DIVERSE DISCIPLINES ONE COMMUNITY

Biomedical Computation REVIEW

Published by Simbios, a National Center for Biomedical Computing

On Simulating Growth and Form

PLUS

Assembling The Aging Puzzle: Computation Helps Connect the Pieces

Spring 2008

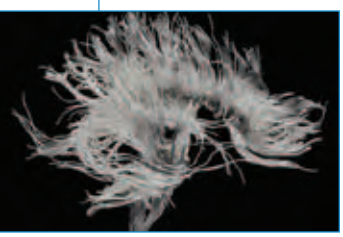
contents

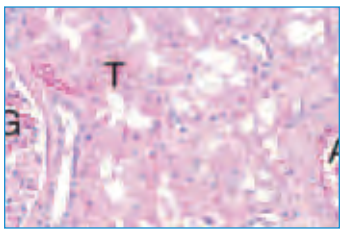
ContentsSpring 2008

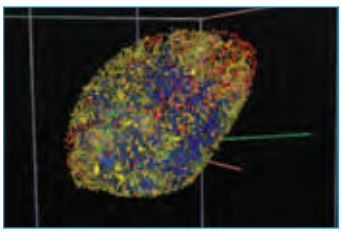
FEATURES

DEPARTMENTS

COVER ART BY
JAMES BENNETT OF AFFILIATED DESIGN







Spring 2008
Volume 4, Issue 2
ISSN 1557-3192

For general inquiries, subscriptions, or letters to the editor, visit our website at www.biomedicalcomputationreview.org

Office
Biomedical Computation Review
Stanford University
318 Campus Drive
Clark Center Room S231
Stanford, CA 94305-5444

Biomedical Computation Review is published quarterly by Simbios National Center for Biomedical Computing and supported by the National Institutes of Health through the NIH Roadmap for Medical Research Grant U54 GM072970. Information on the National Centers for Biomedical Computing can be obtained from http://nihroadmap.nih.gov/bioinformatics. The NIH program and science officers for Simbios are:

Peter Lyster, PhD (NIGMS)
Jennie Larkin, PhD (NHLBI)
Jennifer Couch, PhD (NCI)
Semahat Demir, PhD (NSF)
Peter Highnam, PhD (NCRR)
Jerry Li, MD, PhD (NIGMS)
Richard Morris, PhD (NIAID)
Grace Peng, PhD (NIBIB)
David Thomassen, PhD (DOE)
Ronald J. White, PhD (NASA/USRA)
Jane Ye, PhD (NLM)
Yuan Liu, PhD (NINDS)

GuestEditorial

BY DR. ANDREA CALIFANO

Taking the leap: from single genes to the molecular choreography of the cell

The Human Genome Project has spurred extraordinary developments in our ability to characterize cellular systems in high-throughput fashion. Polymorphism, methylation, gene expression, and proteomics profiles are just a few of the data modalities that we could not have hoped to measure without a genomic sequence anchor. Yet, despite this flood of new data, we are still struggling in our understanding of the complex molecular choreographies that ultimately determine physiological and pathological phenotypes. The reasons are multifold and define some of today's grand challenges of molecular biology and medicine. Two issues, however, perhaps most significantly contribute to our still limited understanding of normal and disease-related cellular processes.

First, while we have now characterized the majority of our gene repertoire and, to some extent, its byproducts (e.g., microRNA, proteins, etc.), cellular processes are not regulated by individual genes but rather by complex webs of molecular interactions that are precisely time- and space-coordinated and exquisitely phenotype specific. Unfortunately, we still know far too little about these interactions. Our estimate, is that only 1 to 20 percent of the full human interactome may be represented in the literature and in databases. Also, as reported at the 2007 Hinxton Interactome meeting in the United Kingdom, even for yeast, as many as 50 percent of the protein-protein interactions reported in individual articles may not be reproducible. Thus, our existing molecular interaction maps constitute a small, partially incorrect, and certainly largely incomplete view of the interactions that *could exist* rather than the ones that *are implemented* in a specific cellular context of interest.

Second, we are still mostly looking at individual genes as the determinants of both normal cellular processes and of their dysregulation in disease. For instance, most genome-wide disease association studies still test polymorphisms either in isolation or within small contiguous genomic regions, rather than in the context of interacting gene products. Even worse, individual factors such as copy number, chromatin methylation and acetylation, SNPs, expression levels, and post-translational modifications, are also being mostly studied in isolation, rather than integrated within pathways. Yet the same function may be regulated or dysregulated by any of these factors individually or in combination, each one contributing a very small fraction of the total penetrance. As a result, the total number of distinct molecular phenotypes may be as high as the product of all genetic/epigenetic variants across all interactions in a set of synergistic pathways: a very high number indeed!

To capture any unifying principles behind the heterogeneity of specific diseases, it will be crucial that we start integrating all available information over reasonably accurate, genome-wide interaction maps. To accomplish this goal, we may need once again to come together as a community of computational and experimental scientists to embark on an even greater challenge than the human genome project: the genome-wide, high-resolution mapping of molecular interactions, within highly specific cellular phenotypes. These maps will provide critical information on how several gene products work together to implement specific cellular functions. Additionally, they will provide an extraordinary resource to integrate disparate information on normal and disease-related tissue. Specifically, we are starting to see that distinct molecular determinants of the same cellular phenotype cluster within relatively compact regions of these molecular interaction maps, rather than being randomly distributed.

While promising approaches have been demonstrated to accomplish this goal both experimentally and computationally, the end-goal of this activity cannot be reached by using either approach alone. Indeed even once the experimental data becomes available, complex, multi-faceted computational challenges will have to be addressed. These range from an appropriate ontological classification of the cellular phenotypes, processes, and molecular species, to the 2D/3D visualization of the intricate interaction networks (possibly in a time- and space-dependent manner), to the development of reverse engineering approaches that leverage both functional and structural data, to the identification of relevant biomedical problems. Sound familiar? It should, because these are precisely the computational challenges that the seven National Centers for Biomedical Computing are currently trying to address. □

DETAILS

Dr. Andrea Califano is a professor of biomedical informatics at Columbia University, associate director of the Herbert Irving Comprehensive Cancer Center, and principal investigator for MAGNet—the National Center for Multiscale Analysis of Genomic and Cellular Networks (MAGNet) at Columbia University.

NewsBytes

Pore Picture Construction

Like puzzles? Here's a tough one: Try figuring out the construction of a nearly 500-piece machine without blueprints or a complete picture. Biologists have now accomplished just such a feat, working out the protein-by-protein structure of an important cellular assembly called the nuclear pore complex. Their success depended on computationally combining incomplete imaging information with bits and pieces of structural data from all sorts of different experiments.

"It is as if we use many tiny lights, each of which shines from a different perspective, to illuminate every part of the whole structure," says **Andrej Sali, PhD**, a coauthor and professor of biopharmaceutical sciences and pharmaceutical chemistry at the University of California, San Francisco. "We are able to use information from many sources, even sources that haven't been traditionally used for structure determination." As described in *Nature* in November 2007, this gleaning strategy should be helpful in determining the structures of many hard-to-pin-down cellular complexes.

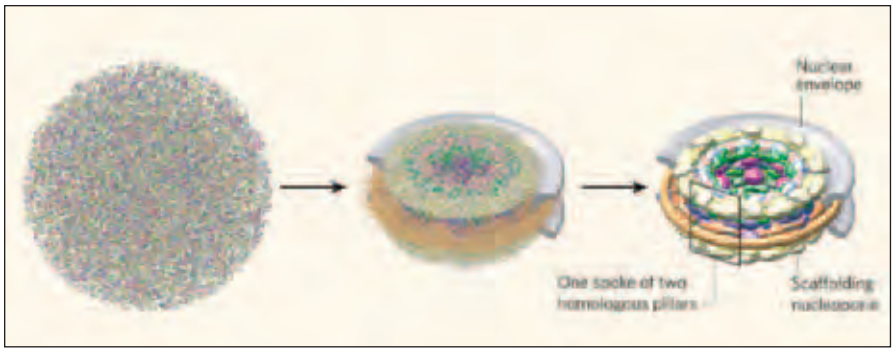
The nuclear pore complex (NPC) is a gatekeeper of the nucleus, a 456-protein assembly in the shape of a thick donut spanning the nuclear membrane. Scientists know the general structure of eight spokes that make up the donut and have a roster of its proteins. But where each protein fits has been difficult to pin down.

The challenge is that no one tool images the details of complexes this size. Electron microscopy reveals the overall shape and outline but not individual proteins. NMR spectroscopy and X-ray crystallography, on the other hand, show individual proteins in stark relief, yet can't be used on the whole assembly.

In collaboration with two groups of experimental biologists at Rockefeller University led by **Michael Rout, PhD**, and **Brian Chait, PhD**, Sali's team found a way of combing structural information from disparate sources and computationally putting all the bits together to create a low-resolution image of the entire complex. They included experimental data from, for example, affinity purification assays (which indicate interactions between proteins) and ultracentrifugation (which reports on protein shape). At least seven different experimental techniques were used to produce structural data.

With the data in hand, they translated each piece into a "spatial restraint"—a mathematical probability of the structure's geometry. For example, one restraint might indicate that protein A very likely interacts with protein W. Then the computer, starting with a random configuration of the proteins, moved them step-by-step in a direction that minimized violations of restraints. This process was repeated until the group had acquired 1,000 optimized structures that each satisfied the restraints. (In total, that took 200,000 trials run on 200 CPUs for 30 days.) The small variations between those 1,000 structures were then combined into a slightly blurry final image.

> "We are able to use information from many sources, even sources that haven't been traditionally used for structure determination," says Andrej Sali.



Starting from a random mess of proteins (456 beads), experimenters ended with the structure of the nuclear pore complex. They did so by directing a computer to move the beads in any direction that minimized pre-programmed structural restraints—as if the proteins were gradually tugged towards proper placement. The final structure is an arrangement with the least cumulative "tug" or structural restraint. Courtesy of Andrej Sali. Reprinted by permission from Macmillan Publishers Ltd: **Nature** ***450:621-622, 2007***

Sali was struck by the simplicity of the final structure. "If you look at electron scanning microscope pictures of the nuclear pore complex, and imagine how many proteins are involved, you think, 'This is a mess! How did this evolve?'" But once the scientists began analyzing the protein architecture, they noticed a number of symmetries and a simple three-layer architecture: one layer to hold the pore to the membrane, one layer to facilitate transport of molecules through the pore, and a final scaffold layer to hold it all together. "It is not hard to imagine the evolution," Sali says.

Establishing the protein architecture

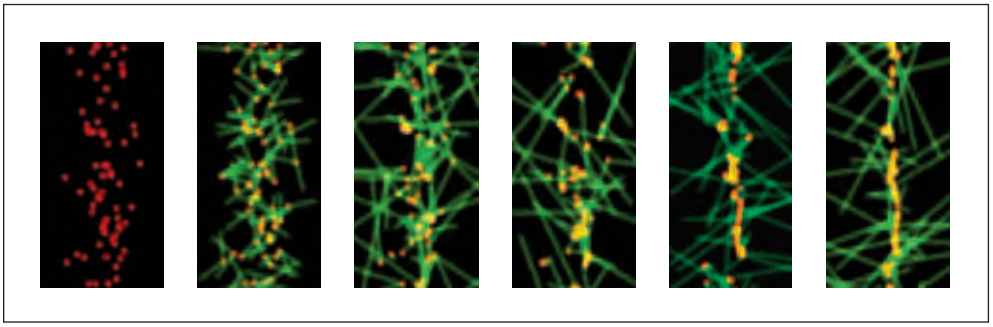
is also a huge step in coming to a better understanding of how the NPC facilitates controlled transport of molecules in and out of the nucleus. The group of **Klaus Schulten, PhD**, director of the theoretical biophysics group at the University of Illinois at Urbana-Champaign, is using this structure to study the transport mechanism. "The recipe that the investigators found for combining many experiments into one picture worked so consistently and so coherently across many independent trial predictions that the results must be true," Schulten says. "Already now the relatively low resolution structure helps us to understand much better how the NPC organizes its complex function." Chait, Rout, and Sali are now working on a high-resolution structure, with detail down to the atomic level.

—By Louisa Dalton

Cell Division's Surprise Twist

During the final step of cell division, a ring of proteins pinches the cell in two—a process often likened to a purse string drawing shut. The analogy evokes a picture of thread-like proteins wrapping around the cell's middle in an orderly fashion. But the mechanics of this "contractile ring"—detailed for the first time in the January 4th issue of *Science*—turn out to be far more intricate and chaotic.

"The answer—which is very exciting and surprising—is that it's a completely random, unguided process that works perfectly," says **Thomas D. Pollard, MD**, professor of molecular, cellular, and developmental biology at Yale University. His team used a combination of computer modeling and high-resolution microscopy to show that ring assembly in fission yeast follows a dynamic "search, capture, pull, and release" mechanism. The general principles are likely to be the same in higher organisms, Pollard says.



Simulations of the assembly of the contractile ring in fission yeast: Nodes (red) sprout actin filaments (green) in a random network. Myosin proteins in one node randomly encounter, capture, pull on, and then release actin filaments growing from another node. Repeated iterations of this process eventually draw the nodes together in a ring. Courtesy of Thomas Pollard.

Their work follows decades of scientific exploration on the topic, he says. Experiments in the 1970s revealed that myosin and actin—the same proteins that make muscles contract—are key players. Genetic studies later identified a complete "parts list" of proteins required (about 50). Recently, scientists observed that the process begins with a broad band of dots—or "nodes"—appearing around the equator of the cell. Pollard's team pinpointed the composition of these nodes; among other proteins, they contain formin, which polymerizes actin filaments, and myosin, which interacts with actin.

> Ring assembly in fission yeast follows a dynamic "search, capture, pull, and release" mechanism.

Their observations suggested a simple and elegant model for ring assembly: Nodes grow actin filaments that are captured by myosins in neighboring nodes to make a continuous chain; then the myosins pull the chain closed. But, a Monte Carlo simulation of the scenario gave disappointing results—instead of forming a ring, the proteins disbanded into isolated clumps. "So we were missing something," Pollard says.

Back in the lab, they carefully measured the movements of fluorescently tagged actin and myosin using high-resolution time-lapse microscopy in live cells. What they saw was unexpected: "The nodes move around in a completely crazy way," Pollard says, "They go at almost 360 degrees. They don't all head to the equator at all. They start and stop."

This suggested a different model of ring assembly where the nodes form transient rather than permanent connections: nodes sprout actin filaments in random directions; these filaments encounter myosins in nearby nodes; the myosins capture, pull on, and then release the actin. Repeated iterations eventually draw the nodes together in a ring.

"You'd swear after two minutes of this 10-minute process, this thing was never going to get there. Even after five minutes, even after seven minutes, it's a mess," Pollard says. "But it turns out that just by this completely random process of searching, getting captured, moving intermittently, and then breaking connections, it always works."

A simulation of this model formed a virtual ring in the same time it takes a

live cell. "The gratifying thing is that not only does it make a ring, but it makes it in 10 minutes—which is actually a big constraint," Pollard says.

"It's fascinating work," comments **Alex Mogilner, PhD**, professor of neurobiology, physiology and behavior and of mathematics at the University of California, Davis. "I think there will be more surprises in the future," he says, "but they nailed the essence of what's going on."

—By Kristin Sainani, PhD

Modeling the Deformable Body

August 2007 saw a surge of new open-source software for simulating musculoskeletal movement. In addition to OpenSim 1.0 (described in the Fall 2007 issue of this magazine), FEBio arrived on the scene. While OpenSim uses rigid body mechanics—simulating the body moving essentially as a series of segments attached at joints—FEBio (Finite Elements for Biomechanics) addresses the other part of the problem. It can simulate how movement deforms and places stresses upon solid parts of the body such as muscles, tendons, ligaments, cartilage and bone.

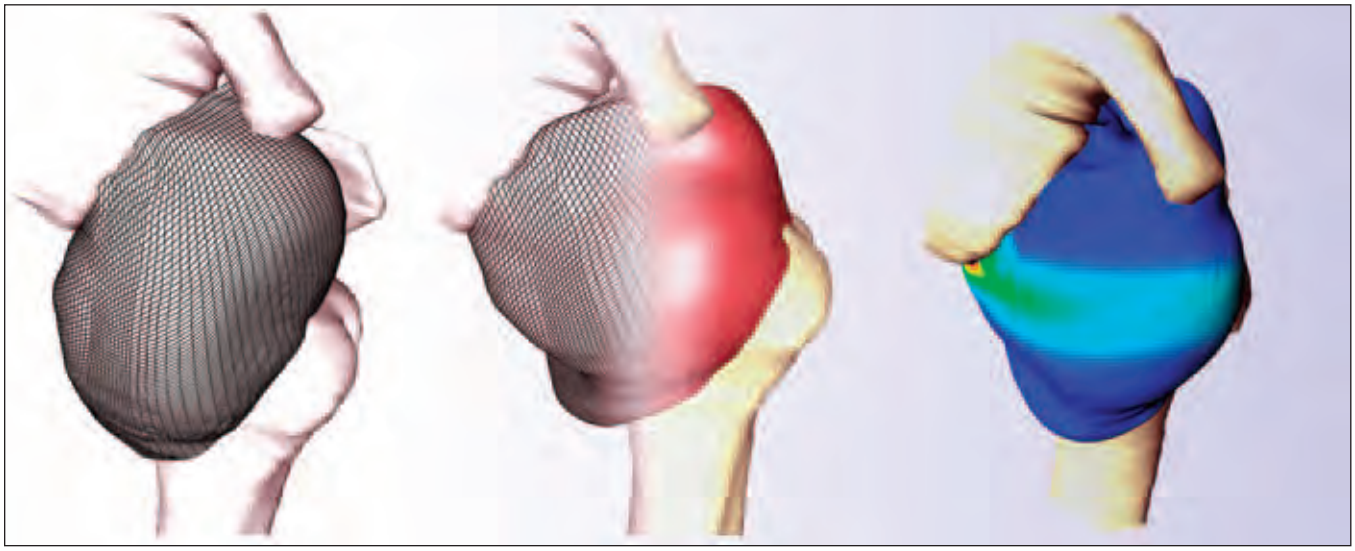
Created by **Jeff Weiss, PhD**, associate professor of bioengineering at the University of Utah, and his colleagues, FEBio already has 200 to 250 users. "Initially we developed FEBio for our use in-house," says Weiss, "but we saw the potential for it to be a really popular tool in the research community and decided to make it available to everyone."

Before now, biomechanics researchers studying the solid mechanics of soft tissue have relied upon costly general-purpose finite-element programs such as Abaqus or LS-DYNA. But because these programs are proprietary, it's hard to add new features to the code. "We saw that as a major shortcoming in our field," says Weiss. So he and his colleagues tailored FEBio to address the kinds of problems that come up in biomechanics.

In addition to FEBio itself, Weiss and his colleagues also released programs that allow users to prepare their models in advance of using FEBio (PreView) and to analyze and visualize the results of an FEBio simulation (PostView). "That's one of the advantages of FEBio," says **Steve Maas**, a software developer who works with Weiss. "You can do your model creation and post-processing on your own computer and use a high performance computer only for the FEBio step."

"Initially we developed FEBio for our use in-house," says Jeff Weiss, "but we saw the potential for it to be a really popular tool in the research community..."

FEBio's users come from many different disciplines including orthopedics, ophthalmology and cardiovascular mechanics. Weiss himself has used FEBio for a variety of research projects including a study of hip stresses in



In this FEBio model of a shoulder capsule—the soft-tissue envelope that surrounds the shoulder joint—the upper arm bone (the humerus) is moved upward and then rotated around its axis. The left image shows the initial undeformed mesh, the middle image shows an intermediate state, and the rightmost image shows the stresses on the capsule in the final deformed state (blue means low stress, red means high stress). During a shoulder examination, clinicians typically move the shoulder in various ways in an attempt to determine the source of a problem. Models of this kind could eventually help clinicians better understand the results of such tests. Courtesy of Jeff Weiss and Steve Maas.

people with displasia and a study of the shoulder capsule. He and his colleagues are also continuing to add new features to FEBio.

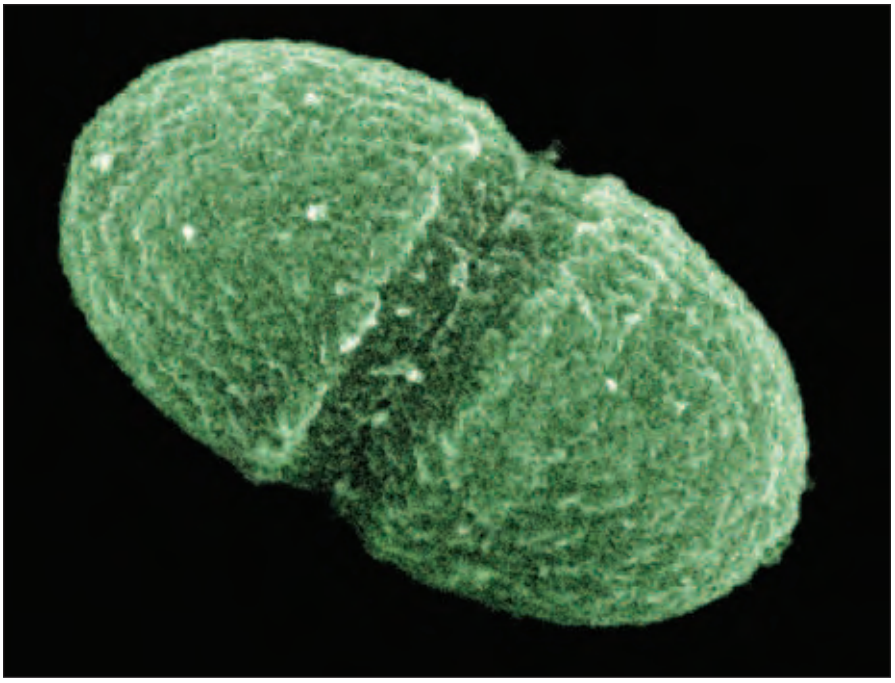
Weiss and one of the OpenSim creators **Scott Delp, PhD**, a professor of bioengineering at Stanford University, have begun a collaboration to link the two programs to address problems that can't be handled by either program alone. Although Delp's group has combined dynamics with finite element approaches in previous work (for example in a study of knee pain), "development of advanced methods in biomechanics would be accelerated if one could use two open-source programs connected in a straightforward way. I'm looking forward to that day," Delp says.
—By Katharine Miller

Discovering The Bugs Within

We are crawling with bugs. It might even be better to say that we are bugs. For every human cell in our bodies there may be ten or even a hundred other cells that aren't human at all. Yet many of these microbes are entirely unknown to science. To change that, the National Institutes of Health has just begun a five-year, $115 million Roadmap initiative called the Human Microbiome Project. It aims to find out what these bacteria, viruses, archaea and fungi are, how they function, and the ways they can keep us healthy or make us sick.

"There have been some tantalizing findings that gut flora influence things like obesity and irritable bowel disorder," says **Jane Peterson, PhD**, associate director of the Division of Extramural Research at the National Human Genome Research Institute and a program director for the project. "Ultimately, what we really want to understand is health as well as disease. What makes us healthy? Our microbes are a part of that."

But learning about these bugs has seemed like an overwhelming undertaking. Part of the problem is simply numbers: thousands of different species of microbes swarm on and in our bodies.



A human gut microbe. This bacterium,* Entercoccus faecalis*, which lives in the human gut, is just one type of microbe that will be studied as part of NIH's Human Microbiome Project. Courtesy: United States Department of Agriculture

The most obvious way to find out what they are is to understand their genomes. Unfortunately, sequencing these microbes is even harder than sequencing our own genome because most of the microbes have an obstreperous unwillingness to grow in isolation in a lab. They will only grow in the particular conditions of, say, our teeth, where they commune with a particular group of other microbes that create an agreeable environment.

Sequencing technology has been improving rapidly, however, bringing the task within reach now. "Metagenomic" techniques have been developed to study the genomes of many different microbes simultaneously, making it unnecessary to culture the microbes in the lab. In addition, modern sequencing machines can now produce millions of sequences in a day, compared to a few thousand in the past, and they do it less expensively.

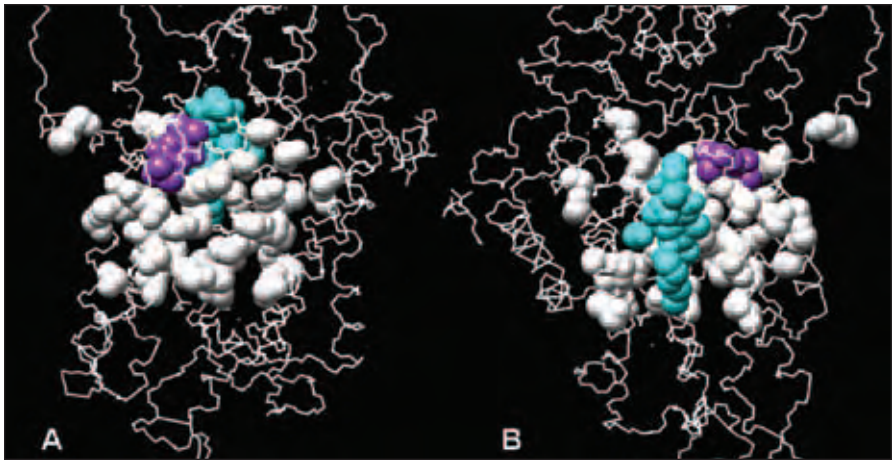
The Human Microbiome Project has already awarded $8.2 million to research groups around the country in 2007, and they currently have six requests for proposals out, due between February and May.

Analyzing the data from all these far-flung groups will require the development of new computational techniques. Genomic analysis already produces such

"Ultimately, what we really want to understand is health as well as disease. What makes us healthy? Our microbes are a part of that," says Jane Peterson.

vast quantities of data that it has pushed the computational capacity to make sense of it all, and the Human Microbiome Project will produce an order of magnitude more data than that. The project aims to coordinate the results from all the different groups, producing a single, publicly-available dataset.

The researchers involved in the project say the most exciting part is that they simply don't know what they're going to find. "You have to expect that there will be very many ways microbes are impacting our health that we don't know and maybe can't imagine at this point," says **George Weinstock, PhD** of Washington University in St. Louis. "We're hopeful it will have an impact on the level of the human genome project."
—By Julie J. Rehmeyer



The algorithm created by Bourne and his colleagues identified a possible Tamoxifen binding site (white spheres) on a protein called SERCA that regulates calcium levels within muscle cells. They also found that two known inhibitors of SERCA bind to areas (shown in purple and blue) within the same zone. This suggests that a side effect of Tamoxifen could be inhibition of this protein, Courtesy of Philip Bourne.

Side Effects *in silico*

Many new drugs carry a risk that they will cause more problems than they cure. That's because a drug intended to bind one protein might also bind others. In an effort to address that problem, researchers have developed a new computational approach that can potentially predict the protein interactions that cause drug side effects. The new algorithm has already provided a possible explanation for some side effects caused by the widely-used anti-cancer drug Tamoxifen. The same approach may also help find new targets for commercially-available drugs.

Traditional drug discovery searches for possible drugs that can bind to a known receptor protein. "We're doing essentially the reverse of that," says **Philip Bourne, PhD**, professor of pharmacology at the University of California, San Diego and lead author of the work published in the November 2007 issue of *PLoS Computational Biology*. "We've already got something that binds to a receptor. The issue is that it doesn't necessarily bind only to that receptor."

To find out what else the compound is binding, Bourne and his colleagues start with a database of potential receptors—what they call the "druggable proteome." They then test whether the compound binds to one or more secondary sites in receptors other than the primary target. Previous attempts to predict such drug-protein interactions have met with limited success. But **Lei Xie, PhD**, a member of Bourne's team, developed a novel algorithm that considers the evolutionary relationship among potential binding sites and also allows the receptor proteins to bend and move.

Combining these new parameters with an analysis of the receptors' shapes and binding characteristics yielded a powerful search tool capable of discovering off-target proteins missed by previous algorithms. Bourne's team then looked at whether the known functions of those off-target proteins could provide a logical explanation for a drug's known side-effects.

Bourne's team applied their algorithm to a family of cancer drugs that includes Tamoxifen. Known as selective estrogen receptor modulators (SERMs), this clan of drugs often causes unwanted side effects such as heart disease and ocular degeneration, both of which involve a disruption in cells' calcium balance. So Bourne's team was not surprised when their algorithm found Tamoxifen could bind a protein that regulates calcium levels within muscle cells (Sarcoplastic Reticulum Calcium ion channel ATPase protein (SERCA)). Specifically, the algorithm predicted that Tamoxifen inhibits SERCA's action (by binding near natural inhibitors' binding sites).

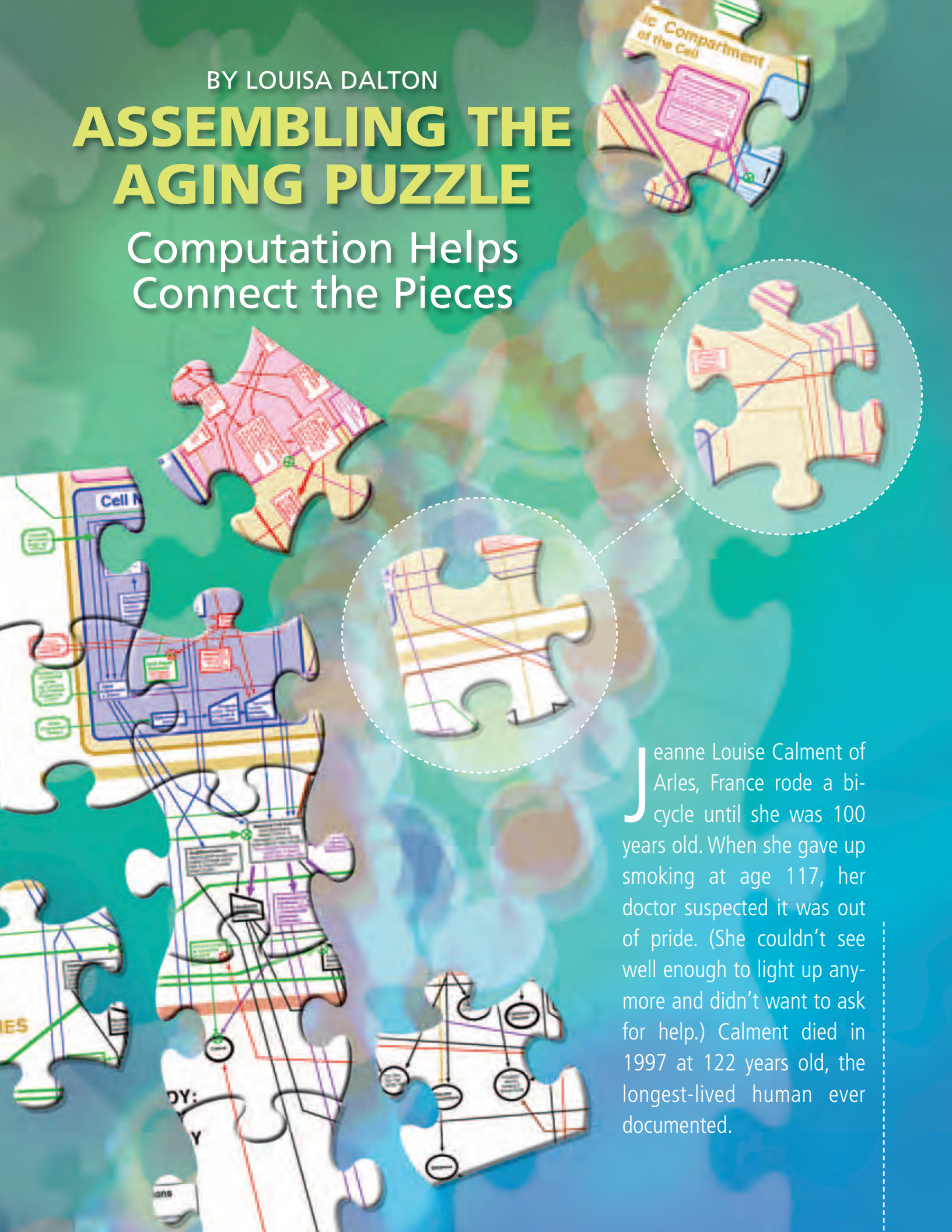
Bourne hopes the algorithm will help identify potential side effects of new compounds before they reach clinical trials, saving enormous amounts of money and time. In addition, the algorithm could help researchers design drugs with fewer side effects and find new targets for already-approved drugs. Indeed, Bourne's group has already found that existing Parkinson's disease drugs may help treat extreme drug-resistant tuberculosis.

"The potential value is huge if one could do this reliably," says **Robert Stroud, PhD**, a professor of biophysics and biochemistry at the University of California, San Francisco. Stroud cautioned, however, that more examples of the algorithm's ability to successfully identify off-target proteins are necessary before any definite conclusions can be drawn.
—Matthew Busse, PhD □

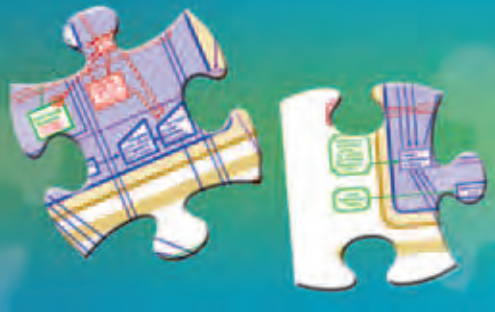
BY LOUISA DALTON

ASSEMBLING THE AGING PUZZLE

Computation Helps Connect the Pieces



Jeanne Louise Calment of Arles, France rode a bicycle until she was 100 years old. When she gave up smoking at age 117, her doctor suspected it was out of pride. (She couldn't see well enough to light up anymore and didn't want to ask for help.) Calment died in 1997 at 122 years old, the longest-lived human ever documented.



Puzzle pieces show portions of a diagram of biological interactions in human aging, from cellular pathways to whole body effects. The diagram was created by John Furber, founder of Legendary Pharmaceuticals. A complete image can be found at http://www.legendarypharma.com/chartbg.html. Courtesy of John Furber.

Calment's extraordinary vigor illustrates why aging researchers are so interested in studying the oldest among us, from spry seniors to long-lived yeast (*Saccharomyces cerevisiae*). We all want to know what keeps them ticking.

Any number of theories have been put forth to explain aging and identify its underlying causes. Many of them sound probable. "But the theories are way, way in front of the data," says **Stuart Kim, PhD**, professor of developmental biology and genetics at Stanford University. "Understanding aging is really in its beginning stages. I feel we aren't anywhere near a good understanding."

One thing is certain: genes play a role. They may or may not direct aging, but they do influence it. In the late 1980s, researchers reported in *Genetics* that a mutant of the *age-1* gene significantly extends the lifetime of the nematode worm (*Caenorhabditis elegans*). Since then, genetic researchers have been pinpointing other single genes that either slow, speed up, or alter the aging process. The count is up to more than 100 in the worm. A few age-associated genes have even been identified in humans.

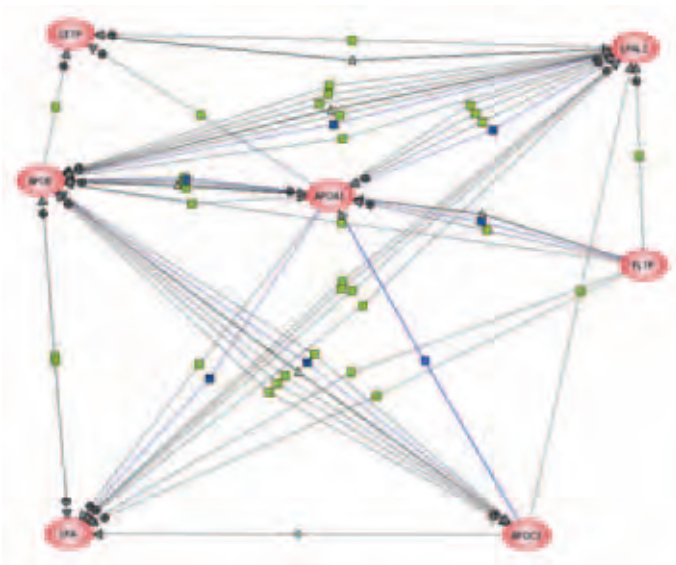
"But no one really knows how they fit together," says **Scott Pletcher, PhD**, an assistant professor in the Huffington Center on Aging at Baylor College of Medicine. The challenge now, he says, is to take that growing stack of largely disconnected single genes that affect aging and start putting them together.

Computational tools are helping aging researchers take that next step. The complexity and variability of aging itself, along with the fragmented nature of researchers' current understanding of aging, call for tools that can help scientists dig through mounds of data to find often subtle connections. Within the fields of comparative genomics, bioinformatics, and systems biology, for example, tools are now being developed and used to help assemble the puzzle of aging.

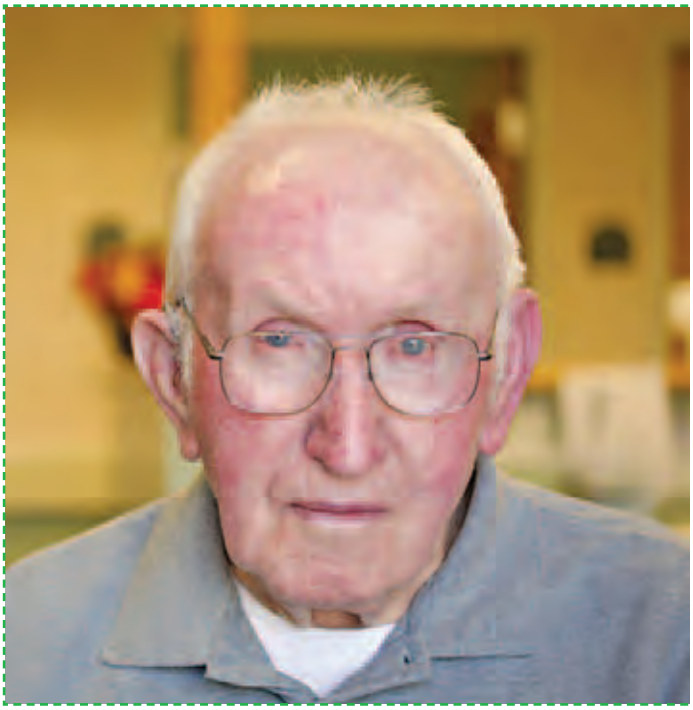
Buffering Genes in Centenarians

Nir Barzilai, MD, director of the Institute for Aging Research at the Albert Einstein College of Medicine, studies a group of more than 400 Ashkenazi Jewish centenarians and their children.

Barzilai and his colleagues have identified a number of longevity genotypes concentrated in this population, including a beneficial variant of CETP (the gene for cholesteryl ester transfer protein) associated with a decreased risk of heart disease. Barzilai has also noticed, however, that while centenarians are enriched with longevity genotypes, they also often concurrently harbor genotypes linked to early deaths from the diseases of old age. It seems a paradox. Barzilai and his colleagues guessed that this might occur because longevity genes buffer centenarians against the effects of the age-related disease genes.



This protein-protein interaction map indicates that the proteins lipoprotein(a) (LPA—lower left) and cholesteryl ester transfer protein (CETP—upper left) indirectly interact. This map was created from a search of the biomedical literature performed by Nir Barzilai and his colleagues after they determined that the favorable (CETP) genotype neutralizes the harmful effects of the unfavorable lipoprotein(a) genotype. The team used the GRID (General Repository of Interaction Database) to build the network and confirmed their findings using Pathway Architect software by Stratigene. Courtesy of Nir Barzilai.



Genes influence lifespan. When asked why they live to be so old, centenarians will say, "What do you mean, live to be so old? My mother was 102, and my grandfather was 108," says Nir Barzilai who studies large groups of the oldest of the old. "Well, we say, 'you have to tell us, did you eat yogurt all your life? What did you do?' ... But they didn't do anything special." Not one of them exercised regularly, a third were obese or overweight, and about a quarter of them smoked. "They have a family history, they didn't interact with the environment, and they are 100 years old. So we expect to find on the genomic level something different between them and 60-year-olds."

To test their hypothesis, they came up with a creative systems approach to help untangle the interaction between longevity genes and their counterparts. The results appeared in *PLoS Computational Biology* in August 2007. If, they said, a gene encourages longevity, it ought to be more common in older populations.

If, on the other hand, a gene makes seniors prone to disease, the individuals with the ill-fated gene will die off, and its frequency in the population should dwindle with age. But if an age-related disease gene first falls and then climbs back up in older populations—showing a U-shaped frequency curve—then a longevity gene in the population could be buffering the effects of the age-related disease gene. The longevity gene protects its holders from the disease gene's harmful effects, allowing the disease gene to remain well-represented in the long-lived crowd.

Barzilai and his coworkers gathered 1.6 million points of data on genes related to cardiovascular disease in 1200 centenarians, their children, and unrelated controls. Then the researchers looked for patterns in gene frequency. The U-shaped frequency curve turned up for a gene called LPA, which codes for lipoprotein(a) and is associated with an increased risk of vascular disease in the elderly. The frequency of the LPA gene decreased until age 80-85, but subsequently increased until, by age 100, its prevalence was similar to that found in 60-year-olds.

Examining frequency trends in subpopulations, the researchers carefully looked for specific buffering of LPA by longevity genes. Their hypothesis was correct. They found that those with the CETP genotype for longevity can also have the detrimental LPA genotype but it will not shorten their lives.

Barzilai writes, "it is most likely that longevity involves a far more complex relationship among longevity and disease genes than the pairwise interactions we have introduced here." But his approach is a start toward unraveling the complicated connections between human aging and longevity genes, and a "proof of concept" for applying high-throughput methods to identify gene-gene interactions such as buffering.

High-Dimensional Biomarker

The growth of data-rich gene and protein profiling tools has instilled some new hope and creativity in the search for biomarkers of aging. A substantial drag on the field of aging is that experiments take so long. A typical experiment involves altering the genes, diet or lifestyle of an animal and observing it till death. "For mammals, that's going to be years," Kim says. "For humans, it's decades!"

"What we need," he says, "is some way to find out right away whether we've made an animal younger or older." In other words, the field needs a biomarker of aging.

The fact that aging researchers don't yet have biomarkers isn't for lack of trying. A ten-year effort spearheaded by the National Institutes of Health in 1988 to find biomarkers of aging in the mouse failed. Many potential measuring sticks have been put forth for both animals and humans, including hormones, body temperature, telomere lengths, and immune cell concentration in the blood. Yet these single candidates generally display too much variation in a healthy, non-aging population, let alone an aging one.

Kim believes he is now onto something with his recent work on the transcriptional profile: a snapshot of the genes being expressed in a group of cells at any one time, usually obtained with microarray technology.

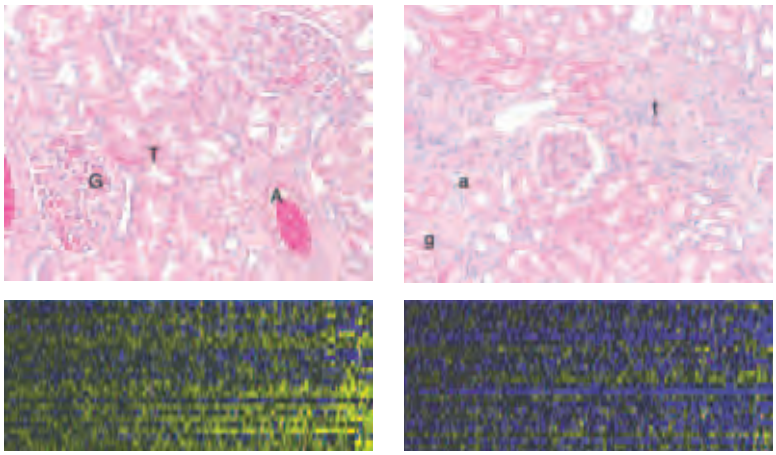
Since a transcriptional profile is a "fairly complex readout from hundreds and hundreds of different markers," Kim says, it provides far more information than a single biomarker could. "It's by integrating all of those different markers together that we have greater power to predict physiological age," he says.

Kim started by looking at old and young human kidneys. Like other filters, the key components of a kidney start to clog as we age: the main filtration unit, the arteries, and the excretion tubules. But some kidneys age much faster than others. Indeed, a physician can simply look at a kidney tissue sample and tell its physiological age—a measure that can be out of sync with chronological age and serves as a more accurate measure of kidney health.

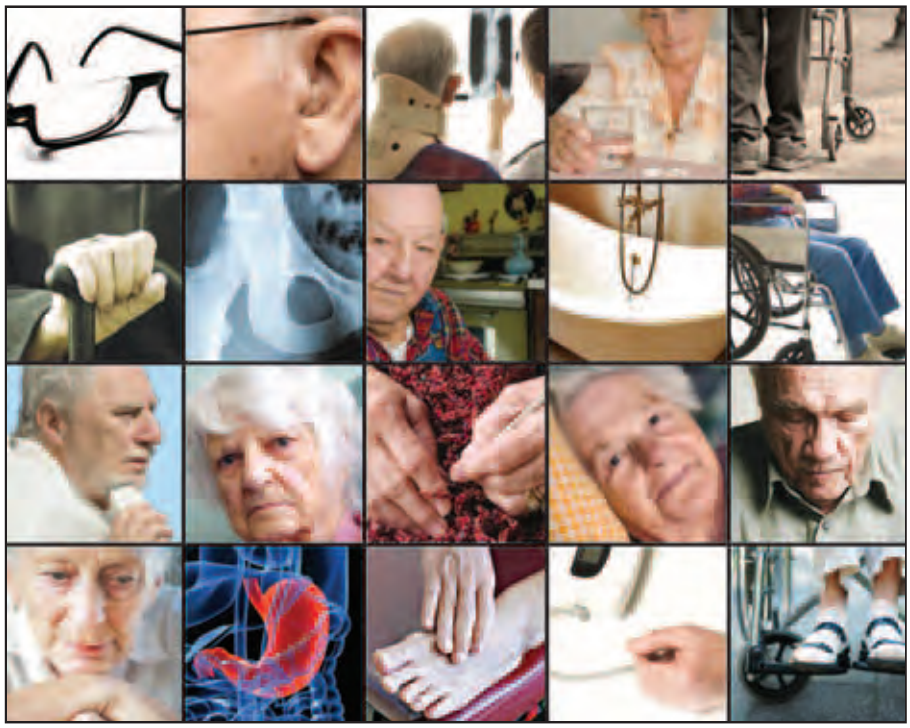
The Stanford group gathered transcriptional profiles from 74 kidneys and found, in a study reported in December 2004 in *PLoS Biology*, that 447 genes (out of 33,000 tested) showed a difference in regulation with chronological age. And when Kim placed the doctor's measures of physiological age against the age indicated by the transcriptional profiles, they correlated. One 78-year-old woman's kidney that looked like it belonged to a 100-year-old displayed a very old transcriptional profile. And an 81-year-old man's kidney that looked "terrific" showed a much younger transcriptional profile.

Kim repeated the transcriptional profile study in human skeletal muscle. Again, the results (published in *PLOS Genetics* in July 2006) showed that the transcriptional profiles reflected real physiological age (muscle deterioration), "not just the birthdate."

Identifying an aging biomarker for a specific tissue like kidney or muscle is helpful. Even more helpful would be a biomarker for a whole organism, especially one that crosses species. In search of a more universal biomarker, Kim also performed a meta-analysis of aging transcriptional profiles in four species. He set his human kidney and muscle data against fly, worm, and mouse data from his and other laboratories.



Kidney age can be tracked by tissue inspection or gene expression profiles. A 29-year-old kidney (top left) with a normal filtering unit (glomerulus, G), tubules (T), and arterioles (A) shows a young gene-expression profile (bottom left). An 84-year-old kidney (top right) with a hardened glomerulus (g), narrow tubules (t), and diseased arteries (a) displays an old gene expression profile (bottom right). The genes shown drop in expression with age. Yellow indicates high expression; blue indicates low expression. Courtesy of Stuart Kim.



To measure aging in individuals, Mitnitski and Rockwood created a frailty index based on the number of aging indicators possessed by each individual. A sample collection of 20 indicators, also called deficits, was published in **BMC Geriatrics** *in February 2002. It includes vision loss, hearing loss, impaired mobility, vascular problem, gait abnormality, impaired vibration sense, difficulty toileting, difficulty cooking, difficulty bathing, difficulty going out, difficulty grooming, skin problems, resting tremor, changes in sleep, difficulty dressing, urinary complaints, gastro-intestinal problem, diabetes, hypertension, and limb tone abnormality. Each indicator is given equal weight in the index, whether it be diabetes or a skin problem.*

For this analysis, Kim and his coworkers designed a set of statistical methods to help overcome a "pervasive methodological challenge in genomics studies"—that the simultaneous reading of tens of thousands of genes increases the possibility of a random, rare event showing up as statistically significant. They categorized 20,000 genes by pathways into 600 gene sets, and then looked for overall trends within the gene sets. Their analysis identified only one pathway that showed the same trend in all four species: Expression of genes in the electron transport chain decreased two-fold over the lifespan of each animal.

Why the electron transport pathway would be the only shared aging pathway across four species is still unknown, Kim says. "But it looks like an unavoidable consequence of getting old."

Other laboratories have completed similar cross-species comparisons, and still others are gathering data for such a comparison. Results have been mixed so far, but if future studies confirm Kim's results, a profile of gene expression in the electron transport chain may be a candidate for a universal biomarker of aging.

Quantifying Aging

In humans, chronological age is often used as a rough measure of the aging process. A young actuary in the early 1800s, studying mortality data in England, discovered that risk of death in populations rises exponentially after puberty, doubling about every seven years. Benjamin Gompertz' curve, called the Law of Mortality, links chronological age to risk of death, and has proven useful to insurance companies and aging researchers ever since.

Yet, says **Arnold Mitnitski, PhD**, "individuals do not die from old age. They die because they become more vulnerable, and this occurs with age." Mitnitski, an associate professor of medicine at Dalhousie University in Canada, and his colleague **Kenneth Rockwood, MD**, a professor of geriatric medicine and neurology, developed a better measure of aging for individuals than chronological age—a measure that takes individual health vulnerability into account. They call it the frailty index.

When they first attempted to quantify biological age, the researchers had access to a large

"What we need," says Stuart Kim, "is some way to find out right away whether we've made an animal younger or older." In other words, the field needs a biomarker of aging.

"Individuals do not die from old age," says Arnold Mitnitski. "They die because they become more vulnerable, and this occurs with age."

Canadian database tracking the medical status of folks as they aged. Mitnitski asked which signs or symptoms listed in the database are essential to tracking aging. Is it incidence of cancer? Is it declining mobility? The researchers tried to use a classical statistical approach to determine which variables were most significant. But the approach assumed that the variables were independent. And Mitnitski and Rockwood's analysis, to their surprise, revealed that every indication of health from declining eyesight to memory loss was just as important as every other one in tracking aging.

Aging is like other complex systems, the scientists realized, in which all variables are so interrelated, they are too difficult to treat separately.

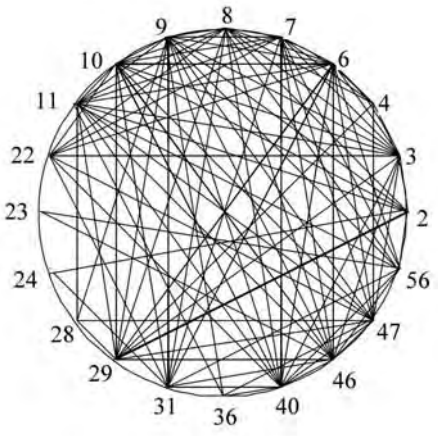
But Mitnitski and Rockwood still wanted to somehow put the variables together into a quantifiable indicator. So they simply counted the characteristics up. They listed every possible sign or symptom of aging, called them deficits, and added them up for each individual.

"Usually nobody does that, just counts," Mitnitski says. But it worked. The number increased with age. And it increased exponentially. It looked, to Mitnitski's surprise, like Gompertz' Law of Mortality.

"It was bizarre that it worked," Mitnitski says. It seemed crazy to give a characteristic like cancer the same weight as a skin condition or a change in sleep patterns. But they found that weighting the deficits was unnecessary. It seems that each of the variables is so highly connected to all the others that it doesn't matter which variables are picked; they all indicate the underlying process of aging.

In order to use the measure across different datasets, the scientists changed their simple sum of deficits into an index measure: a ratio between zero and one. Then they tested the index on datasets from Sweden to Australia. They found it doesn't matter which aging indicators are used—and the more, the better.

Mitnitski and Rockwood's frailty index might have important theoretical implications. Now that Mitnitski and Rockwood have a way of expressing health status as a single number, they are analyzing individual health trajectories over time. Health status tends to go up and down with an overall downward trend, so they have found it can be modeled as a stochastic (random) process. "This makes it possible to apply a powerful mathematical apparatus developed in other fields such as systems theory and engineering to the problems of aging," Mitnitski says. "Our approach opens new horizons in the application of mathematical modeling to aging and health issues."



***The signs and symptoms of aging are highly interconnected. Each node number on the edge of the graph above is one of the deficits listed in the previous image; lines represent statistically significant relationships between deficits. Courtesy of Arnold Mitnitski and* BMC Geriatrics.**

"We start with simple mechanistic models, and then join models together to examine interacting effects," says Carole Proctor.

Aging *in silico*

The complex nature of aging has also been tackled using computational simulation. **Thomas Kirkwood, PhD**, director of the Institute for Ageing and Health at Newcastle University, started making computational models of molecular aging processes more than fifteen years ago.

Kirkwood is motivated in part by a desire to simulate some of the heaps of theories about the mechanisms behind human aging. Some researchers say that hormones are behind aging, others that accumulated damage to DNA leaves the body increasingly vulnerable. Aging could result from growing oxidative damage, an increasing imbalance between launching inflammatory reactions and fighting inflammation, or a shortening of the telomeres (protective caps at the tips of chromosomes). Perhaps aging is a gradual loss of transcriptional control, or a build-up of random damage to the short-lived, low-affinity interactions crucial to operational cellular networks. Perhaps an older body must compromise in its ability to repair itself. Perhaps it is a bit of all of these.

Kirkwood's group has modeled aspects of several of these. For example, they looked at how telomere shortening might be influenced by oxidative stress. They modeled how malfunctioning mitochondria, aberrant proteins, free radicals, and scavengers might interact in the aging cell. And, more recently, they simulated the role of molecular chaperones in aging, and how the ubiquitin-proteasome system impacts protein degradation in aging.

From the start, Kirkwood and his group have created simulations like modular furniture: pieces that can easily be modified, put together, and taken apart again to test the interconnecting pieces underlying aging's complexity. "We start with simple mechanistic models, and then join models together to examine interacting effects," says **Carole Proctor, PhD**, a research associate in the group.

The Newcastle group has also created a model-building platform that allows other researchers, even those not so familiar with modeling, to simulate aging processes. The eventual goal of the Biology of Ageing e-Science Integration and Simulation System (BASIS) project is to put together a simulation of a virtual aging cell,

Less-Fuss Meta-Analysis

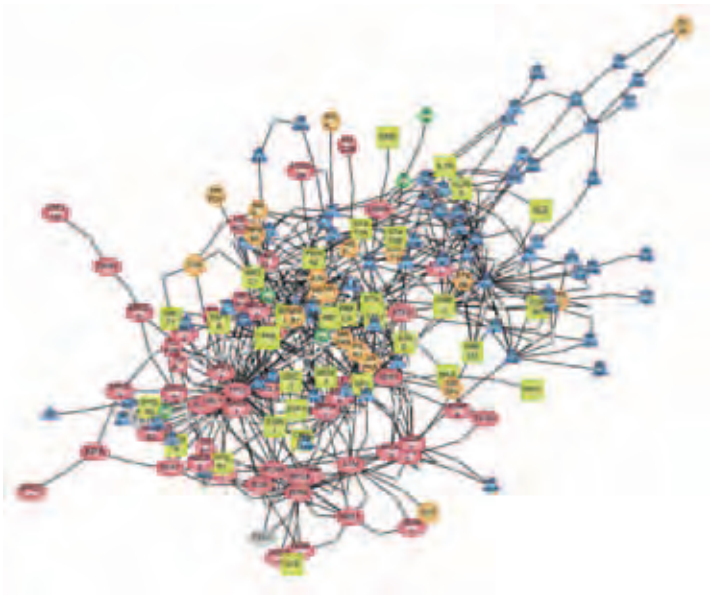
Building the Gene Aging Nexus

Comparing microarray data in the field of aging (and other fields for that matter) can be a nuisance. Different chips or cDNA arrays generate gene expression data in different ways. Even if the array is the same, different parameters set by different laboratories can create widespread variation that is difficult to account for.

Caleb Finch, PhD, professor of gerontology and biological science at the University of Southern California, frustrated when he was unable to even "connect successive papers from the same laboratory," decided to construct a data-mining platform that could assemble databases of primary reads of the chips, convert them into a common format, and offer a variety of statistical tools for cross-examination of the experiments.

He collaborated with a fellow researcher at USC, **Xianghong Jasmine Zhou, PhD**, an assistant professor in molecular and computational biology, and the result is the Gene Aging Nexus (http://gan.usc.edu/public/index.jsp), or GAN. GAN currently holds data from more than 800 microarray experiments in six species (human, rat, mouse, fly, worm, and yeast). And it offers tools to visualize the data with charts and graphs.

When Gene Aging Nexus was introduced in *Nucleic Acids Research* in November 2006, **Chris Patil, PhD**, a postdoctoral fellow studying the genetics of aging with **Judith Campisi, PhD**, at Lawrence Berkeley National Laboratory, gave it a hearty welcome on his aging blog (http://ouroboros.wordpress.com/). "As anyone who's tried to compare experiments performed on different arrays in different cell types (or species) in different labs can attest," Patil wrote, "this publicly available resource will be a huge boon to meta-analysis of gene expression data related to aging."



Although scientists have pinpointed only a handful of human genes that directly affect human aging, one group created a database of more than 200 genes possibly involved in human aging. GenAge is a collection of either human homologs of animal genes known to be related to aging, genes that are part of networks known to be associated with the process of aging, or genes of other interacting partners of age-associated proteins. João Pedro de Magalhães, PhD, a postdoctoral fellow at Harvard University who maintains the database, also helped create a protein interaction network (left) of all the proteins from GenAge known to interact. The colors represent functional groups: red, DNA repair, replication, and condensation; blue, growth and development; yellow, transcriptional regulation and signaling; orange, redox and oxidative regulation and stress response; green, apoptosis; grey, unknown or other. GenAge is available at http://genomics.senescence.info/genes/. Reprinted from "GenAge: a genomic and proteomic network map of human ageing," FEBS Letters 571(1-3) 2004, with permission from Elsevier and co-author João Pedro de Magalhães.

virtual aging tissues (starting with fibroblast connective tissue, intestinal epithelium, and neuronal networks), and virtual aging organisms (starting with *C. elegans*).

Their goal of simulating aging at the cell, tissue and organism level reflects a growing desire within the field of aging to expand from a largely molecular focus to a systems-level exploration.

Aging Systems

Alterations that begin on a molecular level eventually cause the skin to droop, the kidney's filtering to falter, fat cells to redistribute, and bones to become more porous. "Changes percolate upward with age," says **Andres Kriete, PhD**, associate professor for bioinformation engineering at Drexel University.

Kriete believes that quantifying the touches of aging on multiple levels, from the cells to the whole organism, and then integrating that evidence using models and other forms of computational analysis, will be crucial future steps for understanding aging.

"We need a comprehensive point of view," says **Claudio Franceschi, PhD**, professor of immunology at the University of Bologna. "We should continue to do molecular biology studies, because unless you have the molecular biology and genetics studies that identify pathway molecules, then you are lost. But you also need to understand what is going on at more integrated levels. You really need systems and computational biology. Otherwise, you see the details but lose the entire picture."

"You really need systems and computational biology. Otherwise, you see the details but lose the entire picture," says Claudio Franceschi.

Large-Scale Data Mining of Centenarians

For many years, says Franceschi, centenarian researchers studied the genetics of longevity in groups of centenarians about 200 or 300 strong. "But we realized this was not enough," he says. Much larger studies were needed to account for the huge variability among centenarians and gain the statistical power for significant results. Old age is very noisy; many combinations of genetics, lifestyles, cultural factors, and randomness can lead to longevity, Franceschi says.

So Franceschi worked to set up a large-scale project that would eventually include 24 research groups in Europe and one in China working together to sort through the genomes of 5,300 people in 11 European countries. Half of the subjects of the study are siblings over 90 years old. (A few recruited families have five siblings over 90!) GEHA (GEnetics of Healthy Aging) is the largest initiative on the genetics of human longevity ever.

To search for candidate longevity genes, the consortium will perform a linkage analysis of each sibling pair's genomes using 25,000 DNA markers per genome. A linkage analysis calculates which markers are shared more than by chance. "Since they share a trait, which is longevity, we assume that they will share more than 50% of those polymorphisms or markers in their genome that are related exactly to longevity," Franceschi says.

One unique and powerful aspect of the GEHA analysis is that demographers in the consortium will weight the relative significance of each sibling pair's genetic information according to demographic data. For example, the genetic information of two centenarian brothers in Poland will be given slightly more weight than the genetic information of two centenarian sisters in Italy. Men reach 100 less often than women and Poles reach 100 less often than Italians. Because of that, the Polish brothers' genetic data should tell the researchers more about longevity than the Italian sisters' data. "The genetic data will be corrected in a way," Franceschi says, taking into account the odds that each sibling pair had to beat to reach their current ages.

Lack of statistical power is the most common problem in studies of complex genetic traits such as human longevity.

GEHA collaborators will also stratify the data according to health status and cognitive and functional abilities. Some reach 90 in good shape, and some are in very poor shape. "One is associated with healthy aging and the other is associated with aging with some problems," Franceschi says. "We want to account for that."

Franceschi likens GEHA to the large coordinated research efforts of physicists at CERN or the Stanford Linear Accelerator Center. Aging needs such an effort, he says, because lack of statistical power is the most common problem in studies of complex genetic traits such as human longevity. The first results from GEHA should arrive in spring 2008.

On Simulating Growth And Form

BY REGINA NUZZO, PHD

For better or for worse, and on many levels, our tissues never stop growing and changing. While developing from childhood to old age, we grow not only bone, cartilage, fat, muscle and skin, but also toughened arteries, scars for our wounds, and, sometimes, deadly tumors.

As researchers in various fields create computational tools to visualize and simulate growth in all its incarnations, it's clear there's much to be gained. Simulations can teach us how young bodies and faces develop; how an artery compensates for decades of fatty plaque deposits by growing and thickening its walls; how tissue engineers can best coax endothelial cells to develop into organized sheets of skin for burn patients; and how cancerous tumors invade neighboring tissue. Ultimately, computational models of growth may help clinicians and surgeons plan appropriate and patient-specific treatments and interventions for a number of diseases in which growth plays a role.

"Computation is a great tool to study growth," says Ellen Kuhl, PhD, assistant professor of mechanical engineering and bioengineering at Stanford University, "because it lets us understand all those fascinating biological processes we can't otherwise see and predict."

Growing Up: A Change of Face

As we progress from childhood to old age, our faces change. In some ways, these alterations are fairly predictable. As skeletal structures grow and mature, muscle and fatty tissue in our faces increases. Cartilage continues to grow even after facial bone structures have stopped (especially in men), so our noses change shape. And in later years we lose muscle tone and skin elasticity, which can dramatically alter the facial shape.

Though all faces follow this rough trajectory, certain genetic syndromes cause other distinctive patterns. Children with Williams syndrome, for instance, often have particularly full cheeks and lips with a wide jaw. As adults, on the other hand, their faces grow thinner and narrower, while their mouth grows even wider. In fact, experienced clinical geneticists often use their observations of characteristic facial gestalts to make early diagnoses. Yet these observations lack objective, quantifiable evaluation.

Peter Hammond, PhD, professor of computational biology at the University College London (UCL) Institute of Child Health, is working to change that. Hammond and his colleagues are developing computational methods to analyze and visualize variations in three-dimensional face images. The hope, he says, is that pattern recognition tools will support clinicians in their diagnoses of these rare syndromes.

In work published in the *American Journal of Human Genetics* in 2006, Hammond's team investigated facial patterns of four genetic diseases (Williams, Smith-Magenis, Noonan, and 22q11 deletion syndromes), each of which has a characteristic facial gestalt. First, the researchers captured three-dimensional images of facial surfaces from 696 volunteers (roughly half of whom were affected by one of the genetic syndromes) with remote-sensing scanners that use natural light to capture the facial surface.

Then they processed the images—each containing between 4,000 and 20,000 points in three dimensions—to produce a dense correspondence of tens of thousands of points across all of the faces. Next, they analyzed the variability among all the faces using principal component analysis to find the modes of variation that best characterize all the differences. It turns out that about 100 modes could capture 99% of the faces' total shape variation, Hammond says.

This reduced the complexity of the dataset: instead of requiring tens of thousands of points in three dimensions, the essential characteristics of a face can instead be described by a simple linear combination of vectors. The beauty of this, Hammond says, is that each face can now be represented by a single point lying in a high-dimensional "face-space." Important differences between two faces can now be captured by a simple quantitative measure: the distance between two points in face-space.

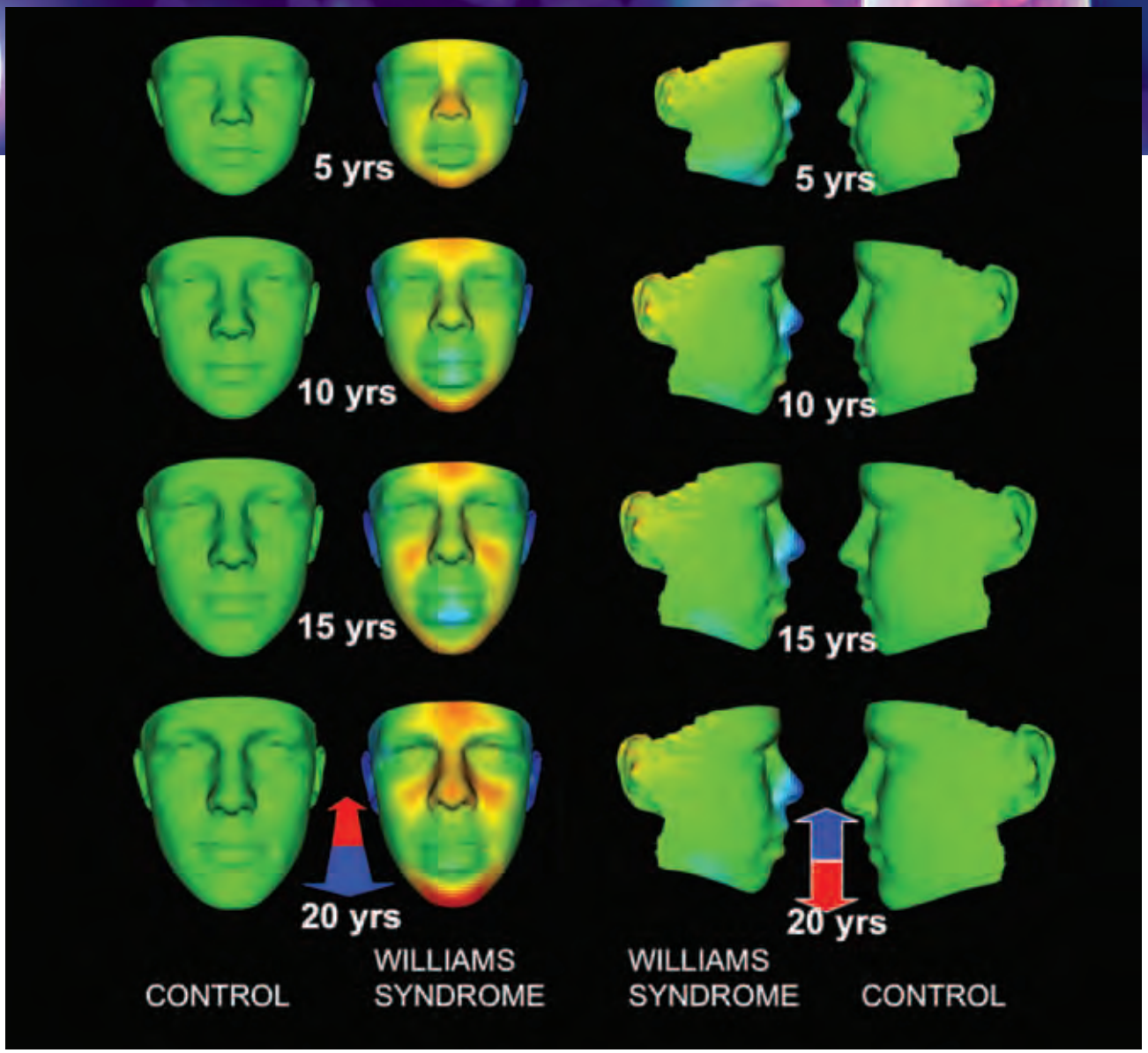
This simplification allows for more detailed facial analysis. In work published in 2003 in *IEEE Transactions on Medical Imaging*, Hammond's team found that changes in a face over time can be nicely expressed as a trajectory through face-space. The methods could also capture subtle features in the faces, even those relating to gender.

Furthermore, in the 2006 study, the researchers found that a single face-space dimension could essentially capture most of the age-related facial growth. So by taking an "average" face and morphing it along this dimension, Hammond says, they were able to construct typical facial growth sequences—one for each of the syndromes as well as the controls.

Hammond's team also used several different pattern recognition algorithms to see if they could discriminate between dense surface models for syndrome and control groups. Each model was developed from a training dataset and then used to classify new test faces. In each case, the algorithms were able to achieve at least a 76 percent success rate—and reached as high as 100 percent for recognizing adults with and without Williams syndrome.

It also turns out that for Williams syndrome, just looking at the areas around the nose and eyes distinguishes affected children from unaffected ones nearly as well as analyzing the entire face. Over time, however, the mouth becomes the most distinctive characteristic: In adults the pattern-recognition algorithms could distinguish affected and unaffected adults very well by focusing only on the areas around the mouth. These discoveries should help researchers develop streamlined clinical applications.

Currently, Hammond and his team are expanding their work to 30 different genetic conditions. But the ultimate goal, Hammond says, is to make these visualization and pattern recognition tools available to clinicians. Before that can happen, however, more data needs to be collected to ensure that the models are trained on a broad range of faces, Hammond says—including both children and adults, affected and unaffected, male and female, and from various races.



Here, simulations of growing faces are colored to emphasize the lack of forward growth (the red areas in cheeks, chin and forehead in the second column) and reduced vertical growth (the blue areas such as the shortened nose in the third column) in the average face of people with Williams syndrome. Courtesy of Peter Hammond.

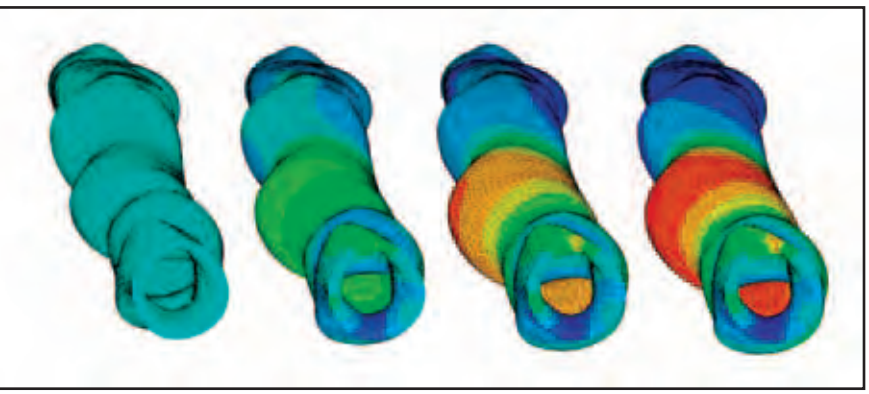
Growing Tough: Arteries Harden

As we age, our blood vessels slowly grow and change in response to the stresses of life. Ultimately, this can lead to hardening of the arteries—atherosclerosis—a disease in which arterial walls grow and change in response to the accumulation of waxy plaque. Eventually, drastic alterations in the arteries not only reduce blood flow but may also lead to aneurysms or blood clots from sudden plaque ruptures.

Treatments for atherosclerosis, while potentially life-saving, sometimes also backfire and induce fast, dangerous growth. Surgically-implanted stents widen the diameter of arteries and increase blood flow, but they can also trigger in-stent restenosis in which the artery narrows once again, sometimes in the span of only a few weeks.

Ellen Kuhl has been looking at how this atherosclerotic growth occurs. She uses computational models to simulate artery walls that respond to changes in mechanical loading from both plaque build-up and stent insertion.

"The golden rule is that the wall of the artery tries to keep the stress at a physiological base level," Kuhl says. "If you increase the base load on the wall [as happens with plaque build-up or stent insertion], the wall will thicken. If you decrease the load, the wall will thin."



A simulated aorta, based on computer tomography data from a specific patient, highlights the effects of stent surgery (shown before surgery as well as 4 days, 16 days, and 33 days after surgery). Cyan shading shows artery walls with normal thickness; blue shading shows thinning walls; red shading show thickened walls. In a real patient, the tissue growth shown in red would have resulted in in-stent restenosis. Courtesy of Ellen Kuhl. Reprinted from Computational modelling of artherosclerosis - A first approach towards a patient specific simulation based on computer topography, by Kuhl et al.,* Biomech. Model. Mechanobiol. 2007;6:321-331*, with kind permission of Springer Science and Business Media.

To study this effect, Kuhl and **Ramona Maas**, then a master's student under Kuhl at the University of Kaiserslautern, first modeled three stages of atherosclerosis in an idealized artery: initial plaque build-up, adiposis (in which plaque is still soft and fatty), and calcification (in which plaque becomes hard and brittle).

In the earliest stage, dramatic tissue growth developed around soft plaque in the simulated artery. This caused a general thickening of artery walls. As the plaque calcified, however, the wall stresses became focused at the boundaries of the plaque, producing a different growth pattern: increased tissue in very small areas—spots that would be prime targets for sudden plaque ruptures. These results were published in 2006 in *Biomechanics and Modeling in Mechanobiology*.

Kuhl and Maas also applied their methods to the more complicated geometry of an actual aorta undergoing stent insertion. They obtained computer tomography images that captured cross-sections of a human abdomen—from beneath the heart down to the legs—in slices 10 millimeters apart. By applying the finite element method to these data they then generated a solid model of the patient's aorta. A uniform pressure inserted at a certain spot in the artery simulated the patient's stent surgery. The team then followed the simulated aorta as it healed. Each time step in their simulation corresponded to 30 minutes (approximately the amount of time required for a stent-implant surgery).

Within 200 time steps—about four days after surgery—stress-induced growth had started to appear in the tissue. The artery walls thickened dramatically, especially in a few extra-vulnerable spots. After one simulated month, the aorta walls had stopped thickening. But in a real patient, restenosis damage would have already been done. Due to forces in the body acting on the outside of the artery (forces not present in the simulation), the walls' growth would have been forced inward, once again narrowing the artery and reducing blood flow.

Eventually, Kuhl hopes simulations like these will help in the design of better stents and will also allow clinicians to simulate the effect of various stent locations and materials on a specific patient's anatomy before doing the surgery.

"Advances in this field have largely been driven by trial and error and have not yet been driven by patient-specific modeling," Kuhl says. "Now we might have the means to eventually say what works best for each patient—and to say why it works."

Next, Kuhl is working to apply these methods to a multi-scale simulation of the heart. Working with **Oscar Abilez, MD**, a postdoctoral fellow in the Department of Surgery at Stanford University and member of the Cardiovascular Tissue Engineering Group, she is exploring the use of computational tools to model how heart tissue changes after a heart attack, and how an experimental treatment that would inject stem cells into the heart might quickly restore normal form and functioning of the heart after the attack.

"Advances in this field have largely been driven by trial and error and have not yet been driven by patient-specific modeling," Kuhl says. "Now we might have the means to eventually say what works best for each patient—and to say why it works."

Growing Together: Social Cells

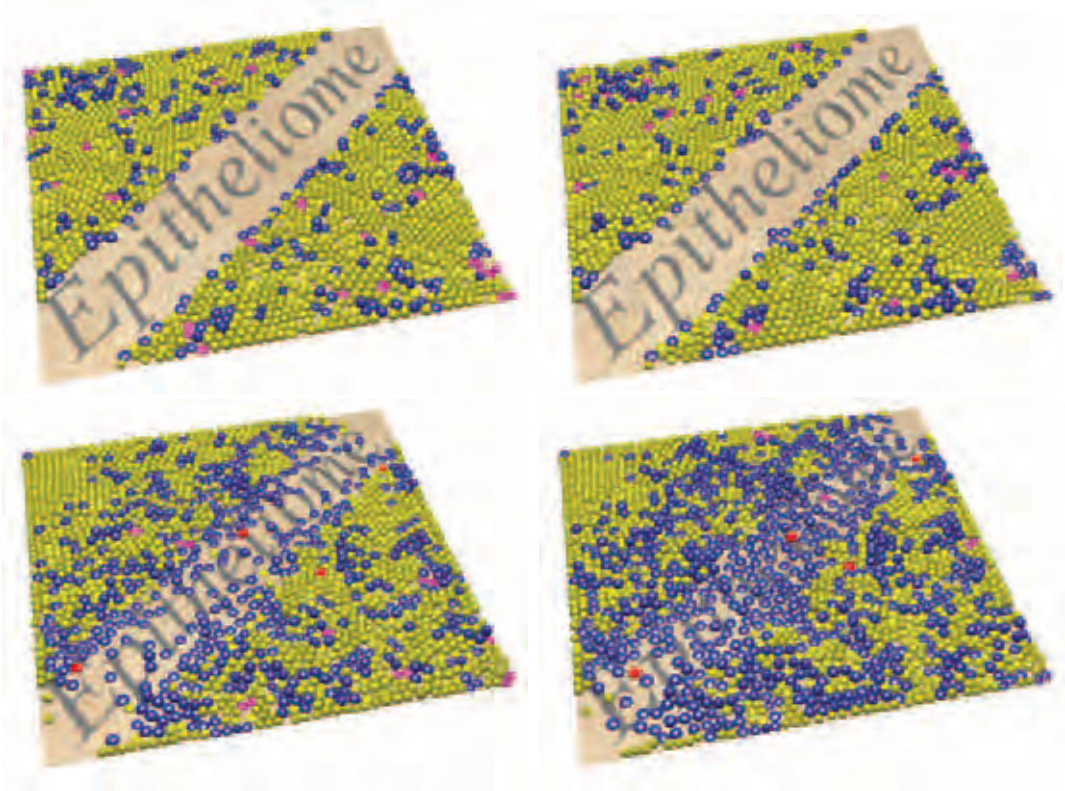
Other computational researchers are focusing on the cellular components of growth—looking at what happens, for example, when growth is desirable (as in wound healing) or undesirable (as in cancer).

In a sense, cells self-organize into tissue in much the way that individuals form a society. The analogy is more than a cute anthropomorphism, however; it also makes surprising sense from a systems-level perspective. Cells form communities and exchange information with their neighbors. They multiply, and they will continue to do so until their neighbors send signals that discourage excessive behavior. Once rebuked, cells will lie quiet until changes in the neighborhood remove these inhibitions and allow growth once again.

Rod Smallwood, PhD, a professor of computational systems biology in the Department of Computer Science at University of Sheffield, refers to this paradigm as "the social life of cells." To study this self-organization, he is developing computational simulations in what he has dubbed the Epitheliome Project. "I wanted to explore how the machinery of individual cells could produce tissue at the next hierarchical level without a blueprint," he says. "That is, how can the interactions of cells produce something greater than individual cells?"

To answer these questions, Smallwood turned to agent-based modeling, a method in which every cell is represented by an individual chunk of software with logical rules that abstract out the details of cells' biochemical lives. In his simulations, these cellular agents each possess individual memory and a physical location in the simulated tissue as well as other physical properties. They can communicate with each other and with their environment—for instance by sensing hormones in extracellular space—and most importantly, make behavioral decisions based on a set of rules.

"This modeling paradigm is so general that we've also



Simulated bladder cells show the healing of a scratch wound in low-calcium (above) and physiologically-normal (below) calcium environments after 2.5 hours [left] and 4 hours [right]. In low calcium, individual cells migrate quickly into the wound, while pressure from dividing cells also slowly pushes more cells behind them. In physiologically-normal calcium, cells don't migrate, and the wound takes twice as long to heal. Courtesy of Dawn Walker.

used it for modeling individual proteins, ant colonies, and the European financial market," Smallwood says. "But the level I'm particularly interested in is that of cells."

Smallwood and colleagues, including **Dawn Walker, PhD**, a post-doctoral academic fellow in the Department of Computer Science, have focused specifically on epithelial cells. Not only are these cells relatively simple and backed up by good in vitro models, Smallwood says, but they also are associated with important clinical applications, such as wound healing, skin grafts and tissue engineering.

"I wanted to explore how the machinery of individual cells could produce tissue at the next hierarchical level without a blueprint," says Ron Smallwood.

In one study, published in 2004 in *IEEE Transactions on Nanobioscience*, Smallwood and Walker explored how bladder cell monolayers heal differently in different environments. Their simulation found that wounds in low-calcium environments healed about twice as fast as those surrounded by physiologically-normal levels of calcium. What's more, the healing mechanism was different for the two cases, Walker says. Under the right conditions in the low-calcium environment, cells at the wound's edge first quickly migrated into the bare area, while a second united front of cells then slowly inched forward into the wound, pushed by physical forces created by cells spreading and proliferating behind them. But in the physiologically-normal environment, individual cells did not usually migrate into the wound, Walker says. Healing occurred only as a function of the much slower united-front process. Experiments with wound healing *in vitro* confirmed these simulation results.

Walker has been working on more detailed simulations to better understand growth and healing in epithelial tissue. In work presented at the Foundations of Systems Biology in Engineering conference in Stuttgart, Germany in September of 2007, Walker and colleagues investigated how different patterns of direct cell-to-cell contact can change whether a cell is likely to undergo growth. This required a set of complicated behavior rules that incorporated mathematical modeling paradigms to capture molecular mechanisms involved in cell-to-cell contact. The results are difficult to verify experimentally, but they suggest that cell-to-cell signaling through epidermal growth factor receptors could explain how calcium can affect cell growth, resulting in the kinds of wound-healing patterns observed previously.

For their simulations, Smallwood's group developed their own freely-available modeling framework, called Flame (www.flame.ac.uk). Their simulations, which typically include about 50,000 cells on a single processor, are run at HPCx, the largest academic computer center in the United Kingdom. Smallwood's group is now working to extend the framework for use on supercomputers as well.

Growing Out of Control: Cancer

Cancer's hallmark is uncontrolled growth, but particular kinds of tumors have especially complex growth patterns. For example, some malignant brain tumors appear to grow by harboring two single-minded focuses: Their tumor cells may invade nearby tissue, or they may divide, but they cannot do both at the same time. The growth process for tumors like these remains largely hidden, in part because imaging still can't capture the advance of single cells.

Tumor simulations might help researchers better understand the mechanisms behind these growth patterns as well as develop new hypotheses for further testing, says **Thomas S. Deisboeck, MD**, assistant professor of radiology at Massachusetts General Hospital and Harvard Medical School and principal investigator of the Center for the Development of a Virtual Tumor (CViT).

Deisboeck and his colleagues **Le Zhang, PhD**, and **Zhihui Wang, PhD**, both postdoctoral fellows at Massachusetts General Hospital and Harvard Medical School, simulate cancer growth using agent-based modeling with a hybrid approach of both continuous and discrete techniques.

Most importantly, Deisboeck says, their model spans scales from the molecular (using proteomic data, for example) to the micro- and macroscopic levels (using imaging data, for example). "The insight that we're after is how perturbations move throughout and across the scales," he says.

In Deisboeck's models, each tumor cell senses cues from its neighbors and from its microenvironment, processed through a gene-protein interaction network with epidermal growth factor receptors (EGFR). Each cell uses this signaling information to choose its behavior—divide, invade nearby tissue, stay quiet, or die—at every point in time in the simulation.

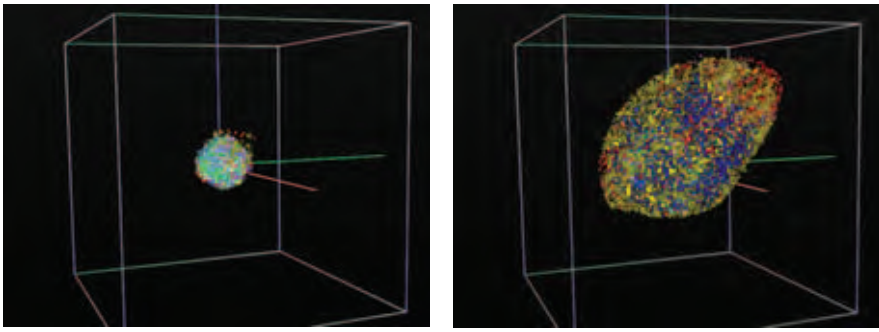
In work published in 2007 in the *Journal of Theoretical Biology*, simulations by Deisboeck showed that the number of cells dedicated to proliferation and migration do not increase steadily. Rather, they tend to oscillate over time. These patterns in turn affect how quickly the entire tumor spreads through the virtual brain.

The group has also extended this model to include brain tumor cells with different characteristics, including a range of EGFR densities. The simulation results, discussed in a paper posted on arXiv.org in 2007 [http://arxiv.org/abs/q-bio/0612037], show that heterogeneity in the tumor on a microscopic level can indeed affect the tumor system's growth patterns. The results also suggest that higher EGFR density levels will lead to some groups of cells switching from proliferation to migration behavior even earlier.

Deisboeck and his team are also using multiscale agent-based modeling to investigate another cancer where EGFR-related signaling pathways seem to play an important role: non-small cell lung cancer. Their early results, published in 2007 in *Theoretical Biology and Medical Modelling*, suggest that tumors will spread more aggressively when more epidermal growth factor is available to the them. This work holds potential for use in drug discovery, Deisboeck says.

His group expects to soon move from "biologically-inspired lattices" to "patient-specific lattices" that will incorporate a patient's imaging data on which to train the simulations, Deisboeck says. In time, he hopes that such cross-scale simulations of tumor expansion and related biomarker evaluations will help clinicians treat cancer patients. For instance, Deisboeck says, in the case of brain tumors, the simulations might help physicians plan for the impact of various treatments, or continuously monitor patients' response to treatment, especially by reducing the number of costly brain scans needed.

"The insight that we're after is how perturbations move throughout and across the scales," Deisboeck says.



Three-dimensional snapshots show the brain tumor simulation at 225 hours and 625 hours. Red circles represent the nutrient source for the tumor; green, blue, yellow, purple, and red shaded dots represent different clonal tumor cell types in proliferation and migration state; grey circles represent cells in quiescence and apoptosis states. Courtesy of Thomas Deisboeck.

Growing Challenges: Data, Methods, Platforms and Patients

If computational models of growth are to find a place in the physician's tool kit, Deisboeck says, they will need to handle not only a huge volume of data but a wide spectrum of different data types, incorporating scales from the molecular to the macroscopic. "Multiscale and multiresolution modeling are one promising way to address that," he says.

In addition, according to Kuhl, models that incorporate a hybrid of approaches will likely fare the best. "Traditional continuum models assume everything is homogeneous and continuous," she says. "But when you look at it at the cell level, you see that's not quite true. We are now able to build these models much better."

Common software platforms will be another challenge for the growth modeling community, Smallwood says. He believes when researchers develop new modeling software, they have an obligation to make it freely available. "Otherwise no one else can replicate your results," he says. "And if other researchers can't replicate them, it isn't true science."

Ultimately, the holy grail of growth modeling will be applying simulations to individual patients. "Simulations of growth are finally becoming powerful enough that we can start to do patient-specific modeling," Kuhl says. "That's where the real pay-offs are, because we can virtually play with possible medical scenarios and find the best ways to treat a patient's personal disease."

BY JOY KU, PHD

An insider's view of biological structures

In March, Simbios released version 1.0 of the SimTK Simulation toolkit. A cornerstone of this release is Simbody, a new piece of the open-source SimTK Core toolkit for physics-based simulation. Simbody solves Newton's equation of motion, F = ma, with a twist: It tackles the equation from the inside out rather than from the outside in.

Simbody provides a new option, says **Michael Sherman**, the toolset's chief software architect. "With Simbody, researchers can do dynamics in whatever coordinates make sense for what they're doing."

Traditionally, the motion equation is solved in a global (Cartesian) coordinate space, from the perspective of someone watching a system from the outside. But Simbody solves the equation in internal coordinate space: The components of a mechanical system are described relative to each other, rather than to a global outside point.

Simbody's Internal Coordinates: Natural and Efficient

This internal coordinate system is a more natural way to model biological structures where highly interconnected parts are the rule. The motion of a hand, for example, is best described relative to the arm's movement. Similarly, the carbon atoms of a benzene molecule are tied to one another in a ring structure.

The method is also more efficient. Because internal coordinates can often be represented with fewer equations, the approach greatly speeds up an analysis. This is especially appealing for simulating large, complex structures over long time scales. **Nagarajan Vaidehi, PhD**, a professor in the division of immunology at the City of Hope Medical Center, says that's why she thinks this work is so important to explore. Vaidehi believes there is great promise for the use of internal coordinates in molecular dynamics simulations, particularly for large proteins. Current approaches take too long to produce the several microseconds of results needed. Internal coordinates could increase the timescales that could be modeled, she says.

Simbody's Code: Elegant

Other software programs that perform dynamic calculations don't measure up to Simbody in one way or another, Sherman says. He designed Simbody to be a best-of-class multibody system. It incorporates code by **Charles Schwieters, PhD**, of the National Institutes of Health for molecular structure determination. This code is based on multibody dynamics theory that **Abhinandan Jain, PhD** and **Guillermo Rodriguez, PhD** at the Jet Propulsion Laboratory developed and used over many years for spacecraft simulation.

"The [Simbody] code is beautiful," says **Paul Mitiguy, PhD**, a consulting associate professor in mechanical engineering at Stanford University and a developer of mechanical simulation software packages. Computationally, Simbody is extremely efficient; it uses templates, takes maximum advantage of the compiler and minimizes memory movement.

Though Simbody has all the hallmarks of thoughtfully designed commercial code, it's not commercial. It's open-source and can be freely downloaded. It's also extensible; researchers can modify it to suit their needs.

Mitiguy describes Simbody as "the Matlab of simulation" because its flexible interface enables it to be the platform upon which many other useful applications are built. "If you really want to get a bang for your buck and computation is an important part of your process, then I think you'd be hard-pressed to do better than Simbody," he says. □

DETAILS

How Do You Get SimTK 1.0 and Simbody?

Simbody is part of the SimTK Core toolkit, an open-source C++ application programming interface (API) to computational tools and algorithms for biological simulations. It has already been incorporated into applications such as OpenSim, an open-source application for modeling and simulating the neuromusculoskeletal system. Release 1.0 of the toolkit, including Simbody, can be freely accessed at http://simtk.org/home/simtkcore by going to the Download section. Source code for Simbody can be found at http://simtk.org/home/simbody. A hands-on workshop to teach researchers how to use SimTK 1.0 was held on March 20 & 21 (http://simbios.stanford.edu/workshop.htm).

Under TheHood

BY NATHAN HAGEMAN

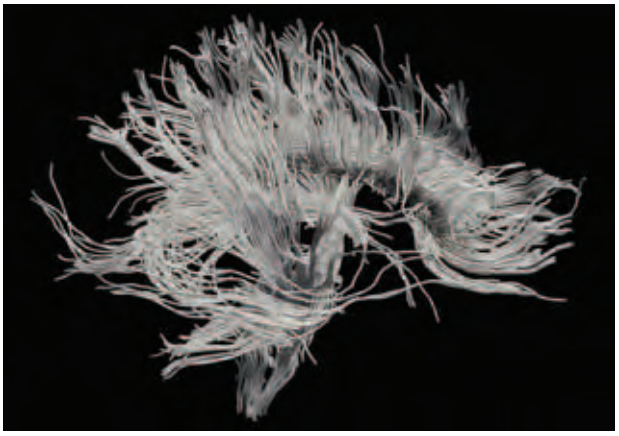
Diffusion Tensor Imaging Tractography: Revealing Connectivity in the Living Brain

One of the major obstacles to studying the human brain has always been gaining access. Until relatively recently, almost all of what we knew about the brain was obtained through post-mortem examination. Few techniques were available for obtaining a clear picture of the structure and function of the living human brain at work. Recent imaging methods, such as MRI, fMRI, and PET, have been partially successful at overcoming this obstacle.

However, while structural methods, such as MRI, allow us to differentiate tissue types, and functional methods, such as PET and fMRI, allow us to characterize regional brain blood flow and metabolism, we still had no clear way of looking at how different parts of the brain connect to each other via the white matter in vivo. The development of diffusion tensor MRI changed that.

Diffusion tensor imaging (DTI) is a non-invasive magnetic resonance imaging based technique used to measure the in vivo local diffusion of water within tissues. This is useful because water molecules move unevenly within fibrous tissues. In the white matter of the brain, the movement of water molecules along axons is much freer than perpendicular to them. DTI, therefore, gives us the local diffusion profile of the brain, which is linked to the local axonal integrity and orientation. Computerized tractography methods have been created to determine likely connection paths through the white matter between different parts of the brain based on this local orientation information. These methods use varied and sometimes sophisticated mathematical techniques to generate regional connectivity maps.

The earliest DTI tractography algorithms were streamline methods that generated discrete paths through the white matter by following the local major diffusion direction at each voxel, point by point, through the image volume. There are several disadvantages to this approach. In particular, random noise in the image causes perturbations in the diffusion tensor, which can have significant effects on the reliability of these streamline methods. More sophisticated techniques have been created to try to correct for the weaknesses in the original streamline model. Others have attempted to reconstruct the connectivity information using entirely different mathematical approaches. For example, some techniques use partial differential equation based models to generate a connectivity map by simulating the movement of some physical substance based on the DTI data. One such method involves simulating a fluid flow through the DTI image volume guided by a pressure force at each voxel identical to its diffusion profile [1]. The fluid dynamics then are directly related to the underlying local axonal orientation and can provide a metric of regional connectivity.



Segmentation of major white matter tracts in the brain using a fluid mechanics-based DTI tractography method. Courtesy of Nathan Hageman.

Probably the greatest limitation of tractography methods currently is due to the resolutions at which DTI data can be acquired. Current voxel resolutions for DTI are about 1-3 millimeters. However, the typical axonal diameter is 1000 times smaller—approximately 1 micron; this disparity can lead to averaging of different fiber populations within the voxel. The fiber architecture within a single voxel can be extremely complex and not well modeled by a single fiber direction. Therefore, even the best tractography methods provide only probable connection paths, and we can not assume that they correctly reflect the true underlying fiber architecture. However, DTI tractography remains a powerful and promising technique. Perhaps one day it will help us reveal the true fiber architecture of the living brain. □

DETAILS

Nathan Hageman is a graduate student in the Laboratory of Neuroimaging (LONI) at the University of California, Los Angeles School of Medicine (Director: Arthur Toga, PhD). His work focuses on developing analysis methods for diffusion MRI and their application in understanding neuropathology.

REFERENCES:

N.S. Hageman, D.W. Shattuck, K. Narr, and A.W. Toga. A Diffusion Tensor Imaging Tractography Method Based on Navier-Stokes Fluid Mechanics. In *Proceedings of the 2006 IEEE International Symposium on Biomedical Imaging*, 2006.

DIVERSE DISCIPLINES, ONE COMMUNITY
Biomedical Computation REVIEW
Published by Simbios, a National Center for Biomedical Computing

Nonprofit Org.
U.S. Postage Paid
Permit No. 28
Palo Alto, CA

Biomedical Computation Review
Simbios A National Center for Biomedical Computing
Stanford University
318 Campus Drive
Clark Center Room S231
Stanford, CA 94305-5444

seeing science

SeeingScience

BY KATHARINE MILLER

The Envelope Please...

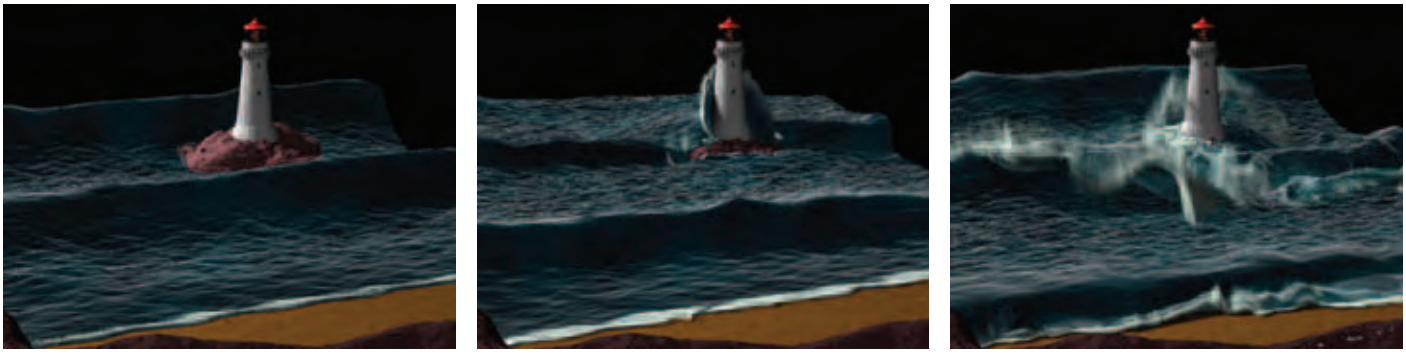
They may not own tuxedos or Dolce & Gabbana gowns, but computational scientists can nevertheless win Oscars.

In February, **Ron Fedkiw, PhD**, associate professor of computer science at Stanford University, together with two scientists at Industrial Light and Magic, **Frank Losasso Petterson, PhD** (a former doctoral student of Fedkiw's), and **Nick Rasmussen**, received a Science and Engineering Award from the Academy of Motion Picture Arts and Sciences.

Their Oscars recognized the group's fluid simulation system which produced impressive and realistic fluid effects in multiple movies including *Terminator 3*, the *Pirates of the Caribbean* movies, *Star Wars: Revenge of the Sith*, and *Evan Almighty*.

The Oscar-winning program simulates water using the same equations others rely on, but, says Fedkiw, the algorithms do a particularly good job of accurately and robustly calculating the interface between water and air. In other words, the spray and bubbles in splashing waves look realistic, while smooth water reflects light in appropriate ways.

The Oscar work also has potential biomedical applications. As Fedkiw puts it, anywhere there's fluid—such as blood—the simulation of moving fluids can be useful. "Most recently we've developed some new solids-fluid coupling algorithms that allow rigid, deformable and even thin shells to be simulated in a two way coupled fashion with fluids," he says. "This could be of use, for example, in simulating heart valves." □



***Oscar-winning Water.** A 2008 Academy Award for Science and Engineering went to developers of a fluid simulation system used in a number of movies produced by Industrial Light & Magic. As shown here, the Oscar-winning program generates realistic-looking waves and splashes using the particle level set method and an additional method that simulates how the spray interacts with itself and the surrounding water. Courtesy of Ron Fedkiw*