis for building a functional network.

In this case as well, the network is predictive. By carrying out 131 microarray experiments in parallel, Baliga and colleagues showed that the network predicts transcriptional responses to many different perturbations that were not used to initially build it. "It suggests that you need a finite number of perturbations to build a useful model," says Baliga. "This has exciting implications for using the approach on other organisms about which not much is known." What is more, Baliga predicts that one should be able to build such networks using data obtained purely from environmental perturbation studies. This would open up the analysis of organisms for which genetic manipulation is still not possible.

The other eventual application of this work is to synthetic biology—the building of cells. Recent reports on the *de novo* synthesis of microbial genomes have received a great deal of attention, but this, as Baliga puts it, is merely molecular engineering technology. "What you need," he emphasizes, "and which the synthetic biologists don't have, are the comprehensive circuit diagrams for a whole cell. And we're providing the circuit diagrams."

**Natalie de Souza**

**RESEARCH PAPERS**
Bonneau, R. *et al.* A predictive model for transcriptional control of physiology in a free living cell. *Cell* **131**, 1354–1365 (2007).
Lee, I. *et al.* A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 181–188 (2008).