Going deep to improve maize transcriptome

April 29 2014

Deep sequencing of the maize (Zea mays) transcriptome provides researchers with a near-complete snapshot. Credit: H. Zell, Wikimedia Commons CC 3.0

A team of researchers from the U.S. Department of Energy Joint Genome Institute (DOE JGI), the University of California, Berkeley, and the Great Lakes Bioenergy Research Center generated an ultra-deep, high quality transcriptome–the fraction of the genome that provides

information about gene activity–of maize.

Employing the novel strategy of assembling a hybrid transcriptome generated through deep RNA sequencing technologies allowed researchers to improve the current gene annotations for maize, revealing several unique features.

With 2.3 billion bases, maize (Zea mays) is one of the largest genomes sequenced and assembled thus far. As a crop that provides fuel, food and animal feed, maize's genetic code offers researchers tools with which they can breed desired traits into other plants. In particular, the transcriptome provides information on which genes are active in maize.

As reported in the March 31, 2014 issue of *Scientific Reports*, the team compared the comprehensiveness of a plant transcriptome produced with the help of increased read depths from RNA sequencing (RNA-Seq) technologies against the previously generated transcriptome that relied on more traditional EST sequencing. The researchers generated a de novo assembly of the sequence data produced by Illumina and Pacific Biosciences platforms with the help of Rnnotator, an automated software pipeline developed at the DOE JGI.

Generating nearly 3 billion RNA-Seq reads, the team produced what they described as "a near complete snapshot" of the maize transcriptome, with over 90 percent of the annotated transcripts after adding nearly 5,000 previously unannotated transcript variants. Deeper sequencing, they noted, allowed them to identify transcripts that encoded proteins with regulatory roles that had not been found with when typical sequencing read depths were employed. The team assessed the quality of the assembly using metrics such as accuracy, completeness and contiguity. They then looked at regions of conserved gene order and content in maize, comparing them against matching regions in the genomes of the closely related grasses rice and sorghum.

"By applying a hybrid assembly approach we compiled a high quality, comprehensive annotation of the maize transcriptome," the team concluded. "These findings highlight the utility of ultra-deep RNA-seq for comprehensive transcriptome annotation in complex large plant genomes, especially for those that lack robust reference genomes…. The insights gained in this study are not limited to plant genomes, but could be generically applied to any large and complex transcriptome."

More information: Martin JA et al. "A near complete snapshot of the Zea mays seedling transcriptome revealed from ultra-deep sequencing." *Sci Rep.* 2014 Mar 31;4:4519. DOI: 10.1038/srep04519.

Provided by US Department of Energy