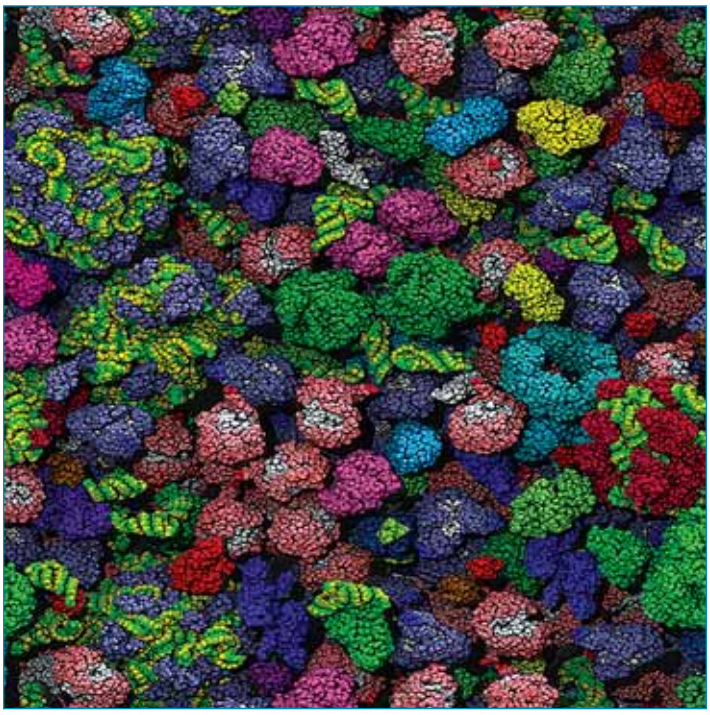
NewsBytes

Simulating Crowded Cytoplasm

In biology textbooks, the carefully rendered cross-section of an *E. coli* cell often resembles a well-organized and spacious apartment, with everything in its place and ample room for movement. But a recent computational recreation of the scene looks more like a Friday night dance floor, with molecules bumped up against one another in every direction. In addition to providing a dramatic, qualitative description of the crowded cytoplasm, this first atomically detailed computational model of *E. coli* innards is also a tool for quantitative predictions of molecular conduct within the cell. The model is described in the March 2010 issue of *PLoS Computational Biology*.



Combining all available known details about the atomic structures and concentrations of 50 of the most common proteins within* E. coli*'s cytoplasm, Elcock and McGuffee created a model of what it might be like inside the crowded cell. They then simulated 20 microseconds of jostling with and without various types of molecular interactions, including crowding (excluded volume effect) and electrostatic and hydrophobic interactions. They then compared the results to experimental observations. Reprinted from McGuffee SR, Elcock AH, 2010 Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm.* PLoS Comput Biol *6(3): e1000694. doi:10.1371/journal.pcbi.1000694.

"This is an attempt to build a virtual lab, in which we can study various biological and biophysical processes as they might occur inside the cell," says **Adrian Elcock, PhD**, coauthor and associate professor of biochemistry at the University of Iowa.

The sea of floating proteins inside every cell is the background against which many cellular reactions take place. Scientists realized years ago that the cytoplasm is generally not an invisible player in those reactions. One of the best-studied examples is macromolecular crowding (also called excluded volume effect). Having large neighbors on every side changes a protein's effective concentration and influences its movement and ability to react. A biological reaction observed in dilute solution can be much faster or slower than the same reaction inside a crowded cell.

To create the model, Elcock and then graduate student **Sean McGuffee, PhD**, started by gathering known structural data for 50 of the most common *E. coli* proteins. They then combined the detailed representations inside a computer model at known cellular concentrations, creating a strikingly dense model of 1008 proteins. The researchers then set that image in motion, running independent Brownian dynamics simulations governed by varying energetic descriptions of intermolecular interactions. The simplest description included only the excluded volume effect: no molecule could take the space of another molecule. The most complex scenario they ran included excluded volume, electrostatic interactions, and favorable short-range hydrophobic interactions. The more complex simulations performed surprisingly well when asked to predict molecular behaviors, such as diffusion and stability, in the *E. coli* cytoplasm.

The model was able to match experimental observations of how quickly green fluorescent protein diffuses in the *E. coli* cytoplasm. And it was able to predict the greater stability of the unfolded state of the protein CRABP, cellular retinoic acid binding protein, over the folded state inside *E. coli*. Although the presence of close neighbors (crowding) generally stabilizes a large folded protein, the specific electrostatic and hydrophobic interactions of unfolded CRABP with other cytoplasmic proteins counteract the crowding effect.

"What this doesn't mean," Elcock emphasizes, "is that crowding effects are unimportant. It means that crowding is only part of the story."

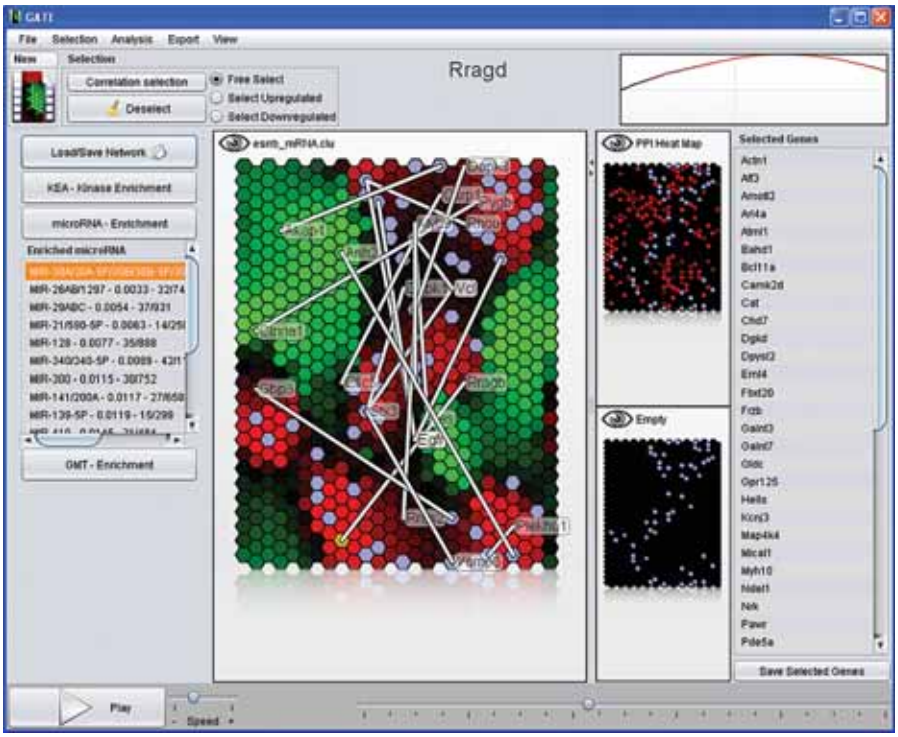
A computational box of 1008 proteins is still a far stretch from the complex *E. coli* cytoplasm, says **Allen Minton, PhD**, a pioneer in the study of crowding effects and researcher of physical biochemistry at the National Institutes of Health. "But there are a lot of questions that only this type of computation can answer," he says. "From a computational point of view, it is a real tour-de-force."

—By Louisa Dalton

Animating Molecular Biology

These days, molecular biologists often gather data over a period of time—observing shifts as they occur inside groups of cells undergoing natural changes. The researchers then face the daunting task of making sense of it all. Now, computational biologists have devised a software program to easily visualize and analyze these mountains of time-series data in animated movie form. While these flicks might never

appear at a theater near you, scientists studying such disparate areas as stem cell development and the microbial communities of the Pacific Ocean will be playing them on their computer screens to explore how all the genes and proteins



This screenshot from the GATE software program shows RNA expression levels from experiments on stem cells that were genetically manipulated to differentiate. Each hexagon represents a single gene; red hexagons are genes with increased RNA levels and green are those with decreased levels. Commonalities among gene annotations are highlighted in blue, and white lines represent known interactions between proteins. GATE movies animate a series of these images to show changes over time. Courtesy of Avi Ma'ayan.

of a cell type or organism change over the timespan of experiments.

"This is a tool that is really useful for interrogating datasets collected as a time series at multiple layers of regulation," says **Avi Ma'ayan, PhD**, assistant professor of pharmacology and systems therapeutics at the Mount Sinai School of Medicine who spearheads the project. "It allows you to form hypotheses for future experimentation very quickly."

The software, called GATE (Grid Analysis of Time-Series Expression), was originally designed to analyze clustered gene and protein expression data taken at various time points during stem cell development, Ma'ayan says. This work, led by **Ihor Lemischka, PhD**, Mount Sinai professor of gene and cell medicine, was published in *Nature* in November 2009. "It was a relatively simple approach but it hadn't been done before," Ma'ayan says. But Ma'ayan's group realized that GATE movies would be even more useful if they could incorporate existing biological data, such as libraries of protein-protein interactions or annotations of genes' functions. The updated software was further described in *Bioinformatics* in January 2010.

The movies GATE generates show a 2-D honeycomb of small hexagons, each representing a single gene or protein and colored red (for increased expression) or green (for decreased). The hexagons are clustered near other genes or proteins with similar behavior patterns in the experiments. When the movie plays, waves of color shift across the grid, representing molecular shifts over the time course of the experimental series. Although GATE was developed for stem cell biologists, its potential applications are broad, Ma'ayan says. Recently, he was contacted by a group at the University of British Columbia that wants to use the software to analyze changes in marine flora and fauna in the Pacific Ocean. In this case, the movies will look at changes both over time and distance, as the researchers sample further from the coast.

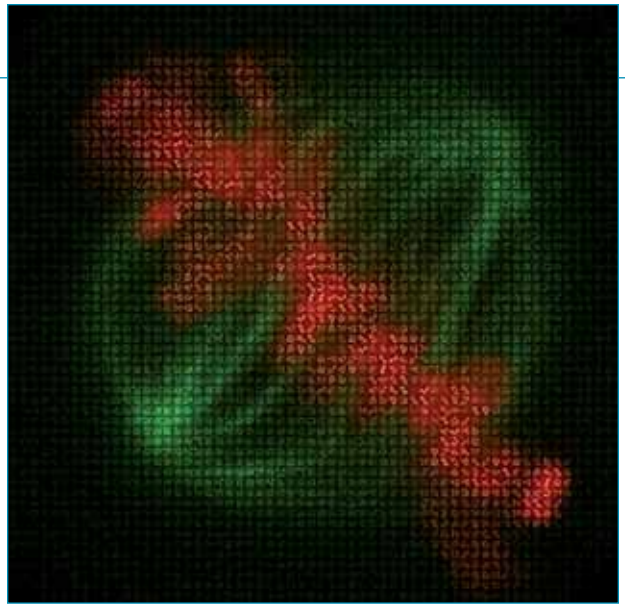
Oliver Hofmann, PhD, a computational biology research scientist at the Harvard School of Public Health, says the technology will be very useful for the field of molecular biology. "It's a very neat way of visualizing time series," he says. "But it's not just a pretty picture you can look at. You can explore it interactively too." It is still difficult to coordinate more than two types of data timecourses in GATE, Hofmann says, and Ma'ayan agrees. He says their to-do list includes plans to better overlay multiple movies.

—By Rachel Tompa, PhD

Capturing Mitosis Genes in Action

During the one-hour drama that is human cell division, many genes enter and exit the stage. Until now, researchers did not know the identities of many of these actors, nor understand their various roles. Now, using a combination of high-throughput screening methods, time-resolved movies and a supervised machine-learning algorithm, researchers have identified 572 genes that are involved in mitosis in human cells. The raw data and images are available to the research community at www.mitocheck.org.

"Researchers can go to the database, do a clustering analysis, and extract the genes that are most interesting from their research question point of view," says **Jan Ellenberg, PhD**, head of the Cell Biology and Biophysics Unit at the European Molecular Biology Laboratory and senior author on the paper pub-



This microscopy image captures the mitotic spindle (green) and the chromosomes (red) of a dividing cell. EMBL researchers videotaped mitosis for 22,000 different gene knockouts. Videos and data for all 22,000 genes are available at www.mitocheck.org. Courtesy of Thomas Walter & Jutta Bulkescher / EMBL.

lished in *Nature* on April 1, 2010.

The research addressed an age-old problem in the study of cell division, Ellenberg says. "We didn't know all the genes or the proteins involved," he says. "So we decided that we had to do this gene discovery ourselves."

First, Ellenberg and his colleagues in the Mitocheck consortium developed the technology to do systematic high throughput screens of multiple samples of all 22,000 human genes and then visually match each knockout to a phenotype. They relied on RNA interference to knock out each of the approximately 22,000 individual genes. They then printed more than 384 of these samples at a time on microarray chips. Because mitosis occurs transiently (approximately once every 24 hours), the researchers developed microscopes to capture movies of each sample from four such microarrays in parallel over the course of 48 hours.

Analysis of so much visual data—nearly 200,000 movies—required supervised machine learning. First, a human expert annotated examples of different morphologies observed within the movies. A computer then extracted a numerical signature with 200 different parameters that it correlated with those characteristics. After iterative training with movies of just 3000 different individual cells, the computer analyzed additional movies and identified phenotypes with 90 percent accuracy. The researchers also developed new distance measures for clustering algorithms to categorize the differences in cell division behavior.

The scale of these experiments and the use of time-lapse imaging over two days are "unparalleled and nothing short of phenomenal," says **Anne Carpenter, PhD**, director of the Imaging Platform at the Broad Institute, who was not involved in the research. "[The insights into mitosis are] just the tip of the iceberg of the knowledge that will be extracted from this single experiment," Carpenter says.

The researchers' next project, called Mitosys, will explore the molecular activity of the 572 mitosis-related genes.

—By Sarah A. Webb, PhD

Cells' Collaborative Middle Management

Like corporate and governmental organizations, cells rely on middle managers to keep things running smoothly. These "middle managers" function as a critical bridge that controls the flow of information traffic. According to recent research, however, the middle managers often partner with one another, ensuring that the failure of one manager doesn't bring down the entire organization. Moreover, this partnering becomes more extensive in more complex organisms.

"Understanding the system isn't about the function of the individual parts. [It's about] understanding the importance of these information flow bottlenecks and how natural systems get around them," says **Mark Gerstein, PhD**, professor of bioinformatics at Yale University. He and his colleagues have been studying networks of genes and transcription factors to describe the information flow within cells.

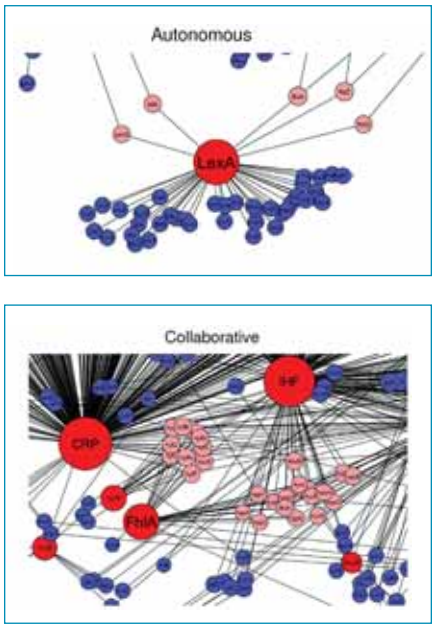
The work serves as part of a larger effort within Gerstein's group to develop real-world analogies to explain how biological systems use and process information. Previously, the group had shown that hierarchies in biological regulatory systems resemble directed social structures such as governments and corporations. That study, published in *PNAS* in 2006, found that "middle managers rule," Gerstein says. Transcription factors in the middle layers of the networks have the most regulatory interactions with other genes. "The genes in the middle are much more essential. If you knock them out, the organism is much more likely to die."

In a paper published online in *PNAS* in March 2010, Gerstein and his colleagues took that work a step farther, seeking to understand how cells avoid failure at the sites of middle manager bottlenecks in five species ranging from *E. coli* to humans. First, they identified which genes are regulated by other genes in each organism. They then stacked the levels of regulators in hierarchies and placed them between two extreme types of social hierarchies, autocratic and democratic, and showed cellular regulatory hierarchies have "intermediate" structures. They found that, in all five organisms, coregulation happens most at the middle level and least at the bottom. And more complex organisms exhibit more collaborative, "democratic" regulatory structures with more interconnections. For example, yeast has about one regulator for every 25 targets whereas in humans the ratio is much smaller, about one to 10.

"The parallels between government structure and regulatory network structure are provocative," says **Trey Ideker, PhD**, associate professor of medicine and bioengineering at the University of California, San Diego, who was not involved with the study. One question, says **M. Madan Babu, PhD**, an investigator in the MRC-Laboratory of Molecular Biology at the University of Cambridge, is the function of these hierarchies within a cell. "Are they really important? Or

are they something that is emergent because of the complexity of the system and has no consequence whatsoever?"

Regulatory networks are definitely important for organism function, Gerstein notes. So the question of whether the networks emerged in response to complex roles or the system's complexity allows organisms to carry on these complex interaction is a "chicken and egg type of issue."



Diagrams of hierarchical networks: In an autocratic network, such as the military, there is a clear chain of command. In a democratic network, many members interact and regulate each other. And in an intermediate network, such as exists within a law firm and many cells, the hierarchy shares features of both types. As biological organisms become increasingly complex, their organization becomes more democratic.

—By Sarah A. Webb, PhD

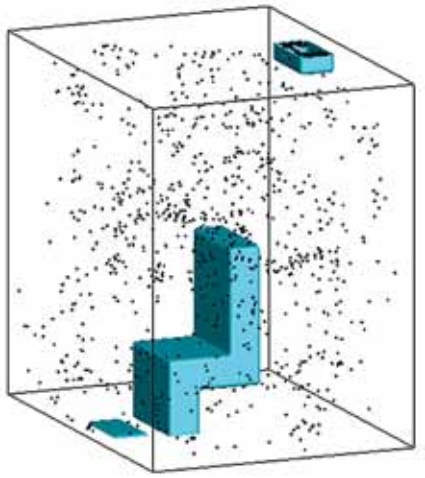
Hot Bodies a Lure for Unseen Specks

We can't see them, but tiny particles—dust, pollen, microbes, and the like—swirl around us in complicated, turbulent pathways. New numerical simulations suggest that, at least in tiny indoor spaces, our body heat may pull them even closer, where they have a better chance of eventually landing in our lungs.

"The conventional wisdom is that the thermal plume from your body protects you from particles falling from above," says **John B. McLaughlin, PhD**, professor of chemical and biomolecular engineering at Clarkson University and coauthor of the study. "We found that, in our small room at least, that is not true." Such findings can help engineers design better ventilation systems, McLaughlin says. "Studies have shown that schoolchildren learn more and office workers are more productive in environments where the concentration of particles in the air is very low."

Airflow dynamics are notoriously tough to model computationally, largely because of the huge range of physical scales in equations for turbulent fluids. McLaughlin and his colleagues used a direct numerical simulation approach that offers accuracy but requires intensive computational resources. Their computational models of airflow and particle paths were built in a 4.8-square-meter virtual room at two-centimeter resolution using three-millisecond time steps over about three minutes of total simulated time. In each simulation, a mannequin sits motionless in the middle of the room. A stream of air suffused with particles—each with the density of sand and about the size of a grain of pollen—shoots up through a floor vent in front of the chair. Particles fan out throughout the room, with a ceiling vent as the only exit.

In simulations where the mannequin was bestowed with realistic body heat, researchers could see the hot air surging off the body and interacting with particulates. This thermal plume pulled rising particles directly into the mannequin's breathing zone. At the same time, the plume blocked the path of particles traveling near the ceiling, forcing them to fall down into the mannequin's personal space, doubling the trapping effect of the plume. The work was presented in March 2010 at the American Physical Society meeting in Portland, Oregon.



The positions of 2-micrometer particles inside a 20-degree-Celsius room with a mannequin heated to 25 degrees Celsius, three minutes after particles were released through a floor event. In this simulation, 31 out of 1000 particles fell directly onto the mannequin's warm body; none managed to leave the room through the ceiling vent. Yet when the mannequin was the same temperature as the room, no particles fell onto the body, and 160 out of 1000 particles escaped. Results were similar for simulations with 10-micrometer-diameter particles.

"The computational and the experimental go hand in hand when studying complex turbulent flows such as those around human beings," says **Mark N. Glauser, PhD**, professor of mechanical and aerospace engineering at Syracuse University, whose empirical results helped guide McLaughlin's modeling. Fundamentally, experiments can help validate computational models and give physical insights that spur new simulations. "Then the simulation tools can be used to probe a broader range of parameter space 'virtually,' as well as look in more detail at flow physics," Glauser says. For example, the models from McLaughlin's team can track individual particles in a turbulent flow—a feat that's nearly impossible in real-life experiments.

—By Regina Nuzzo, PhD

Brain Folding

In the four months before birth, a fetus's brain grows from a smooth tube of neurons into a highly crinkled, convolved mass of tissue. Because the cerebral cortex has a surface area nearly three times as big as that of its skull cavity, scientists have proposed that this real-estate-space squeeze is what drives the brain's folding process. Now results

from a computational three-dimensional geometric model agree that the skull does help guide the wrinkling—but they also suggest that a growing brain folds up regardless of its container.

"Mechanical constraints imposed by the skull are important regulators," says **Tianming Liu**, **PhD**, assistant professor of computer science at University of Georgia and lead author on the study, which was published in May 2010 in the *Journal of Theoretical Biology*. "But our simulations indicate that skull constraint is not necessarily the dominant mechanism."

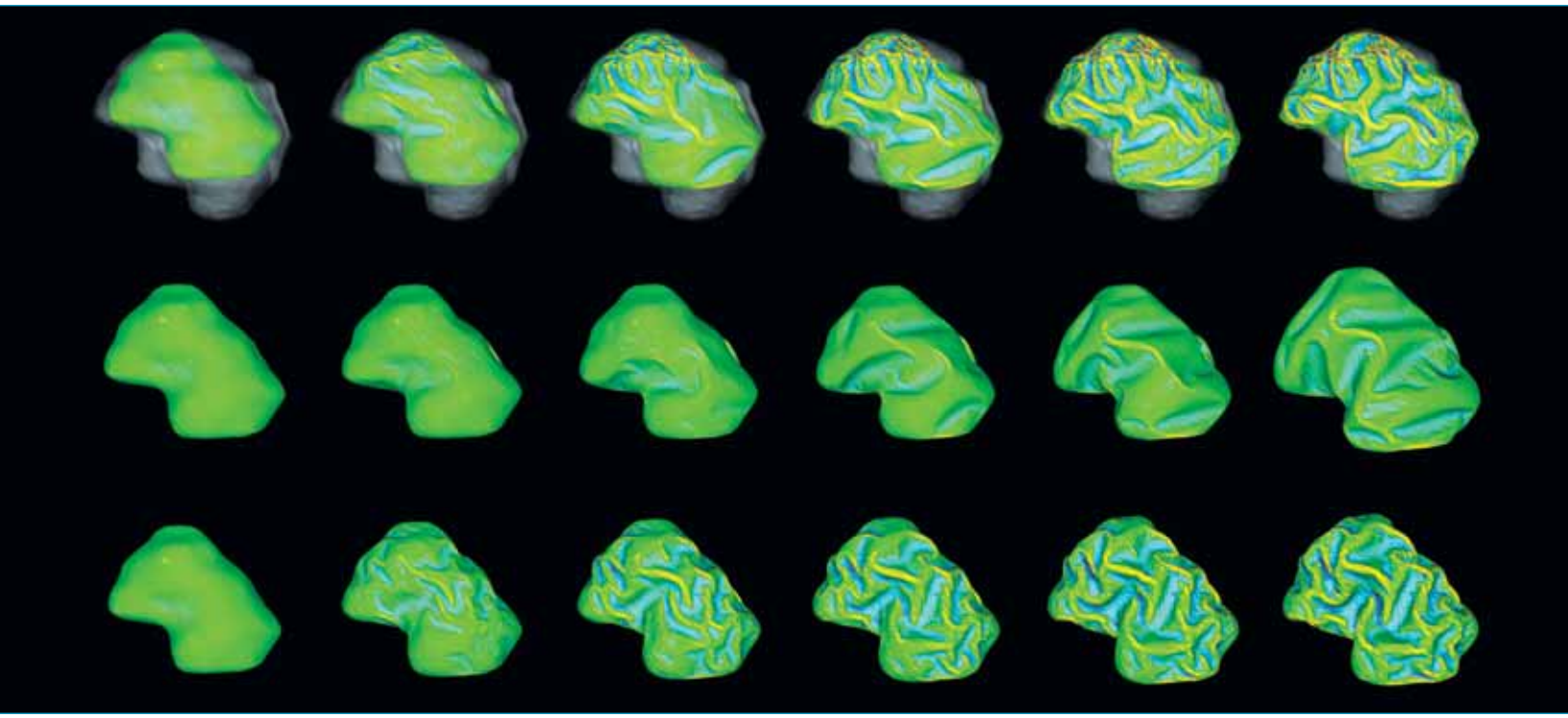
The computational model underneath the simulations had two main features: geometric constraints of the skull, and partial differential equations that model biological processes driving the growth of neurons. To start off the simulation, researchers used MRI data from the brains of two human fetuses; then solutions to the differential equations guided the changing surface geometry of the cortex.

The team simulated how the cortex grows under various conditions: without a skull, with a skull of fixed size, and with a skull that grows at the same time as the brain does. As expected, brains grown in a skull were more convoluted than those allowed to develop unfettered. But even without a skull to confine it, a cortex will still fold in on itself, results showed. This happens as a natural response in a fast-growing cortex, as the tissue attempts to reduce the increasing mechanical tension among axons, dendrites, and neuroglia, Liu says.

Tweaking other parameters in the model revealed how cellular growth affects these folding patterns. When neurons themselves grow rapidly — during synapse development and neuron dendritic projection, for example— the cortical folding increases dramatically too. And when certain areas of the cortex grow more quickly than others, this imbalance subtly shapes what kinds of folds become most prominent.

Computational models can help explain normal brain development as well as what happens when things go wrong, says **Bernard S. Chang**, **MD**, assistant professor of neurology at Harvard Medical School. For example, in some forms of microcephaly, the brain surface is almost completely smooth with no folds; in others, the folding is normal. "A model that predicts how folding is affected by the skull's physical constraints might help us to understand why some patients have one form and not another," he says. Since animal models don't capture the complexity of the human brain, and doing repeated MRIs of developing fetuses for research isn't feasible, Chang says, "we need to rely on these theoretical models as tools to help us understand what we're observing clinically."

—By Regina Nuzzo, PhD □



Growth of the cortex under different assumptions. From left to right, the images show simulated development of the cortex over time. The cortex grows (a) within a skull of fixed size, (b) without a skull, or (c) within a skull that also grows at the same time (which corresponds to a fetus' developing skull). Major cortical folds developed much earlier and faster during simulations with skull constraints. But the cortex increases its surface area and convolutes itself to reduce the fast-growing internal tension, with or without skull constraint.