

BREEDING A HARDIER CABERNET: SMRT® SEQUENCING PROVIDES DETAILED VIEW OF GRAPE GENOME



At the University of California, Davis, Dario Cantu is applying long-read PacBio sequencing to the heterozygous genome of the Cabernet Sauvignon grape. Now, his team has access to whole genome data that could help guard against the effects of climate change and disease.

Dario Cantu is on a mission to protect and improve wine grapes. These plants are under tremendous pressure from disease and climate change, but Cantu believes that advances in genomics, including Single Molecule, Real-Time (SMRT) Sequencing, could provide important insights into better ways to breed and grow wine grapes.

Cantu's lab is based at the University of California, Davis, near the heart of California's wine country. He and his team have used SMRT Sequencing to study grape pathogens and recently completed a highly contiguous assembly of the Cabernet Sauvignon grape variety. According to Cantu, the research community is entering a golden age for studying these plants; wine grapes are often first-generation crosses of other varieties, and their complex genomes are notoriously difficult to assemble due to high levels of heterozygosity. Comprehensive resolution of these genomes will enable the first genomics-based approach to breeding wine grape varieties for optimal growth traits, including resistance to disease and drought.

Currently, intentional breeding for wine grapes by crossing varieties is guided by the pursuit of optimal flavors. Making plants hardier has simply not been a factor. "What we grow are the worst genotypes for many reasons, but they are the best for flavor," Cantu says. "They are susceptible to diseases and soil-borne pathogens. They are absolutely not drought tolerant. But we still grow them because they have the names we want to associate with our wines."

Using resources such as his new grape genome assembly, Cantu envisions a future where breeders can apply detailed genetic knowledge to maintain or improve grape flavor while also making the plant more resilient to temperature change, lack of water, and common pathogens. Beyond this, he says a better understanding of grape genetic history may lead to revelations about human migration patterns as wine production has been linked with human civilization throughout our history.

The Genomics of Wine

Cantu's focus on wine grapes began in 2012. "I work on the vineyard ecosystem," he says. "I sequence the host, the plant, and all the organisms associated with the plant."

A plant biologist, Cantu was surprised by how little advanced breeding had been applied to wine grapes over the years. "For wheat, corn, or rice, there's a different variety for every region, and there's a new variety every two or three years," he explains. "We grow grapes everywhere in California, but the genotypes are the same. The Cabernet Sauvignon grown in Central Valley is the same Cabernet that is grown in Napa and Sonoma."

As the effects of climate change become more severe, Cantu predicts that optimal growth of wine grapes will require the same kinds of frequently updated, region-specific varieties used for most crop plants. His lab is applying genomics to grapes, their beneficial microbes, and their pathogens with precisely this goal in mind.

"We can easily develop new varieties to take some traits from wild species that are more tolerant to drought or disease, but we want those varieties to be competitive," Cantu says. "Despite the fact that we know a lot about flavors, we know basically nothing about the genes in the grape vines that are responsible for those traits. There is a huge need for flavor genetics, as well as for developing markers of drought and disease resistance."

A near-term benefit of developing high-quality genomic resources including a contiguous and complete assembly will be a clearer view of the genetic and geographic history of wine grapes. "All the varieties we use are basically crosses of other varieties," Cantu says. For instance, Cabernet Sauvignon came about by crossing Sauvignon Blanc and Cabernet Franc, which itself is only one generation removed from Merlot. Sequencing and *de novo* assembly of several of these varieties will provide clues to their phylogeny and geographical spread around the world, and can also be used to select markers to properly differentiate clones.

Ultimately, this information will help unravel the complex traits contributing to flavor. "The idea is to reconstruct the phylogenetic tree of flavor for wine grape varieties to identify the unique alleles and the shared genes that define certain flavors," Cantu says.



At the University of California, Davis, Dario Cantu is applying SMRT Sequencing to explore the entire vineyard ecosystem, including the grape varieties and their associated pathogens. Photo credit: Gregory Urquiaga/UC Davis.

SMRT Approach

His team has been working with PacBio long-read sequencing for some time now, applying it first to the microbial communities that live around grapes. "PacBio has been our sequencing platform of choice for all the microbial work we do – all these trunk diseases and grape diseases," Cantu says. For most fungal pathogens they sequenced, the lab was able to reconstruct entire chromosomes into one-contig, telomere-to-telomere assemblies. The team has also performed metatranscriptomics as well as deep sequencing of ribosomal RNA. The end result was a profile of the gene expression showing how these organisms interact with each other and their grapevine host.

Now, Cantu's team is taking advantage of more recent P6-C4 chemistry to generate longer SMRT Sequencing reads for the grape genomes themselves. They knew from experience that their grape of interest – *Vitis vinifera* cv. Cabernet Sauvignon

– was highly heterozygous. "Every short-read assembly we tried broke because of the heterozygosity," Cantu says.

But with PacBio's extraordinarily long reads, high accuracy, and diploid-aware assembler, the scientists finally made inroads. They sequenced the genome with the PacBio RS II System, using 20 kb and 30 kb libraries, and emerged with a highly contiguous assembly marked by a contig N50 of 2.17 Mb. Their first attempt at using SMRT Sequencing for this grape immediately represented a 10-fold improvement over a previous Sanger assembly, Cantu says.

In an evaluation of assembly completeness, the team aligned mRNA sequences previously annotated to the existing *V. vinifera* genome reference. They not only found that the majority of genes were present in the PacBio assembly, but could now distinguish the haplotype phasing.

The FALCON and FALCON-Unzip assembly tools were critical to the success of this project. "With diploid genome assembly and haplotype phasing, these algorithms were what we needed because now we can look at the Cabernet Sauvignon alleles – the gene copies that come from either the Cabernet Franc or the Sauvignon Blanc – rather than looking at the haploid consensus," Cantu says. "That's a breakthrough in grape genetics, or in any genomics field where you are dealing with hybrids."

Cantu's lab is now sequencing the parent grapes of Cabernet Sauvignon and expects to generate high-quality results for those, too. With the current chemistry and continually increasing read lengths combined with FALCON and FALCON-Unzip, he says, "we are moving the grape genomics field forward at an unprecedented pace because we have the best tools in our hands." Despite the incredibly complex genomes, Cantu ultimately aims to assemble these and other important grape varieties into one-contig-per-chromosome sequences.

"Now we can look at the Cabernet Sauvignon alleles – the gene copies that come from either the Cabernet Franc or the Sauvignon Blanc – rather than looking at the haploid consensus," Cantu says. "That's a breakthrough in grape genetics."

"The complexity of the phenotypes of grapes requires a huge effort from a genetic perspective," says Cantu, noting that the results his team has produced already will go a long way toward enabling faster crop improvement. "These tools are really revolutionizing the way we can do grape genetics."

