

Strategy to derive a sequence assembly for plant genome of bread wheat

April 16 2015



The genome of bread wheat is five times larger than that of humans and contains many repeat sequences, making it very challenging for plant researchers to obtain a reference sequence for this important crop. Credit: Jordi Paya Canals via Flickr, CC-BY-SA 2.0

Through a combination of high-throughput sequencing, high

performance computing, and genetic mapping, DOE JGI researchers have derived a sequence assembly for the highly repetitive plant genome of bread wheat (*Triticum aestivum*).

The researchers pointed out that the strategy they developed and applied to the polyploidy species [bread wheat](#) is also applicable to any species, for as long as a mapping population can be constructed.

Genome researchers have spoken often of the challenges of sequencing plant genomes with more than two paired sets of chromosomes (polyploidy). For example, in the case of bread wheat (*Triticum aestivum*), multiple hybridization events have led to a genome that is five times larger than that of humans and more than 80 percent of it made up of repeat sequences. To generate a whole-genome shotgun assembly of such a complex genome, researchers need to be able to assemble short reads coming off next-generation sequencing platforms and also use significant computational resources to manage the data.

In the January 31, 2015 issue of *Genome Biology*, an international team of researchers led by U.S. Department of Energy Joint Genome Institute (DOE JGI), a DOE Office of Science user facility managed by Lawrence Berkeley National Laboratory, described a strategy that employs the above criteria. Using short read sequencing technology, the team assembled over 9-billion basepairs (Gbp) of the 16 Gbp bread wheat genome using Meraculous, a whole genome assembler for next-generation sequencing data geared for large genomes, and access to the Edison supercomputer at the National Energy Research Scientific Computing Center (NERSC), another DOE user facility, saving the team days of compute time.

The team's work follows the International Wheat Genome Sequencing Consortium's release of the bread wheat genome in *Science*. DOE JGI Chief Informatics Officer and former Plant Program head Dan Rokhsar

noted that he and his colleagues took part in the IWGSC effort, contributing a dense genetic map.

In the current paper, the team highlighted some of the differences between the IWGSC's sequence, and their own whole-genome shotgun draft genome assembly: "At first glance, the summary statistics of our assembly might look unimpressive... However, the fraction of the genome in assembled contigs is in the same range as the chromosome-by- chromosome shotgun assembly of IWGSC (9.1 Gbp vs. 10.1 Gbp), suggesting that the problems are intrinsic to the [wheat genome](#) and short read datasets." They added that their whole-genome shotgun approach allowed them to anchor more than 7.1 Gbp of the genome to chromosomal locations, while from the previous reference sequence, only 4 Gbp is anchored.

The impact of the team's whole-genome shotgun strategy was highlighted in a separate Genome Biology article, in which the authors compared the DOE JGI's plant genome sequence assembly against the bread wheat sequence released in Science last year. They noted that, "The combination of whole genome shotgun sequencing and linkage mapping by skim sequencing produced a better genome assembly than both the chromosome arm-based assembly and a previously described whole genome shotgun sequencing assembly approach." Additionally, they concluded that "[this method] ... should also be applied to draft genome sequencing of all species, both diploid and polyploid.

The team also compared the utility of their whole-genome shotgun assembly of the bread wheat plant [genome](#) to the "highly fragmented assembly" of barley, which is a DOE JGI Community Science Program project. They noted that the barley assembly has enabled the development of reference-based genetic mapping and cost-efficient resequencing strategies, among other benefits.

The longer-term impact of this work is that this improved strategy for dealing with large and complex genomes (frequently characteristic of plants that, as potential biofuel feedstocks, are of major interest to DOE) will accelerate and advance the sequencing of, and exploitation of, plant genomes.

More information: "A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome." *Genome Biol.* 2015 Jan 31;16(1):26. [DOI: 10.1186/s13059-015-0582-8](https://doi.org/10.1186/s13059-015-0582-8)

Provided by DOE/Joint Genome Institute

Citation: Strategy to derive a sequence assembly for plant genome of bread wheat (2015, April 16) retrieved 28 April 2024 from <https://phys.org/news/2015-04-strategy-derive-sequence-genome-bread.html>

<p>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.</p>
--