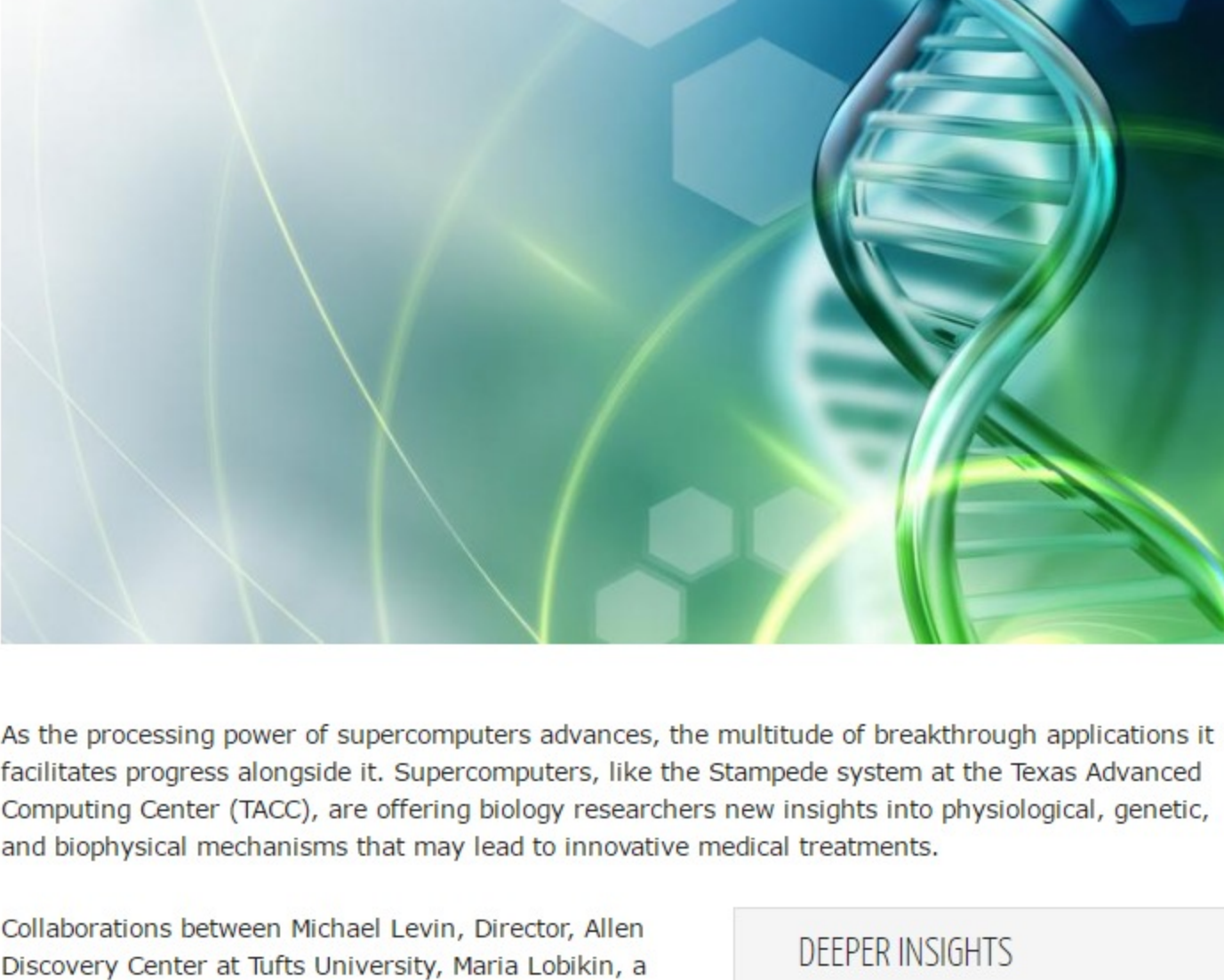


Supercomputers

Supercomputers Use Machine Learning to Gain New Insights into Complex Cellular Processes

Wed, 06/21/2017 - 4:22pm by Rob Johnson



As the processing power of supercomputers advances, the multitude of breakthrough applications it facilitates progress alongside it. Supercomputers, like the Stampede system at the Texas Advanced Computing Center (TACC), are offering biology researchers new insights into physiological, genetic, and biophysical mechanisms that may lead to innovative medical treatments.

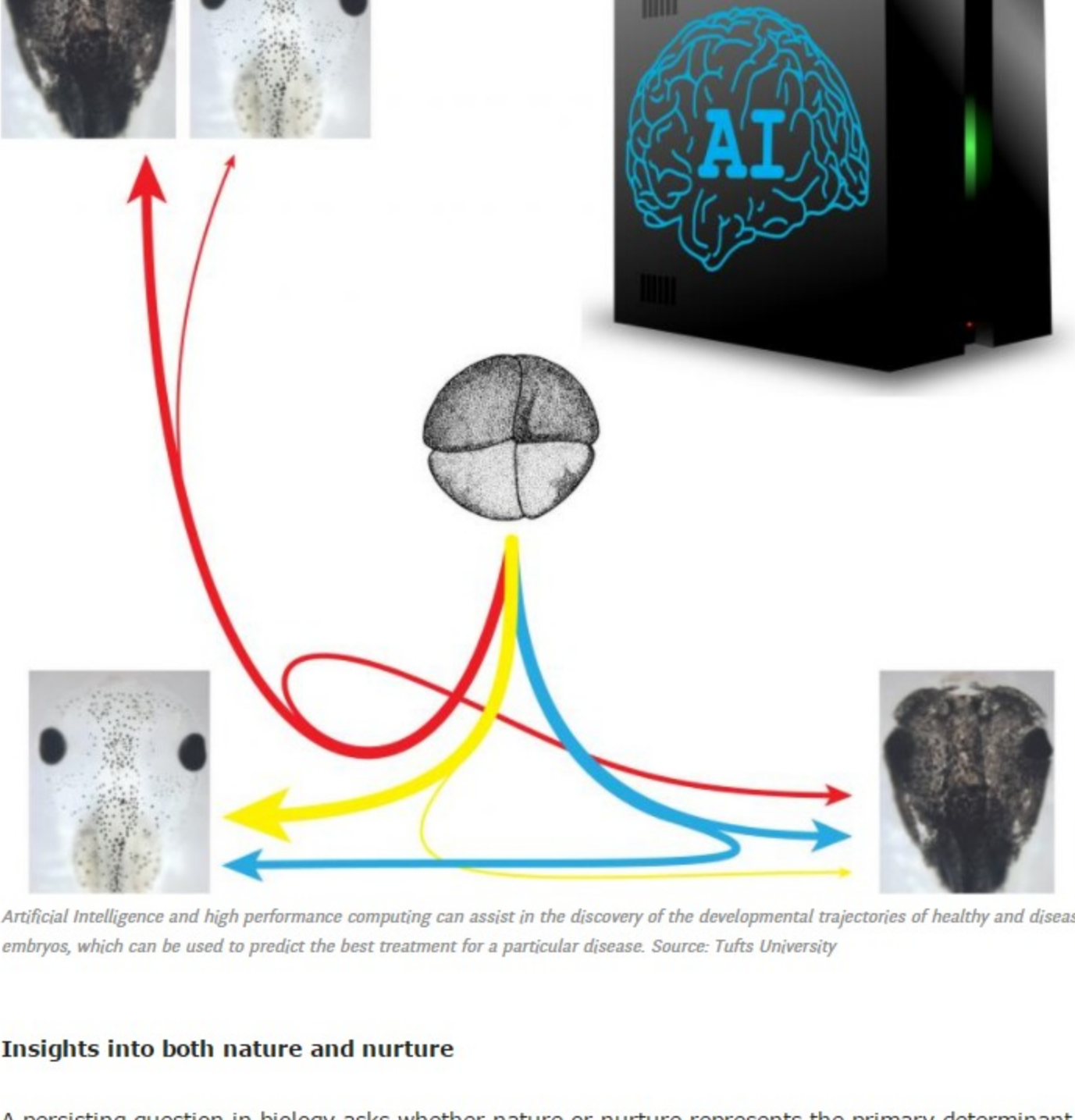
Collaborations between Michael Levin, Director, Allen Discovery Center at Tufts University, Maria Lobikin, a post-doctoral researcher in the Levin lab, and Daniel Lobo, Assistant Professor at the University of Maryland, Baltimore County, and their teams are working together using the power of Stampede to perform the detailed calculations necessary to unlock nature's secrets. Their pioneering work seeks to understand better the workings of signaling systems in living tissues, and how cellular activity is coordinated for the control of anatomical structure and function.

With greater insight into this process, Levin and his team can ascertain how a flatworm, for example, determines where its head or tail will form as it regenerates itself after severe injury, or how cancerous tumors form and grow.

In a sense, the cells perform their own biological version of "computations" which determine how they will behave as a small part of a living being's overall anatomy. By understanding how cells' "computations" occur, these scientists can consider ways to utilize that information. With knowledge of the mechanisms dictating how cells cooperate to maintain normal structure and suppress tumorigenesis, the future may hold insights that lead to new cancer therapies, or further research the field of regenerative medicine, applicable to humans.



Rob Johnson
Owner, Fine Tuning LLC, a strategic marketing and communications consulting company in Portland, Oregon.



Artificial Intelligence and high performance computing can assist in the discovery of the developmental trajectories of healthy and diseased embryos, which can be used to predict the best treatment for a particular disease. Source: Tufts University

Insights into both nature and nurture

A persisting question in biology asks whether nature or nurture represents the primary determinant of biological variation among individuals of the same species. "Both," explains Levin as he describes the complex interaction of the two elements in a living being. "If we use an analogy comparing biological processes to a computational system, DNA specifies the hardware, and electrical transmission among cells is akin to the software. Electrical impulses are not only occurring in the brain; they are transmitted among cells throughout the body. DNA determines the expression of electrically-active proteins in cells, but an individual organism's electrical dynamics are unique and based on external factors as well as internal ones. In nature, living cells are programmed to perform computations and make decisions about patterning, healing, and regeneration of the body. If we can tap into those regulatory algorithms using mathematical models, we may find ways in the future to re-program cancer cells, and induce regenerative repair - as we are already doing in some animal model systems. We see our research as an important step toward a better understanding of these processes, hopefully enabling other scientists to expand on our findings toward biomedical applications."

Cracking biology's code

Even complex and dynamic processes taking place in living cells can be described mathematically if the appropriate algorithms and models are applied. However, the deluge of functional and molecular data create an enormous challenge for those dedicated researchers trying to crack this code and propose a model consistent with the ever-increasing body of known facts. To understand the system well enough to be able to infer specific therapeutic interventions, a researcher must compare different models, and see which best maps to the results observed in nature. The concept is straightforward, but as the saying goes, the devil is in the details. How does one sort through billions of possible models to find the proverbial needle in the haystack - the model that provides meaningful insights into a living organism's biological processes? It is a daunting task indeed.

Data from many lab experiments with tadpoles were fed into the Stampede supercomputer, which uses machine learning in the quest to find the best mathematical model. Once the supercomputer reverse-engineers the mechanism modeling the observations from the experiments, Levin, Lobo, and their team can validate the results in living tissues using various drugs and reagents. Eventually, the technique may lead to custom treatments for people. "There is a gap between gathering data, and extracting actionable wisdom from it," notes Levin. To make sense of these intricate processes, we need to perform billions of simulations to identify patterns and form meaningful, bigger-picture insights. Powerful tools like supercomputers capable of machine learning help us extract that type of information. Only then can we make better sense of the data to form useful hypotheses applicable to biology and medicine."

Empowering the Stampede

The complex modeling process is possible thanks to the advanced technologies under the rhetorical hood of TACC's Stampede supercomputer. Supported with funding from the National Science Foundation (NSF), and other organizations, Stampede's infrastructure utilizes the latest hardware and software technology for advanced machine learning, accelerated by Intel® processors. Stampede is comprised of 6,400 Dell C8220 compute nodes that are housed in 160 racks; each node has two Intel E5

8-core (Sandy Bridge) processors and an Intel Xeon Phi 61-core (Knights Corner) coprocessor. According to the November 2016 Top500.org listing, this level of performance earned the supercomputer a number 17 ranking among the fastest systems in the world. With other innovative technologies working in tandem, Stampede delivers a whopping ten quadrillion operations (10 petaflops) per second.

Onsite at the University of Texas at Austin, TACC's team of supercomputing experts also partner with other researchers around the globe. In addition to the work undertaken by Levin and Lobo, Stampede currently supports over 2,000 other users, and 1,000 research projects in fields like chemistry, geophysics, and energy. In total, over one billion CPU hours have performed detailed calculations in support of advanced research. By configuring Stampede for a particular scientist's needs, the TACC team can help researchers perform simulations, calculations, or modeling in virtually any scientific discipline.

Levin added, "With Stampede's performance, we can obtain biological insights not possible without the aid of a supercomputer. While our particular research today focuses on biophysics in tadpoles, machine learning gives us a very advanced tool from which to base our work. Today's research with tadpoles helps us understand anatomical disorders and disease manifestations. It will take a lot of work to go from tadpole experiments to the development of human applications. However, the models we are evaluating uncover many of the mysteries of basic biology, so there's a direct path between models today and practical human uses in the future."

Daniel Lobo reiterates the challenge of their current research, the scope of the computational power required to obtain the needed data, "Our algorithms running in Stampede can identify and isolate a precise mathematical model that explains the observed cellular behaviors in both healthy and diseased living organisms. Without this machine learning capability and supercomputer power, this task of automatically reverse-engineering explanatory models directly from experimental data would be impossible." A challenge of this magnitude requires careful coordination among the nodes of a supercomputer, applying the workload to the nodes most appropriate for the task. Describing his work with Stampede, he elaborates. "A system as powerful as TACC's allows us to set up a master node which iteratively designs the complex mathematical models that are possible explanations for the experimental data at hand. In turn, the master node offloads the computing tasks of evaluating the accuracy of these candidate models to other nodes in the cluster which do the actual simulations. Intel's Xeon processors help those nodes make short work of a daunting task."


Coming on strong: Stampede 2

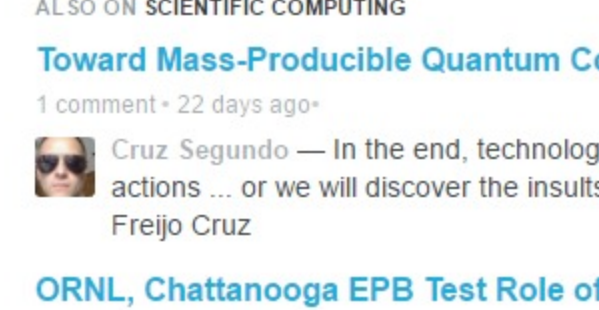
In May, TACC began supporting early users on their next generation Stampede2 system, made possible by a \$30 million grant from NSF. The new system will be fully deployed to the research community later this summer. Stampede 2 will undergo several upgrades to give the new system nearly double its predecessor's computing power, delivering a peak of 18 petaflops. Increases in performance are enabled by the inclusion of the latest generation Intel Xeon and Intel Xeon Phi processors and enhanced networking though the Intel Omni-Path Architecture. With these latest Intel technologies, Stampede 2's augmented capacity will help enable the next generation of researchers to drive breakthrough scientific discoveries in simulation, modeling, artificial intelligence and machine learning.

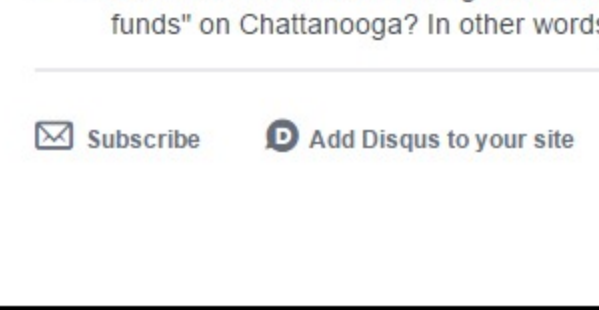
Thanks to the combined effort of TACC's team and researchers, it seems the gap between supercomputers' advancements in artificial intelligence and the natural processes they emulate is narrowing. With ever-increasing supercomputer power, increasingly intricate research projects will ultimately give up their secrets in support of human breakthroughs. Levin echoes the optimism about supercomputers' potential today, and the many ways they can lead to possible health-related innovations of the future. "Machine learning and artificial intelligence made possible by supercomputers are not just for robotic tasks and number crunching. They help us understand the living world better."

Rob Johnson spent much of his professional career consulting for a Fortune 25 technology company. Currently, Rob owns Fine Tuning, LLC, a strategic marketing and communications consulting company based in Portland, Oregon. As a technology, audio, and gadget enthusiast his entire life, Rob also writes for TONEAudio Magazine, reviewing high-end home audio equipment.

RELATED READS

- 

A Spatiotemporal Model to Unlock The Secrets of the Human Heart
- 

Efficient Simulation
- 

Optimizing Physics, Computational Science at Princeton University
-

Dell EMC Expands HPC Portfolio with Advances in Cloud, Software and Systems

0 Comments

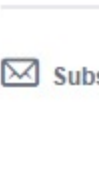
Scientific Computing

1 Login

Recommend

Share


Sort by Best



Start the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS



Be the first to comment.