

IPK scientists: New method for genome assembly in barley provides excellent results

Gatersleben, 17.03.2021 Genomes differ between the individuals of one species and we can learn a lot about diversity in our crops by comparing genomes of different varieties. However, researchers that want to study many genomes need a fast and reliable method for sequence assembly. An international research team led by the IPK Leibniz Institute has now investigated a new DNA sequencing method. The results, which have now been published in the journal "The Plant Cell", are very promising. The scientists now hope to be able to use the method for assembling other barley genomes in the future.

The genome encodes the entire genetic information of an organism. It is stored as a DNA nucleotide sequences in the nucleus of each cell of an organism. The DNA is organized into chromosomes, structures large enough to be visible under a microscope.

The genome of the crop plant barley consists of about five billion base pairs. Sequencing and assembling a genome means reading out all its nucleotides and arranging them into a computer-readable text. Genome assembly used to be very difficult. For example, it took over ten years to assemble the sequence of the barley reference genome. "Therefore, we wanted to know whether a new method of DNA sequencing with high accuracy and long read lengths could simplify genome assembly in resequencing, but also the initial sequencing of a reference genome," explains Dr. Martin Mascher, head of the Independent Working Group on Domestication Genomics at IPK and member of the German Center for Integrative Biodiversity Research (iDiv).

The international research team used PacBio HiFi sequencing. HiFi sequencing is a method to read long DNA sequence segments of up to 25,000 base pairs, so-called "reads", with high accuracy. The reads were assembled into sequences, which represent entire chromosomes. The scientists also made genome assemblies from other types of sequence reads generated with alternative methods such as short reads or long-reads with high error rates.

When they compared the different assemblies to each other, the HiFi provides the best quality genome sequences. The HiFi assembly approach contained more complete genes than the other assembly comparison methods. Even when the genes were present in multiple, nearly identical copies, HiFi succeeded in distinguishing and accurately assembling them. HiFi also managed to correctly reassemble almost all of the non-coding and highly repetitive sequence regions between the actual genes. "The most important thing for us, apart from the accuracy, is the speed. The HiFi sequencing works very fast, so we were able to complete highly accurate and complete genome assemblies within a few days", explains Dr. Martin Mascher.

The results now open up new possibilities for the scientists. Genome assembly will remain a challenge in the future. However, the new HiFi approach and their many years of



Scientific Contact
Dr. Martin Mascher
Tel.: +49 39482 5243
mascher@ipk-gatersleben.de

Media Contact
Christian Schafmeister
Tel. +49 39482 5461
schafmeister@ipk-gatersleben.de

experience in sequencing and assembling cereal genomes will allow IPK researchers to reliably and quickly assemble the genomes of cultivated barley and related forms. "We will use HiFi sequencing to assemble the genomes of further barley varieties. Our gene bank at the IPK Leibniz Institute in Gatersleben will provide us with further cultivars and related wild forms for this purpose. Our goal is to record the genetic diversity that exists worldwide", explained Dr. Martin Mascher and already describes the next challenges. "We will search for previously unknown genetic variation, such as resistance genes, in order to further enrich the genetic the genetic possibilities of the for cultivated barley and to be able to support new breeding."

Original publication: Mascher *et al.* (2021), Long-read sequence assembly: a technical evaluation in barley. *The Plant Cell*. DOI 10.1093/plcell/koab077

Figures (for free use):

<https://ipk-cloud.ipk-gatersleben.de/s/sWFg3W5bCRXFZb3>



The genome of the crop plant barley consists of about five billion nucleotides. When the international research team led by the IPK compared the different assemblies to each other, the HiFi assembly performed best.

Photo: Eva Siebenhühner/ IPK