Genomics & Proteomics

Systems Biologists Push Technology to the Edge

• Systems biology is coming of age with the help of key technologies in the fields of informatics, robotics, microfluidics, and nanotechnology. In turn, systems biologists are demanding newer, faster, smarter technologies to carry out their experiments. • **By Catherine Shaffer**

It is perhaps debatable which came first, science or technology, but there is an interdependence between the two. Advances in science stimulate new technological developments, which in turn stimulate more advances in science. The development of science and technology is not merely a miracle of modern life which we take for granted, but a function of the mechanism of human thought.

Our curiosity about the physical world leads us to form hypotheses, to test them, and to put the results to a practical use. Thus, there is a closer link than we may realize between scientific theory and the gadgets we find at the electronics store. In fact, the nature of those gadgets depends on the nature of the questions we ask about our environment, and the methods we use to answer those questions.

Although science is split into many different disciplines, scientific theoreticians tend to be attracted to universality. Over roughly the past half century, a sort of über-science has arisen that connects such disparate subjects as physics, biology, economics, and even marketing. This science goes by various names: "systems theory," "complex systems," or "systems science." All refer to the study of systems made of interrelated or interconnected parts.

The figure represents the actual biological influences within a cell as obtained from analysis of gene expression data. It illustrates the influence of gene groups involved in key cellular level functions. (Source: Yaneer Bar-Yam, PhD, New England Complex Systems Institute)

Biological life is a perfect example of a complex system, and in fact, early systems theory was inspired by biology. The biological sciences, however, pursued mechanistic and causal experimentation into the 1970s and 1980s, and have not until recently begun to take advantage of systems theory. This may be due to the limitations of technology. The scale of computing power necessary to analyze a biological system has been largely unimaginable until recent times, and even now, systems biologists chafe at the limitations of technology.

According to Paul Thomas, PhD, senior director of computational biology at Applied Biosystems, Foster City, Calif., tools for the study of biological systems need to fulfill three roles: making perturbations to a biological system; measuring perturbations in a biological system; and interpreting the data. In practice, systems biology experiments and technology have an iterative relationship, wherein experimental results stimulate the use and development of technology, and this new technology in turn gets used in new experiments.

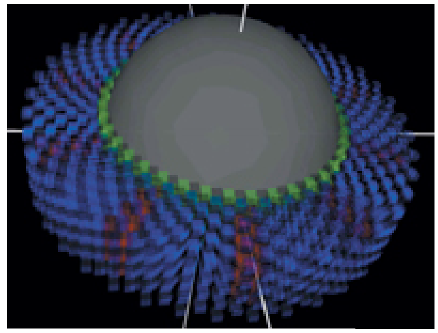
The microarray

The DNA microarray was the first tool for analyzing many genes simultaneously, and is thus well-suited for systems-based experiments. The method allows the analysis of thousands of genes by hybridization to a target molecule of choice. It is popular for discovery-based science, the sort of mass data mining and screening that occurs in pharmaceutical research.

Yaneer Bar-Yam, PhD, president of the New England Complex Systems Institute, Cambridge, Mass., subjected one such large-scale discovery screen to a mathematical analysis that produced a robust description of the interactions of groups of genes [B. D. Bivort *et al., Proc. Natl. Acad. Sci.,* vol. 101, pp. 17687-17692 (2004)].

Using murine expression array data contributed by the Alliance for Cell Signaling, Bar-Yam and his collaborators assigned genes in murine B lymphocytes to 12 major groups using a self-organizing map (SOM) algorithm.

The SOM is a technique invented in another branch of complex systems studies: neural networks. It reduces the dimensions of a data set to a manageable size. Bar-Yam's team then calculated expression levels for the 12 gene groups over times ranging from 30 minutes to four hours, using 33 linear equations. This global analysis revealed regulatory influences of ATP-generating genes in mitochondria, glycolytic genes, and chromatin- reorganizing genes, among others.



A quantum coreworld after a fluctuation as computed by a Quantum Virtual Machine (QVM). Bricks in the coreworld are arranged into neighboring regions which can be distributed to different physical processors for simulation. Green bricks enter the biosphere and provide a source of nutrients. Red bricks are exchanged during fluctuations stitching together the various regions. The work permits the computational investigation of biological systems on temporal and spatial scales inaccessible to other approaches. (Source: Alexander Wait, Harvard University Department of Biophysics)

Bar-Yam's research spans many scientific disciplines, including such seemingly unrelated subjects as health care systems, sociology, military conflict, and evolution. This diversity of interests is not only common in systems theory, but pretty much required.

"People come to our courses in complex systems and see the world in a totally different way," says Bar-Yam. "It has to do with the perspective of how things work together, trying to discover the patterns of behavior of the system as a whole. The advantage of having a lot of tools is it lets you answer a wider range of questions. If you only answer the questions with a particular tool, that's what's going to limit your understanding."

High-throughput technology

Many of the technologies used in systems biology are high-throughput in nature. Tools used in proteomics include mass spectroscopy; separations techniques such as microcapillary extraction and solid phase extraction; and software for analyzing data, mapping protein interactions, modeling protein structure and folding, and sequencing peptides.

Moving up from the molecular scale to the cellular scale, automated cell sorters have become an important technology in systems biology. In a cell sorter developed at the Institute for Systems Biology (ISB), Seattle, cells are forced into a thin stream of water and analyzed individually by a laser beam, at a rate of 40,000 cells per second. Other up-and-coming high-throughput technologies include genotyping and DNA sequencing.

John Aitchison, PhD, associate professor at the ISB, studies the biogenesis of peroxisomes using high-throughput technology. Peroxisomes are intracellular organelles whose primary function is beta-oxidation of fatty acids. Because of intimate links to oxidative metabolism, defects in the peroxisomes can be catastrophic. Technologies used in Aitchison's lab include whole-genome transcriptional arrays, complementary large-scale mass spectrometry, isotope-coded affinity tag-MS (ICAT), and high-throughput genetic and phenotypic screens.

The ISB was founded in 2000 by Leroy Hood, PhD, who is said to have coined the term "systems biology" in the 1980s. The institute's mission is to probe the mysteries of human biology. "Systems biology ultimately will be the foundation of medicine," says Aitchison, "predictive, preventive, personalized medicine."

Aitchison's research on peroxisomes not only holds out hope of a cure for peroxisome-related diseases, but provides a biological model. "We can learn biology by studying [peroxisomes]," he says. "That's the point. What makes them a really good systems biology model is that they're dynamic . . . understanding molecular networks and propagation of networks through cells is really what we're doing."

Computer technology
The roles filled by computers fall into roughly two categories: data analysis and model building. Applied Biosystems has installed about 180,000 instrument systems in 100 countries for use in genomics, agriculture, forensics, informatics, molecular diagnostics, and many other specialties. Their product, PANTHER 5.0, is a freely available database that allows scientists to browse biological pathways or processes associated with protein families (https://panther.appliedbiosystems.com).

"What you really need as a researcher is an encyclopedia, with entries for each of the molecules you're looking at," says Thomas of Applied Biosystems. "This is a really big encyclopedia with 30,000 entries for genes alone, and twice as many transcripts, 50,000 to 100,000, and even more proteins. Instead of having to look at 30,000 gene entries, you have about 250 bioprocess categories. All of a sudden, the problem is tractable for a human being again."

One example of a systems biology experiment using the PANTHER technology is a collaboration with Johan Kuiper, PhD, of the Leiden/Amsterdam Center for Drug Research, Netherlands. They compared mRNA expression in liver tissue in mice that were fed a high cholesterol diet with mice fed a normal diet, and found which genes were expressed at higher levels under the influence of cholesterol. Using the PANTHER statistical tools, they found many influences that were expected and already established scientifically. The analysis also pointed them toward some unexpected and far-reaching changes in other metabolic pathways, including carbohydrate metabolism and even amino acid metabolism.

Utility of modeling
Changing scales is a hallmark of the systems method, whether it's zooming out, as in Bar-Yam's work with murine expression arrays, or zooming in to look at single molecules or scales even smaller and more abstract. Often, doing science at these changing scales requires the use of computer models. Modeling tends to be a controversial topic in the life sciences. Many biologists don't understand the benefit or

Nanostrings Attached
As a young field, systems biology is currently most active in universities and research institutes such as the Institute for Systems Biology (ISB), Seattle. Yet, there is a small but growing commerce in technology and tools for systems biology. One of the areas of greatest potential is in nanotechnology.

Nanostring Technologies, Seattle, was spun out from the ISB to develop an automated, robust, highly multiplexed single-molecule detection device to identify and count individual target molecules in a biological sample by attaching a tag, called a nanostring. Although the technology is currently proprietary, Nanostring Technologies chief operating officer Amber Ratcliffe was willing

Richard Lenski, PhD, professor of microbial ecology at Michigan State University, East Lansing, Mich., has attracted quite a bit of interest in his ongoing experiment in bacterial evolution. He has been following the evolution of a colony of *E. coli* bacteria since 1988, involving tens of thousands of generations. In research featured in the journals *Science* and *Nature*, Lenski described how the bacteria adapted to different living conditions, how they differentiated, even what changes occurred on a molecular level. By its nature, evolution submits well to a systems approach.

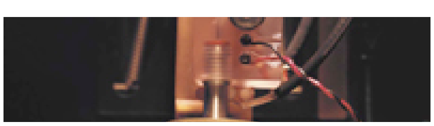
In an evolution of Lenski's experiments with *E. coli*, he became interested in digital evolution, and began collaborating with Chris Adami of the California Institute of Technology (Caltech), Pasadena, Calif., and Charles Ofria of Michigan State University, working on evolution in the Avida software platform. In Avida, developed at Caltech, simple computer programs mutate and evolve in a simulated environment. Through this marriage of biology and computer science, Lenski and Ofria have been able to study evolution in a system wherein all of the variables can be controlled and changed to an even greater extent than the *in vitro* evolution experiments with *E. coli*.

A common complaint about computer simulations is that experiments conducted under such a comprehensive set of controls produce artificial results which are not relevant to the real, biological world—that, in fact, biology is too complex to model on a computer. "A simulation performs the instructions that we give so that we can see whether the behavior of the system as we understand it is commensurate with what we know or what we expect," says Bar-Yam. "It's simply a tool for recognizing whether or not our understanding is correct. If it's wrong, that is in many cases the best result, because it forces us to rethink what we know, or what we think we know. People think that a simulation that doesn't give expected results is a failure, but it's actually a key step to understanding the system."

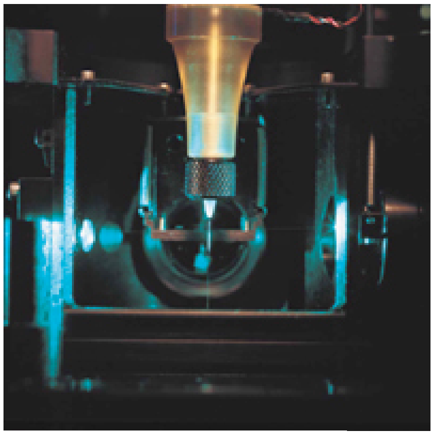
to reveal that nanostrings will be a reagent mixed with a sample that will allow researchers to scan the sample for molecules of interest in much the way a grocer scans the barcodes on packages of food. The development of this technology has been driven by dissatisfaction with conventional microarray technology, in which 100,000 copies of each RNA are necessary for detection. Nanostrings will detect each individual copy.

"Traditional biology looking at one gene, trying to understand how affects one gene . . . We're the next generation to that," said Ratcliffe. "We're a tool that will enable people to continue doing systems biology where they're looking at many things simultaneously."

Computer modeling can also allow researchers to go in unexpected and highly speculative directions. At Harvard University, Cambridge, Mass., graduate student Alexander Wait studies the behavior of digital organisms under the direction of mentor George Church. Like Avida, Wait's quantum coreworld is an environment in which digital organisms compete for limited resources. Unlike Lenski and Ofria's digital creatures, however, Wait's are quantum mechanical. The goal of this research is to model quantum artificial life, and ultimately to engineer quantum-mechanical biomolecular machines, with real-world practical applications such as drug delivery, early cancer diagnosis, or monitoring RNA expression levels in human beings.



In research summarized in *Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* in 2004, Wait describes experiments in



The ISB-invented cell sorter processes 40,000 cells per second. (Source: The Institute for Systems Biology)

which quantum and non-quantum digital life forms compete for nutrients inside the quantum coreworld. The results show that the quantum organisms have some selective advantage —that they might possibly, in fact, be "cleverer" than ordinary digital life forms. Although very little scientific credence is given to the idea that large-scale quantum effects could be important in biology, Wait's research gives insight into what biological life could look like if such was the case.

"I see systems biology and modeling as being complementary to engineering," says Wait. "One will accelerate the other. People underestimate the suddenness of what is to come. There will be a radical transition from no appreciable synthetic biology capacity to the possibility of building very complicated biomolecular machines. When this transition will occur is not obvious, but it could very well be in our lifetimes."

Fusion of technology, biology

The word "technology" has many different meanings. It brings to mind computers, automobiles, cell phones, and other electronic or mechanical devices that people have come to rely on in daily life. Yet stripped to its core, technology is nothing but a set of tools humans use to interact with their environment. Archeologists study humankind's earliest technology of stone blades, woven baskets, and bone needles. These innovations allowed human beings to live more comfortable and successful lives in the stone age.

It's clear that our ability to understand and manipulate our environment continues to expand with the development of new technologies, and that it is intimately and iteratively linked with the nature of the questions asked and the organization of those questions.

Systems biology is the field of study with the greatest stake in new and developing life science technologies, and the most likely to push the boundaries of technology in search of greater understanding. It is an irony, then, that systems biology and technology are at opposite ends of the scientific spectrum: basic research and applied science. It implies that the relationship between the two is not a linear continuum, but a circle, made of interconnected and interwoven parts—a complex system, continually increasing in complexity.

Shaffer is a freelance writer based in Ann Arbor, Mich.

Glimpsing the Future

I joke that we will know that our work on digital organisms has succeeded when the evolved programs write our papers for us. While I'm kidding at one level, just saying something like that forces us to think more deeply about our long-term goals and visions. Most working biologists think of computer science as a tool to help them handle and analyze huge data sets. That's very important, but the real excitement, which has the potential to radically influence our future world and technology, is more subtle and remains hidden to most people. That excitement lies in the fact that more and more scientists and engineers are discovering and harnessing fundamental principles of biology to evolve new technologies."

— Richard Lenski, PhD, professor of microbial ecology at Michigan State University

---