Maize is an amazingly diverse crop. A study in 2009 demonstrated that half of the genome sequence and one-third of the gene content between two inbred lines of maize were not shared. This diversity, which is more than two orders of magnitude larger than the diversity found between humans and chimpanzees, highlights the inability of a single reference genome to represent the full pan-genome of maize and all its variants. Here we present and review several efforts to characterize the complete diversity within maize using the highly accurate long reads of PacBio Single Molecule, Real-Time (SMRT) Sequencing. These methods provide a framework for a pan-genomic approach that can be applied to studies of a wide variety of important crop species.

### Improved Genome Annotation

In addition to improving the maize reference genome for a more comprehensive view of genetic diversity, an effort to fully characterize the transcriptome with long reads was undertaken. Using the PacBio Iso-Seq method for RNA sequencing, researchers were able to produce >100,000 non-redundant full-length isoforms, 57% of which were novel. These isoform sequences were able to correct gene models and identify previously missed genes, boosting the quality of the genome annotation. These benefits were largely due to the fact that sequencing CDNA with PacBio generates full-length transcripts with no assembly required.

### Discovering New Variation

The version 4 reference genome, with its highly contiguous sequence, allowed for a genome-wide evaluation of structural variants in two lines of maize, K11 and W22. The optical maps showed only 32% of the K11 and 39% of the W22 maps were alignable to the B73 reference, highlighting the tremendous genetic diversity found within Zea mays.

### Conclusion

- A single reference genome is **NOT** enough to represent the genetic diversity within a species or clade
- PacBio SMRT Sequencing provides a workflow for characterizing the pan-genome of crop species, including de novo assembly, isoform sequencing for exploring transcriptome diversity, and structural variation detection

### References


Acknowledgements

The authors would like to thank everyone associated with the PacBio blog, 15 Jan 2018.