CRAY®

The Pawsey Centre's Cray® XC40™ Supercomputer "Magnus" Helps Battle Major Cause of Hunger in East Africa

Organizations
University of Western Australia
School of Chemistry and Biochemistry
Australian Research Council Centre of Excellence in Plant Energy Biology
Crawley, Australia
www.chembiochem.uwa.edu.au

Pawsey Supercomputing Centre
Kensington, Australia
www.pawsey.org.au

PAWSEY supercomputing centre

Scientific Field
Evolutionary Biology

Applications
MrBayes

About the Pawsey Centre & Magnus
The Pawsey Supercomputing Centre supports researchers with supercomputing, data and visualization services across a range of scientific fields. "Magnus" is a petascale Cray® XC40™ supercomputer and the most powerful system in the Southern Hemisphere.

"Magnus is changing the world in agricultural development."

— Dr. Laura Boykin
Assistant Professor
University of Western Australia

Background

Devastation sometimes comes in the tiniest packages. The silverleaf whitefly registers at barely a single millimeter. Yet an infestation of the sap-sucking insects can destroy entire crops — and livelihoods.

In East Africa, the whitefly, or *Bemisia tabaci*, attacks the cassava plant, a crucial food source for the region. A slow-growing crop, farmers rely on cassava to bridge the nutritional gap between growing seasons after other crops such as bean and sweet potatoes have been consumed. However, the whitefly transmits two cassava-killing viruses — cassava mosaic disease and cassava brown streak disease. Together, these viruses will completely wipe out a year's product. For a family, a whitefly infestation suddenly means no food. For the region, it can mean widespread economic hardship and famine.

Bemisia tabaci is a worldwide agricultural pest costing global agriculture billions of dollars a year. But because of their particularly devastating effect on East Africa, a team of researchers at the University of Western Australia, led by Assistant Professor Laura Boykin is focusing their research efforts there. "It's a massive problem," says Dr. Boykin. "I'm one of 15 principal investigators working on a new project whose mission is to give farmers a cassava plant that's resistant to the viruses and the whiteflies."

The team is funded by the Bill and Melinda Gates Foundation and was awarded time on the Pawsey Supercomputing Centre's Cray® XC40™ "Magnus" supercomputer. Underscoring the urgency of the whitefly problem, Dr. Boykin also won a 2015 TED Fellowship in support of her work.

CASSAVA KILLER: The silverleaf whitefly transmits viruses that destroy cassava plants and render their roots inedible. Photo courtesy of Laura Boykin

Challenge

For decades, scientists assumed they were battling a single silverleaf whitefly species. In reality, they're battling a species complex of at least 34 morphologically indistinguishable species. "It's only been in the last seven years or so that people have started to do sampling of the region," says Dr. Boykin. "The more we sample, the more we realize there are tons more species of whitefly in Africa than we ever thought."

Dr. Boykin and her team are employing genomics, supercomputing and evolutionary history to understand the enemy's genetic differences. This knowledge will help African farmers and scientists distinguish between harmless and invasive whitefly species, develop management strategies and breed new whitefly-resistant strains of cassava.

The computational challenge for the team is in making sense of the vast amount of genomic data their sequencing machines produce. "We have the task of trying to make sense out of billions of base pairs — billions of As, Ts, Gs and Cs at a time," says Dr. Boykin.
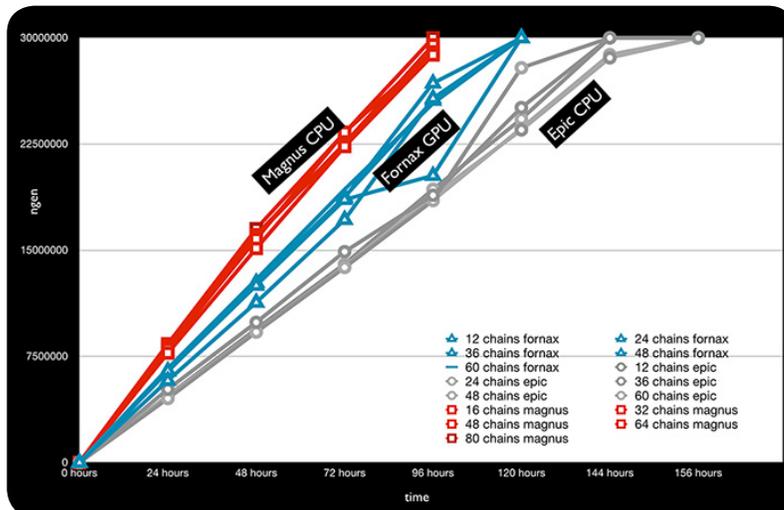
Solution

Using Magnus, a petascale Cray XC40 supercomputer at the Pawsey Supercomputing Centre, the team is generating phylogenetic trees of whitefly species from around the world. Phylogenetic trees represent evolutionary relationships, or genealogy, among species. For this project, the genetic datasets involved thousands of base pairs. Even with only 500 whiteflies in a dataset, the possible relationships between these flies run into the octillions (a 1 followed by 27 zeros) — a calculation impossible without a supercomputer.

The team is running MrBayes — a program for Bayesian inference and model choice across a wide range of phylogenetic and evolutionary models. It uses Markov chain Monte Carlo methods to sample the massive evolutionary tree space. (A Markov chain is a random process that undergoes transitions from one state to another on a state space.)

RACE TO SAVE CASSAVA: Magnus handled more Markov chains in less time than its closest competitor. Image courtesy of Laura Boykin

Given the large size of the genetic datasets and sophisticated computing techniques involved, the project is highly computationally intensive. Magnus is well suited for this kind of problem. Equipped with multiple processor technologies, a high-performance network, distributed operating system and a productive programming environment, the Cray XC40 supercomputer excels at large-scale computations and reduces processing times. Additionally, Cray architected the XC40 system to be easily customized and upgraded. Magnus consists of eight compute cabinets with 384 blades and four nodes per blade. Each node supports two 12-core, Intel® Xeon® E5-2600 v3 "Haswell" processors for a total of 35,712 cores. The compute nodes communicate over Cray's high-speed, low-latency Aries interconnect.

So far, the team has analyzed an entire genetic region for all the global samples. "We've done benchmarking against our other systems and Magnus outperforms them," says Dr. Boykin.

Between 16 and 80 Markov chains traversed the tree space on Magnus in just under 96 hours. The next best performing system handled about half as many chains in 110 hours. "I feel it's the Aries interconnect that makes it so much faster. Our analyses are suited for nodes talking to one another," says Dr. Boykin.

But perhaps most importantly, Boykin and her team are making meaningful progress toward distinguishing damaging whiteflies from others and arming scientists with the information they need to develop management strategies. "Magnus is changing the world in agricultural development," she says.

For more information:

The Boykin Lab
News and research information about Dr. Boykin's work on whiteflies
www.lauraboykinresearch.com

Whitefly Database
Complete collection of curated 3′ mtCOI whitefly data
www.whiteflybase.org

WORLDWIDE AUDIENCE: Dr. Laura Boykin won a 2015 TED Fellowship for her whitefly research. Here, she discusses how Magnus and the "amazing" Pawsey Centre staff are aiding this urgent work. Photo courtesy of TED/Ryan Lash