

A new movement aims to integrate biology, mathematics, and engineering; even if its objective is hard to define, it is all the rage in the academic world

## Tracing Life's Circuitry

In September, Harvard University opened its medical school's first new department in 20 years. Its focus: systems biology, one of the hottest—and most elusive—new fields in biology. The nearby Massachusetts Institute of Technology (MIT) had already started a Computational and Systems Biology Initiative with 80 faculty members. And the Weizmann Institute of Science in Rehovot, Israel, is following suit, planning its own systems biology institute. All are forging links between biologists, mathematicians, and engineers. Systems biology is fast becoming the academic topic du jour.

Fueling this movement is a gush of data. New technologies have inundated researchers with a deluge of information on genes, proteins, cellular dynamics, and organisms' responses to mutations and the environment. But they haven't explained what makes whole organisms tick. Systems biologists are taking on that challenge, relying heavily on mathematics and statistics to integrate data into a more complete picture of how biological networks from cells to whole organisms function. They are building models and making predictions about how biological systems will behave; the ultimate goal is to understand deep mysteries—such as how cells divide, animals develop, plants flower, and humans breathe.

There's a new world coming, says one of the proponents, David Galas, a molecular biologist at the Keck Graduate Institute in Claremont, California. "It seems to me," he says, "this is the beginning of real biology."

**Bewildering beauty.** This rendering shows the protein actin's various chemical states (green) and the molecules and reactions that alter those states (yellow).

### Data deluge

Indeed, the systems approach may be just what biology needs to keep afloat. The amount of genomic data in the public database at the National Center for Biotechnology Information doubles every 18 months. There are now some 1200 person-years of experimental data on the role of a few genes in the development of the fruit fly alone, says theoretical biologist Garrett Odell of the University of Washington, Seattle. All of this is forcing people to adopt "a new way of doing biology," says Manoj Samanta, a computer scientist at NASA Ames Research Center in Moffett Field, California.

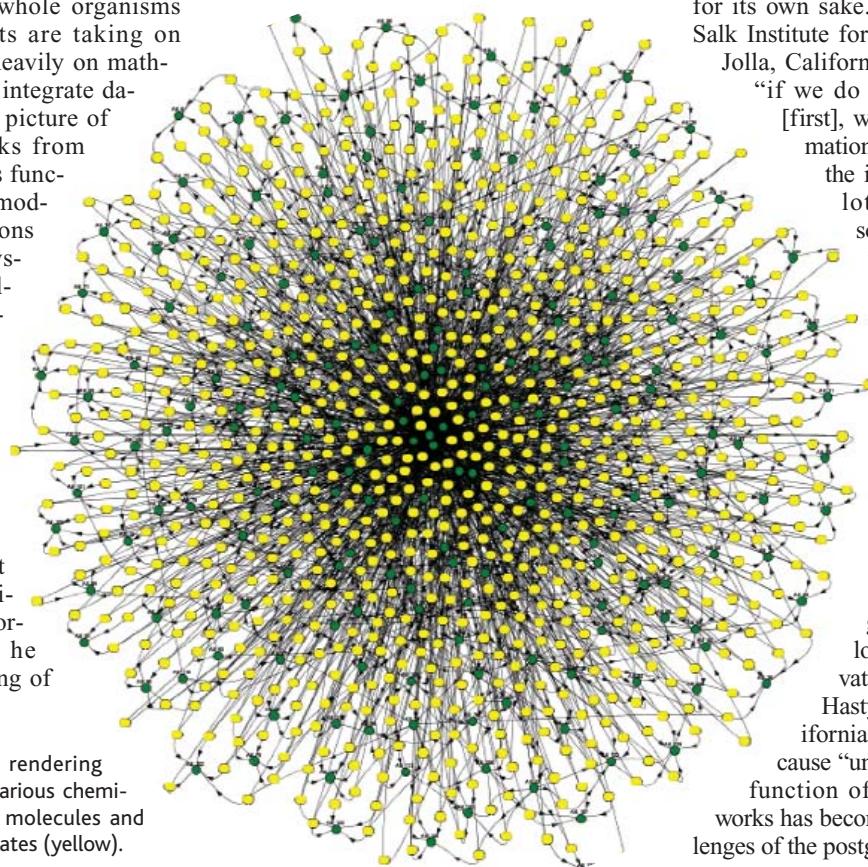
Given the sense of urgency, systems biology seems to be on everyone's lips. In early November, the National Science Foundation (NSF) and the National Institutes of Health (NIH) held workshops on the topic. In

St. Louis, Missouri, more than 500 researchers swapped ideas and research results at the 4th International Conference on Systems Biology. "There is a real possibility of building knowledge from [the] molecular level to the system level," says NSF's Mitra Basu, a computer engineer.

Challenges lie ahead, however. For one, there needs to be true collaboration. Engineers and computer experts need to remember that "you can't do systems biology without a detailed understanding of the biology," warns Leslie Loew, a cell biologist at the University of Connecticut Health Center (UCHC) in Farmington. Leaders in this field are struggling to develop common standards, communicate across disciplines, and overcome bewilderment among department heads and grant reviewers. At the same time, Sydney Brenner warns, the movement needs to take care not to get infatuated with data for its own sake. A Nobel laureate at the Salk Institute for Biological Studies in La Jolla, California, Brenner suggests that

"if we do not define the problem [first], we won't know what information is important." He rejects the idea that one can "make a lot of measurements and something will come of it; I think that's rubbish."

Nonetheless, the maturing of molecular biology, new data-intensive techniques, increased computer power, and new algorithms "have changed people's attitudes about how they think about [biological] problems," says Galas. Researchers are increasingly willing to apply engineering concepts to biological systems. Such innovation is essential, says John Hasty of the University of California, San Diego (UCSD), because "uncovering the structure and function of genetic regulatory networks has become one of the central challenges of the postgenomic era."



CREDIT: JOHN CARSON/UCHC

## Back to the future

Although the pioneers may constitute a small circle, “systems biology should not be and will not be an elite club,” says Arcady Mushegian of Stowers Institute for Medical Research in Kansas City, Missouri. Many scientists may already fit under the movement’s banner. Biochemists, for example, have spent decades piecing together metabolic pathways. Neurobiologists have traced neural networks in moths, squid, and other organisms. In truth, “systems biology is in the eye of the beholder,” concedes proselytizer Leroy Hood, president of the Institute for Systems Biology in Seattle.

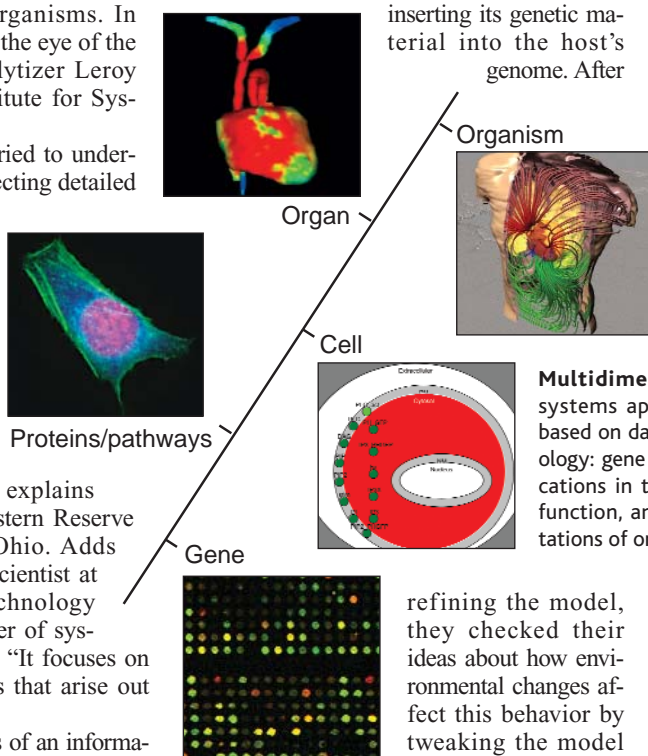
Physiologists have long tried to understand how life works by collecting detailed quantifiable information on biological stimuli and responses. But modern systems biology is different, say proponents. “It’s not just cataloging parts or even interactions, but asking, ‘How do new functions arise? How do you look for stuff that’s not evident [from] the details?’” explains Joseph Nadeau of Case Western Reserve University in Cleveland, Ohio. Adds Hiroaki Kitano, a systems scientist at the Japan Science and Technology Corp. in Tokyo and a pioneer of systems biology in his country: “It focuses on the dynamics and properties that arise out of interactions.”

Couching biology in terms of an information system and analyzing the system mathematically is an essential part of the new approach. Alan Hodgkin and Andrew Huxley pioneered this technique 50 years ago, using a mathematical description and circuit diagram for nerve cells in squid, with some success in predicting how the neurons worked. Models have grown more detailed since then, and researchers have learned that the more data they incorporate, the more accurately they can predict complex biological responses such as changes in gene expression. Eventually, says Sean Athey, a computational biologist at the University of Michigan, Ann Arbor, “these models will allow us to get at fundamental theories.”

The 1990s saw a steady growth in the sophistication of the systems approach. “What really propelled systems biology was the Human Genome Project,” Hood explains. Essential to that effort was new software that could analyze whole genomes and other large biological data sets. It was the dawn of high-throughput biology, first applied to DNA sequencing and later to studies of gene expression and protein interactions. According to

Hood, the genome project demonstrated that biology is really about information systems.

Although the transition was gradual, systems biology by the mid-1990s had become a mathematics and data-intensive world. In 1995, Harley McAdams and Lucy Shapiro of Stanford University School of Medicine built a model of electrical circuits to examine certain behaviors of a phage, a bacterial virus. It helped them determine why a phage kills the host cell rather than inserting its genetic material into the host’s genome. After



**Multidimensional biology.** The systems approach builds models based on data from all levels of biology: gene expression, protein locations in the cell, models of cell function, and computer representations of organs and organisms.

refining the model, they checked their ideas about how environmental changes affect this behavior by tweaking the model and seeing if their intuition was right.

Two years later, Stanislas Leibler of Rockefeller University in New York City and his colleagues melded quantitative experimental data with modeling to predict the ability of certain bacteria to orient themselves to specific chemicals; they used quantitative information such as the concentration of the attractant in their analyses. The data-rich analysis revealed that chemotaxis was driven by feedback within the phage-chemical network and not by the concentration of individual molecules. Since then, Leibler and others, particularly Naama Barkai and Uri Alon of the Weizmann Institute, have continued to refine the data and the models. Leibler did more than describe a system; he “really explained the design principle,” says Marc Kirschner, chair of Harvard’s new department. “He showed that this system did some things very well—and some things very poorly.”

In another data-intensive study 3 years ago, Hood and his colleagues tackled gene activity in yeast’s sugar metabolism. They changed the yeast’s diet from raffinose to

galactose, knocked out one gene at a time, and watched to see how gene expression and protein activity changed. Aided by microarrays, mass spectrometry, and computers, they identified 1000 relevant genes and grouped them according to how they and their messenger RNA changed during the experiment. Hood’s group merged the results with the protein data to build a model explaining how yeast adapts to a galactose diet, and the model correctly mimicked experimental results. The work illustrated that “it’s now possible to really do systems biology seriously,” says Galas.

Model building is one of the chief objectives of systems biologists, who dream of being able to make verifiable predictions about living things just as physicists make predictions about matter. Eric Davidson, a developmental biologist at the California Institute of Technology (Caltech) in Pasadena, says that the model he created of sea urchin development has made accurate predictions about how regulatory DNA alters its activity in response to genetic mutations.

Davidson’s group uses a variety of experimental data: descriptions of developmental pathways, known genes, and regulatory DNA discovered through experiments that disrupted normal gene function. The researchers have pinpointed connections between various genes and regulatory regions. For example, they have detected proteins from the mother that stimulate the formation of the endoderm in offspring. Then they observed repressor proteins setting limits on where endoderm appeared.

It’s a big step to go from modeling gene networks to modeling organs and other large systems, but that’s what Case Western’s Nadeau is trying to do. He and his colleagues first measured various aspects of heart function in about 20 mouse strains. Using a sophisticated statistical technique, they got results that fit with what was already known about how traits change in concert—such as the correlation between ventricle mass and size during contraction and expansion. “It was a proof of principle,” Nadeau says. “Sooner or later, we are going to have to understand how different organs interact in health and disease.” He is now using this approach to evaluate traits that affect responses to high-fat diets.

## First principles

Like Nadeau and Davidson, many others are sifting through molecular networks in search of coherent patterns of stimulus and



response. C. H. Luo and Yoram Rudy of Case Western, for example, have begun using a model of the heart for clinical studies, helping evaluate congestive heart failure patients for gene therapy. Stanford's Deborah Gordon is using a systems approach on ants to study how individuals "know" when to alter their activity according to the colony's needs. Developmental biologists are modeling how limbs form in embryos; plant researchers are looking at the timing of root development.

Others are trying to develop explicit rules for these networks. Uri Alon, a physicist-turned-biologist at the Weizmann Institute, is trying to apply circuit theory to biological systems, taking the field a step closer to engineering. His analysis of a transcription network in bacteria, for example, identified specific network components—often called motifs—that were used repeatedly in studies that range from neural networks to whole ecosystems.

This work "raises the possibility that complex network behavior can be broken down and understood in terms of individual motifs," which serve an important purpose, Alon says. They can buffer against misinformation and, in the case of transcription networks, help make sure that genes are expressed in the right order and at the right time.

"The similarities between evolved circuits and engineered circuits raise the hope that there are deep laws of nature" that unite living and designed systems, Alon adds. But, he cautions, there are differences. Physical systems "just sit there," whereas "biological systems produce beautiful machines that dance, work perfectly despite cellular noise, and dissolve when they are done." He thinks the "rules" of biology will remain somewhat elusive.

### All systems go?

Although systems biology has come a long way since the days of Hodgkin and Huxley, it is still in its childhood. And because the field has been sprouting independently in labs across the globe, it lacks coherent

standards for experimental procedures and computer programming. That, say its proponents, could be crippling in the long run.

Some also worry that researchers have little incentive to share all relevant information and maintain databases. Already, the Internet is full of ghost sites that have languished untended after their creators lost interest or support. "We don't want data to go away when the grant is over," says UCSD bio-engineer Shankar Subramanian. Databases and simulation programs need to be more compatible and convenient, says Eugene Bruce, a biomedical engineer at the University of Kentucky, Lexington. There are so many data on the Web in so many places, "I really get a sense that people are a little overwhelmed," Bruce adds.

Help may be on the way. Several groups are maintaining high-quality databases on complex processes, such as cell signaling. The Signal Transduction Knowledge Environment maintained by *Science* is an expanding database that makes cell signaling easily accessible. Two others are running: the Alliance for Cellular Signaling led by the University of Texas Southwestern Medical Center in Dallas and the Kyoto Encyclopedia of Genes and Genomes in Japan. Meanwhile, the Systems Biology Markup Language (SBML) consortium led by Mike Hucka of Caltech is helping establish common terms. SBML seems to be establishing a "de facto standard," says Kitano of the Japan Science and Technology Corp.

About 30 simulators and analysis tools now follow its guidelines.

### Hearts and minds

One of the biggest remaining obstacles is also one of the most intractable: integrating disciplines. Even though universities have made some progress building bridges, it's still hard to get mathematicians to "understand what biologists really need" or to get computer scientists and molecular biologists to speak a common language, says Bruce.

At the NSF and NIH meetings, some researchers argued that only if changes are made in graduate and medical school curricula will the field move forward. They also recommended that peer reviewers be educated about how to evaluate systems biology proposals. The two agencies, they said, should support short courses in systems biology as well.

Some of these and earlier recommendations have already led to action. This fall NIH established two study sections focused on computational biology. At the same time, NIH and NSF are scrambling to beef up funding. NIH, for example, has set aside \$12 million in 2004 for computational biology centers to develop tools for experimental researchers. In 2005, NIH plans to fund both young and seasoned investigators who want to plunge into quantitative work.

On another front, NSF is pushing computer scientists to take a look at biology. Next year it will award its first grants to researchers to develop computational models driven by biology. The budget has not yet been established. It's an exciting prospect for computer people, NSF's Basu explains. Most software programs are primitive compared to living systems, and insights from biology could greatly improve their quality.

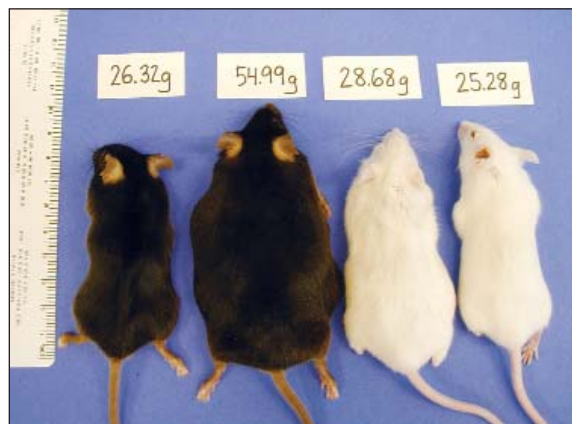
Major institutions are not waiting. Harvard's new center should grow over the next decade to 25 faculty members, says Kirschner. Training students will be a key focus. MIT has received a gift of equipment from IBM and a \$16 million NIH Centers of Excellence grant to jump-start its program. Over the past 2 years, UCSD has hired eight faculty members, all focusing on systems biology and bioinformatics. In November, Stanford licensed a new Web-based computer program to boost its ability to analyze complex genomic data. At Princeton University, David Botstein is pushing his genomics institute to focus more on systems approaches.

Andreas Wagner of the University of New Mexico in Albuquerque sums it up: "[Systems biology] is in fashion right now," and everyone "is jumping on the bandwagon."

—ELIZABETH PENNISI



**Models make a difference.** Mice (below), sea urchins (above), and other animals that can be genetically altered help researchers make sense of the role of individual genes in life's circuitry.



CREDITS: (TOP TO BOTTOM) SHERRY BALLARD/COPYRIGHT CALIFORNIA ACADEMY OF SCIENCES; J. MADEAU/CASE WESTERN RESERVE UNIVERSITY