Researchers produce first complete computer model of an organism

July 20 2012, By Max McClure

The Covert Lab incorporated more than 1,900 experimentally observed parameters into their model of the tiny parasite Mycoplasma genitalium. Illustration: Erik Jacobsen / Covert Lab

(Phys.org) -- In a breakthrough effort for computational biology, the world's first complete computer model of an organism has been completed, Stanford researchers reported in the journal *Cell*.

A team led by Stanford bioengineering Professor Markus Covert used data from more than 900 scientific papers to account for every molecular interaction that takes place in the life cycle of Mycoplasma genitalium – the world's smallest free-living bacterium.

By encompassing the entirety of an organism in silicon, the paper fulfills a longstanding goal for the field. Not only does the model allow

researchers to address questions that aren't practical to examine otherwise, it represents a stepping-stone toward the use of computer-aided design in bioengineering and medicine.

"This achievement demonstrates a transforming approach to answering questions about fundamental biological processes," said James M. Anderson, director of the National Institutes of Health Division of Program Coordination, Planning and Strategic Initiatives. "Comprehensive computer models of entire cells have the potential to advance our understanding of cellular function and, ultimately, to inform new approaches for the diagnosis and treatment of disease."

The research was partially funded by an NIH Director's Pioneer Award from the National Institutes of Health Common Fund.

From information to understanding

Biology over the past two decades has been marked by the rise of high-throughput studies producing enormous troves of cellular information. A lack of experimental data is no longer the primary limiting factor for researchers. Instead, it's how to make sense of what they already know.

Most biological experiments, however, still take a reductionist approach to this vast array of data: knocking out a single gene and seeing what happens.

"Many of the issues we're interested in aren't single-gene problems," said Covert. "They're the complex result of hundreds or thousands of genes interacting."

This situation has resulted in a yawning gap between information and understanding that can only be addressed by "bringing all of that data into one place and seeing how it fits together," according to Stanford

bioengineering graduate student and co-first author Jayodita Sanghvi.

Integrative computational models clarify data sets whose sheer size would otherwise place them outside human ken.

"You don't really understand how something works until you can reproduce it yourself," Sanghvi said.

Small is beautiful

Mycoplasma genitalium is a humble parasitic bacterium, known mainly for showing up uninvited in human urogenital and respiratory tracts. But the pathogen also has the distinction of containing the smallest genome of any free-living organism – only 525 genes, as opposed to the 4,288 of E. coli, a more traditional laboratory bacterium.

Despite the difficulty of working with this sexually transmitted parasite, the minimalism of its genome has made it the focus of several recent bioengineering efforts. Notably, these include the J. Craig Venter Institute's 2009 synthesis of the first artificial chromosome.

"The goal hasn't only been to understand M. genitalium better," said co-first author and Stanford biophysics graduate student Jonathan Karr. "It's to understand biology generally."

Even at this small scale, the quantity of data that the Stanford researchers incorporated into the virtual cell's code was enormous. The final model made use of more than 1,900 experimentally determined parameters.

To integrate these disparate data points into a unified machine, the researchers modeled individual biological processes as 28 separate "modules," each governed by its own algorithm. These modules then

communicated to each other after every time step, making for a unified whole that closely matched M. genitalium's real-world behavior.

Probing the silicon cell

The purely computational cell opens up procedures that would be difficult to perform in an actual organism, as well as opportunities to reexamine experimental data.

In the paper, the model is used to demonstrate a number of these approaches, including detailed investigations of DNA-binding protein dynamics and the identification of new gene functions.

The program also allowed the researchers to address aspects of cell behavior that emerge from vast numbers of interacting factors.

The researchers had noticed, for instance, that the length of individual stages in the cell cycle varied from cell to cell, while the length of the overall cycle was much more consistent. Consulting the model, the researchers hypothesized that the overall cell cycle's lack of variation was the result of a built-in negative feedback mechanism.

Cells that took longer to begin DNA replication had time to amass a large pool of free nucleotides. The actual replication step, which uses these nucleotides to form new DNA strands, then passed relatively quickly. Cells that went through the initial step quicker, on the other hand, had no nucleotide surplus. Replication ended up slowing to the rate of nucleotide production.

These kinds of findings remain hypotheses until they're confirmed by real-world experiments, but they promise to accelerate the process of scientific inquiry.

"If you use a model to guide your experiments, you're going to discover things faster. We've shown that time and time again," said Covert.

Bio-CAD

Much of the model's future promise lies in more applied fields.

CAD – computer-aided design – has revolutionized fields from aeronautics to civil engineering by drastically reducing the trial-and-error involved in design. But our incomplete understanding of even the simplest biological systems has meant that CAD hasn't yet found a place in bioengineering.

Computational models like that of M. genitalium could bring rational design to biology – allowing not only for computer-guided experimental regimes, but for the wholesale creation of new microorganisms.

Once similar models have been devised for more experimentally tractable [organisms](#), Karr envisions bacteria or yeast specifically designed to mass-produce pharmaceuticals.

Bio-CAD could also lead to enticing medical advances – especially in the field of personalized medicine. But these applications are a long way off, the researchers said.

"This is potentially the new Human Genome Project," Karr said. "It's going to take a really large community effort to get close to a human model."

More information: www.cell.com/abstract/S0092-8674%2812%2900776-3