from *SIAM News*, Volume 36, Number 10, December 2003

Some Assembly Required

By Barry A. Cipra

The fundamental unit of natural selection is the gene. Or is it? Garrett Odell has an alternative candidate: the gene *network*. In an invited presentation at the SIAM Conference on Applications of Dynamical Systems (Snowbird, May 27–31), Odell, director of the Center for Cell Dynamics at the University of Washington's Friday Harbor Laboratory, described his group's research on the way genes team up to get things done.

The researchers have found that gene networks are surprisingly robust: In simulations at least, networks survive and function over a wider range of parameters than anyone could have expected. The forgiving nature of networks, by enabling organisms to survive variations in the environment and to experiment with the genetic material they pass along to their progeny, could be central to the tenacity of life and to evolution.

In the central dogma of molecular biology, the genome is essentially a parts list. The parts themselves are proteins; the list consists of genes, recorded in an "owner's manual" made up of DNA. For human beings, the list has about 30,000 different items. (Before the completion of the Human Genome Project, most researchers guessed the number to be three or four times larger.)

What makes life interesting is the way genes and their products interact. Genes spend a lot of their time regulating each other. More precisely, the protein product of one gene may stop the expression (into messenger RNA, the errand boy of genetics) of some other gene or promote the expression of yet another. Protein and mRNA concentrations wax and wane in complicated, interdependent ways. Biochemists have mapped out mathematical models for many of these interactions. These models will, the scientists hope, give insights into how cells work.

"The idea is that we can approach understanding how the whole genome works by breaking it down into groups of genes that interact strongly with each other," Odell explains. Once researchers identify and understand these network modules, the next step will be to figure out the interactions within networks of networks, and so on, "until we eventually understand how the whole genome works, many years from now."

Ultimately, the models translate into systems of differential equations describing the rise and fall in the levels of participating mRNAs and proteins. The rate constants and other parameters of the differential equations depend (in part) on the precise proteins involved, and hence on individual genes. DNA can be thought of not just in terms of the proteins it spells out, but also in terms of the functional networks in which the proteins participate. It's the networks, rather than mere individual genes, that Odell proposes as fundamental units of natural selection: While perhaps not as concrete as strings of A's, C's, G's, and T's, networks are discrete, identifiable objects that are passed down from one generation to the next. In particular, mutations that affect rate constants in the differential equations may matter far less than mutations that disrupt the network.

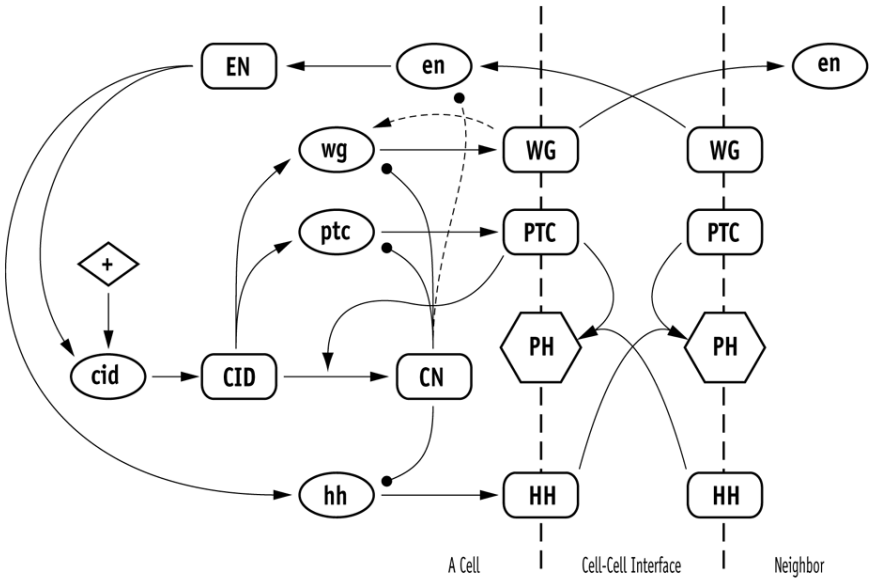

Figure 1. *Missing Links. When only the well-established links (solid lines) among proteins (upper-case) and messenger RNAs (lower-case) were used, simulations of a gene network in Drosophila failed to reproduce the qualitative behavior of the actual net-work. Two additional links (dotted lines), suggested by recent experimental evidence, did the trick. Reprinted with permission from "The segment polarity network is a robust developmental module," von Dassow et al.,* Nature*, vol. 406, 13 July 2000, pp. 188–192.*

Odell and colleagues George von Dassow, Eli Meir, and Edwin Munro have investigated this idea in the fruit fly, Drosophila, the long-time workhorse of genetics. They have focused on a network that controls segment polarity in the fruit fly embryo—basically telling the developing insect which end is up. The network involves the mRNA and proteins of five genes, called wingless, engrailed, hedgehog, cubitus interruptus, and patched (see Figure 1).

"We have fixed our attention on networks of genes that experimenters have studied exhaustively, and where they know most of the genes and most of the connections between them," Odell says. These tend to be small networks, with just a few genes, "because they're the easiest to isolate and understand." The modular hypothesis looks good: "If the genome wasn't partitioned into smallish networks, these experimental approaches wouldn't have succeeded."

They began by examining the known interactions (solid lines in Figure 1). The differential equations implied by the model involve

about 50 rate constants and other parameters, representing things like binding rates and the half-lives of mRNA and proteins—none of which is known, even to within an order of magnitude. Many of the parameters could range over factors of a thousand, "so it's a big parameter space," Odell points out. "We're not talking about ten percent variation, we're talking about a hundred thousand percent variation."

Using Ingeneue, a software package they designed for the purpose (available at www.ingeneue.org), Odell and colleagues searched this enormous parameter space for values that might produce solutions consistent with the required behavior of the network. They found none.

Their next step was to add more links to the network (the two dashed lines in Figure 1), consistent both with recent experimental evidence and with a logical analysis of why the network based on the known links wasn't producing the right behavior. This time they hit the jackpot.

To be precise, out of 240,000 randomly chosen combinations of parameter values, Ingeneue got 1192 "hits"—that is, networks that did more or less what they needed to do. A 1-in-200 success rate may not sound so great, but it is when you consider it's taking place in a 48-dimensional space (48 being the exact number of parameters): It means that each parameter, on average, has a roughly 90% success rate (since $(1192/240{,}000)^{1/48} = 0.895$). Moreover, the suc-cessful hits were not restricted to any particular region in parameter space. Odell likens the set of successful parameter combinations to the holes in a "48-dimensional swiss cheese."

"We couldn't believe this result. We thought there must be an error in our computer software," Odell says. Another group in Sweden found the results so suspect that they rewrote the computer code from scratch—only to get the same results. In retrospect, though, it makes sense. After all, sexually reproducing organisms have two copies of each gene, called alleles, and these alleles are often different, which implies variable rates of expression. In particular, experimental biologists have systematically created fruit flies with only one good allele for particular genes. This by itself should halve the rate of expression, yet the flies continue to thrive. "These networks have to be phenomenally robust in order for that to work," Odell says.

The robustness of the results is an indication that nature may rely on network complexity to offset the effects of mutations and compensate for environmental "noise"—of particular concern for a developing embryo. Indeed, the researchers found the success rate for a simple, "engineered" network with a mere 27 parameters to be less than 1 in 50,000; a more complicated network—which includes four more genes (called frizzled, armadillo, pangolin, and sloppy paired) and brings the number of parameters to about 75—raised the rate to about 1 in 100. Odell compares the likelihood of engineering such a network to that of building a working radio from a random set of resistors and capacitors.

The stability of complex networks resides in their topology, Odell says. While they are robust to changes in parameter values, they are sensitive to changes in structure: Adding or subtracting random connections almost always kills the previously good parameter sets. This suggests network topology as a new basis for evolution. Nature's ceaseless biochemical experiment, that is, can be traced not only by the point mutations that fiddle with the structure of individual proteins, but also by network mutations that occur in the rare cases in which an altered protein forges a new link or severs an old one (or combines two previously unrelated network "modules") without immediately killing the organism.

Evolution by topology may take time to catch on among biologists—if it ever does. Odell thinks they should be enthusiastic about the theory. After all, he points out, making accurate measurements of half-lives of proteins and other such parameters is extremely difficult. "For years biologists have been looking to identify the genes and their qualitative role, and they haven't paid any attention at all to measuring rate constants," he says. "That's a kind of bold thing to do, given how complex the technology of a living organism is. In retrospect, they were right to ignore the quantitative details, because they appear not to matter. It sort of confirms the wisdom of experimental biologists in not squandering their time measuring quantitative details." In other words, he jokes, "this result congratulates them for their laziness."

Barry A. Cipra is a mathematician and writer based in Northfield, Minnesota.