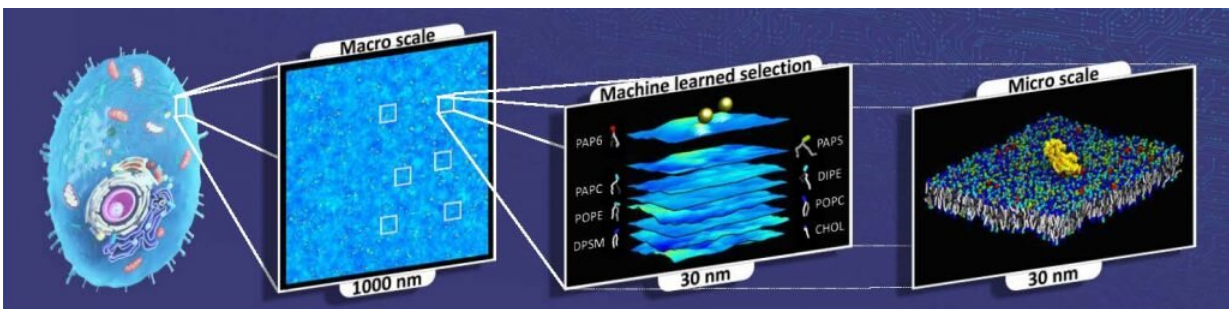# Scientists tap the power of high-performance computing to understand cancer growth

January 12 2021, by Andrea Schneibel

Addressing many important biological questions requires large length- and time-scales, as well as molecular level details. Shown here is a simulation of protein-lipid dynamics for a 1 μm x 1 μm membrane subsection at near-atomistic resolution. The images were acquired using the Multiscale Machine-Learned Modeling Infrastructure (MuMMI). Credit: Tim Carpenter, LLNL

Researchers are running thousands of simulations on Summit to identify how a specific protein triggers up to 30 percent of all human tumors

A multi-institutional team of scientists is using Summit, America's fastest supercomputer for open science, to get a better understanding of how certain proteins signal body cells to reproduce uncontrollably, triggering cancer.

Summit is part of the Oak Ridge Leadership Computing Facility

(OLCF), a US Department of Energy (DOE) Office of Science User Facility located at DOE's Oak Ridge National Laboratory (ORNL).

The effort is one of several projects awarded hours on the supercomputer by DOE's Office of Advanced Scientific Computing Research's (ASCR's) Leadership Computing Challenge (ALCC) during the 2019–2020 period, and again for 2020–2021.

Led by scientists at Lawrence Livermore National Laboratory (LLNL), the team recently completed 400,000 Summit node-hours of simulations in an attempt to gain knowledge of RAS, a family of proteins that, when mutated, has been found to drive about 30 percent of human tumors, including some of the deadliest ones, such as 95 percent of pancreatic cancers and 45 percent of colorectal cancers.

They did this by running a total of 2,600 simulations, up from their originally planned 600.

"This translates into hundreds of thousands of [computational] jobs, all running simultaneously, tracking different data sets across multiple systems of the supercomputer itself. Without a system like Summit, this would have not been possible," said Felice Lightstone, Biochemical and Biophysical Systems Group leader in the Biosciences and Biotechnology Division at LLNL.

RAS: a team player, until it's not

Humans and RAS proteins are no strangers to each other. In fact, RAS proteins are present in all the cells of our bodies, with the exception of red blood cells, and play a significant role in signal transduction, a vital process that affects the biological function of a cell.

"From that perspective, we really need RAS proteins just for normal

living," Lightstone said.

In some people, however, certain RAS proteins mutate into what scientists call oncogenic RAS. These mutated proteins are always in the "on" or "active" state, constantly signaling the other proteins they bind to to grow nonstop. This is how cancer begins.

Gaining knowledge of this initiation process could eventually assist in the development of therapies that target oncogenic RAS proteins so they can be "turned off," said Lightstone, while allowing the other normal-functioning RAS proteins to continue the job they need to do in a healthy system.

This would be a groundbreaking achievement because several mutated RAS proteins—which are very small, balloon-shaped structures—have been deemed "undruggable" by scientists for several decades now.

"Now that we have better biochemical and biophysical techniques, and with the possibilities simulation offer, we believe that we could try to eventually get to a druggable RAS," Lightstone said.

To understand how the oncogenic RAS protein is activated in a cell membrane—as well as how such interaction with the membrane influences the activity of other types of proteins called effectors which promote cell growth—the team had to look at thousands of possible combinations, a massive task that's only possible with the help of a supercomputer such as Summit.

"Summit has allowed us to run our own workflows, composed of a very heterogeneous set of software tools that we have been tuning and developing to do this kind of research at such large scales," said Harsh Bhatia, a computer scientist at LLNL's Center for Applied Scientific Computing.

They also faced another challenge: Biology researchers frequently deal with large ranges of timescales and length scales, from atomistic and molecular scales to macroscopic scales that make simulations spanning the full range of interactions very difficult. Reactions unleashed by RAS proteins occur at atomistic levels, and simulations of these reactions tend to be too short and too small in spatial extent to model the full impact on the larger biological system.

To solve this issue, the team developed a multiscale capability called MuMMI, which stands for Multiscale Machine-learned Modeling Infrastructure. MuMMI allowed them to take information from shorter, smaller simulations and feed it into a larger model to scale it up into longer timescales and larger systems.

"This helps us ask other questions, like does this oncogenic RAS protein aggregate with other RAS proteins? Or does it have lipid interactions that can either govern or be influenced by that aggregation, which is important to whether or not subsequent proteins bind to it?" Lightstone said. "There are hundreds, if not thousands of possible combinations we need to look at."

The team is using Summit to get a statistical sampling from tens of thousands of simulations, which in turn helps them benchmark their models for further improvements.

The results of this first ALCC allocation will also allow the team to train machine learning (ML) models that will be used to create an algorithm to run multiscale simulations during the next phase of the study, which is already under way with a new ALCC allocation of 600,000 Summit node-hours for the 2020–2021 period.

For this new phase, scientists will add a new layer of complexity to the study by taking two different types of an effector protein called RAF

and binding them to the RAS protein to explore how they interact with the [cell membrane](link) to issue the first signal of cancer initiation.

"To parameterize the new macro model and collect input structures for the next MuMMI simulation campaign, we used over 20 milliseconds of molecular dynamics simulation data, analyzing the orientational and lipid dependence of RAS and RAS-RAF complexes," explained Helgi Ingolfsson, computational biologist at LLNL's Biochemical and Biophysical Systems Group.

Collaboration for the win

Even though exploring the role of RAS proteins in the development of many types of cancers has been a research priority for decades—the National Cancer Institute (NCI) has been operating its RAS initiative since 1970—it wasn't until now that scientists had a chance to use the power of supercomputers in their research.

About four years ago, DOE opened its doors to scientists wanting to use high-performance computing to solve questions in the realm of biology. With this in mind, Lightstone's team approached the NCI to see how they could participate in the effort as part of its Cancer Moonshot project, an initiative that aims to accelerate the discovery of cancer therapies within seven years. (The program was funded in December 2016 with the approval of the 21st Century Cures Act.)

The first time the team tested its RAS research capability was on the IBM-built Sierra, a supercomputer that joined the ranks of LLNL in 2018. Gaining an allocation on Summit has allowed them to continue their collaboration with NCI.

"We always have to keep that in mind that this application and the science to answer the biology is important. And just as equally important

is all the science and engineering that's going in developing the capability," Lightstone said.

More information: Francesco Di Natale et al. A massively parallel infrastructure for adaptive multiscale simulations, *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (2019). DOI: 10.1145/3295500.3356197

Provided by Oak Ridge National Laboratory

Citation: Scientists tap the power of high-performance computing to understand cancer growth (2021, January 12) retrieved 2 May 2024 from https://medicalxpress.com/news/2021-01-scientists-power-high-performance-cancer-growth.html