

# Researchers use TACC supercomputers to uncover the genetic roots of Yellow Canopy Syndrome

August 8 2017

---



Yellow Canopy Syndrome (YCS) is a condition affecting sugarcane in Australia, the cause of which has not been identified. YCS causes yellowing of the mid-canopy leaves and reduces the amount of sugar that can be extracted from the cane. Researchers are using TACC's supercomputers to investigate possible clues to YCS's cause in genetic data. Credit: Annelie Marquardt, Sugar Research Australia

Since 2011, a mysterious illness known as Yellow Canopy Syndrome, or YCS, has afflicted Australian sugarcane. The condition causes the mid-canopy leaves of otherwise healthy plants to rapidly turn yellow to a degree that the plant's sugar yield can decrease by up to 30 percent.

In recent years, the syndrome has spread across the continent. Losses are estimated at around \$40 million and growers fear it could ruin the industry in Australia.

"At the start of the project, there were many possibilities but little evidence to suggest the cause," says Kate Hertweck, an assistant professor of biology at The University of Texas at Tyler (UT Tyler) and a member of the team of researchers exploring the causes of the disease. "It could be a physiological reaction caused by water or nutrients in the soil. Or it could be a biological cause, like an insect, virus or fungus."

Whereas some researchers use field experiments and microscopy to investigate the disease, Hertweck and her collaborators from Sugar Research Australia and the University of Queensland are pursuing a genomic approach, using next-generation RNA sequencing to compare and analyze genetic data from affected and unaffected plants from diverse field locations over a three-year time span.

"Sugarcane is an important agricultural crop," says Kate Wathen-Dunn, senior technician at Sugar Research Australia. "It also has one of the most complex genetics known, with multiple and variable numbers of each chromosome."

In part because of its complexity, sugarcane did not previously have a reference genome available as a starting point of comparison for researchers, so Hertweck and Wathen-Dunn set out to create one, specifically for the plant's transcriptome: the messenger RNA molecules expressed by the genes of an organism, which determine what proteins

the plant will produce.

"With this transcriptome reference, we could compare the Yellow Canopy Syndrome and control samples taken at different times from different varieties and from different growing regions," Wathen-Dunn says. "As the amount of data involved was enormous, the only way to do this assembly was on a high-performance computing cluster."

Hertweck and her team turned to the supercomputers at the [Texas Advanced Computing Center \(TACC\)](#), based at The University of Texas at Austin, to perform their large-scale investigations. TACC runs several of the largest supercomputers in the world, which support thousands of U.S. researchers each year.

Transcriptome assemblies take RNA molecules that have been fragmented and sequenced, and puts them back in order. The process is always computationally-intensive, but when there are many samples—as was the case with the sugarcane research—it can be particularly unwieldy. The team gathered RNA sequence data from 70 leaf samples and used multiple algorithms and multiple subsequences to create a de novo assembly.

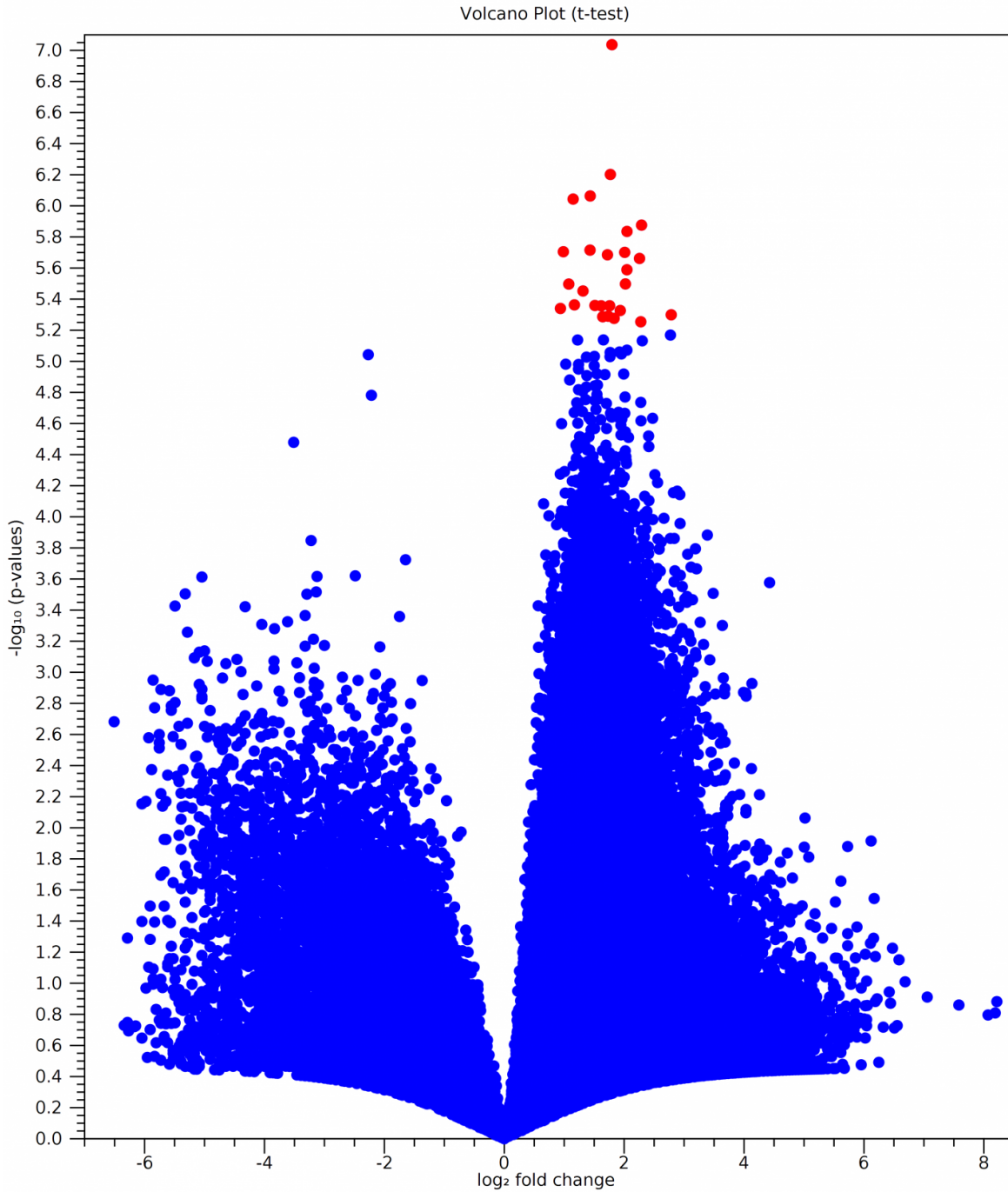
"Even in a compressed form, the file sizes for transcriptome assemblies are enormous," Hertweck says. "I started realizing that I needed much larger computing resources than I had available."

The reference transcriptome they created allowed the team to explore how different samples express different proteins, which provide clues to YCS's root cause.

So far, Hertweck and her team have performed preliminary differential gene analyses on a subset of the data and are using the assemblies to assess a variety of hypotheses for what could be causing the disease.

"If it's bacteria, then there are genes that might be expressed. If it's a virus, separate genes might be expressed," she says. "We've detected some differences that could be a sign of a bacteria, but these are sometimes also related to lab contaminants."

Further investigations will determine whether the gene expression relates to the disease's true cause or is a false signal.



A volcano plot of the change in transcript expression in YCS plants, showing the amount of expression change and its importance, with the most important transcripts shown in red. Transcripts were de novo assembled from 595 million RNAseq reads using the Stampede cluster on TACC. Credit: Kate Hertweck and Kate Wathen-Dunn

They have also found several signals of physical (or abiotic) stress in the data, which require further investigation.

"Abiotic stress is very important in symptom expression, and YCS-affected sugarcane plants appear to be more sensitive to these stresses," says Wathen-Dunn. "The fantastic computing resources available at TACC allowed us to pursue our research into the cause of YCS, and to make new discoveries about the metabolism of sugarcane."

Hertweck and Wathen-Dunn presented their research at the [2016 Evolution Meeting](#) —the joint annual conference of the Society for the Study of Evolution, the Society of Systematic Biologists, and the American Society of Naturalists. The work was also presented at the [2016 Australian Bioinformatics and Computational Biology Society \(AB3ACBS\) Conference](#).

They will present their latest results at the [2017 International Tropical Agriculture Conference](#) in November.

The ability to delve into the transcriptome of the sugarcane using TACC resources has impressed Wathen-Dunn's Australian collaborators.

"They thought it was amazing that we have computers that can manage all of this," she says. "It has encouraged them to have some of their employees get trained to interact with high performance computing resources."

## **RESEARCHING AND TEACHING WITH TACC SYSTEMS**

Hertweck first learned about TACC when she applied for her faculty position at UT Tyler. The job postings mentioned that researchers would be able to use TACC systems through the [University of Texas Research Cyberinfrastructure \(UTRC\)](#) initiative, which, since 2007, has provided researchers at any of the University of Texas System's 14 institutions access to TACC's resources, expertise and training.

Hertweck attended the TACC Summer Institute in 2015 and has used TACC's advanced [computing resources](#), including [Stampede](#) and [Lonestar 5](#), ever since for genomic studies of grasses, lilies, irises, orchids and fruit fly genomes.

"Part of the struggle for a small regional university like mine in being able to attract strong candidates is the fact that we have no clusters locally here on campus," Hertweck says. Easy access to TACC resources addresses that problem.

"There is a huge desire for that type of resource. People are very interested in being able to take advantage of it," she says. "My job was posted with a desire of bringing someone in who could take advantage of those resources and open the door for other researchers to use them as well."

Hertweck has done just that. In addition to her own research, Hertweck acts as a champion for high-performance computing among her colleagues and teaches Bioinformatics courses to undergraduates and graduate students using TACC's Jetstream cloud computing system. In her classes, students explore sequence data from bacteria that are being investigated for future biofuel applications and other genetically interesting species.

"Entire classes of graduates and undergraduates run basic analyses on TACC systems. They do assemblies and basic analyses and see what the

sequences show," Hertweck says. "They get most excited about the ability to analyze new data and find cool things. I tell them: 'You're going to have a piece of information that no one has ever seen before.'"

Helping researchers and students uncover new facets of nature—that's exactly what TACC's systems are built for.

Provided by University of Texas at Austin

Citation: Researchers use TACC supercomputers to uncover the genetic roots of Yellow Canopy Syndrome (2017, August 8) retrieved 25 April 2024 from <https://phys.org/news/2017-08-tacc-supercomputers-uncover-genetic-roots.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.