

## Harnessing Genomics to Exploit Maize Genetic Diversity for Crop Improvement

Alex Brohammer



As one of the most widely grown crops around the world, maize (*Zea mays ssp. mays*) provides calories as a staple food source, and has contributed to our current understanding of genetics as a model organism. Maize has undergone a remarkable transformation from its wild progenitor, teosinte (*Zea mays ssp. parviglumis*) to the high-yielding elite varieties grown in farmer's fields today. The steady increases in maize yield over the past century have been due in large part to advances in genetics via plant breeding and the deployment of hybrid varieties as well as improved agronomic practices

and biotechnology (Troyer, 2006). The success of the hybrid seed industry is contingent upon the ability to harness heterosis, which refers to the phenomenon where hybrids display increased levels of growth, survival, or fitness relative to their inbred parents. Aside from heterosis, much of the success of current germplasm is due to greater tolerance to stresses and improved nitrogen uptake and use efficiency (Duvick, 1999). These gains would not have been possible without the rich amount of genetic diversity present throughout the maize genome.

Genetic diversity can arise from a number of sources including sequence level variation, transposable elements, structural variation, epigenetics, and quantitative expression differences. In the past, studies focused on small-scale sequence variation (single nucleotide polymorphisms and small insertions/deletions) in the context of a single reference sequence. However, using array comparative genomic hybridization (aCGH), Springer et al., found evidence for extensive copy number and presence/absence variation throughout the genome (Springer et al., 2009). Building on this work, Lai and colleagues used whole-genome resequencing to characterize variation beyond the context of the

reference genome assembly in a set of 6 elite inbred lines important in Chinese breeding programs (Lai et al., 2010). This study revealed several hundred genic sequences present or absent across the subset of lines and displayed patterns of heterotic group specificity. Transcriptome sequencing has also been used to investigate differences in gene content between maize lines. Using RNA-seq of seedling tissue across 21 diverse inbred lines, 1,321 non-B73 loci were identified, of which a portion again showed heterotic group specificity (Hansey et al., 2011). In an expanded panel of 503 lines, nearly 9,000 non-reference loci were identified (Hirsch et al., 2014).

Despite these recent advances, there still remain a number of unanswered questions. Using a large dataset consisting of both RNA-seq and resequencing data across 35 diverse inbred lines and their hybrid progeny as well as extensive phenotypic data, the proposed research aims to address (1) different types of variation present in the dataset at both the genome and transcriptome levels, (2) the predictive nature of the transcriptome to the maize pan-genome, and (3) the relationship between genome and transcriptome diversity observed heterosis for a series of traits throughout development.

1. Duvick DN. 1999. Heterosis: feeding people and protecting natural resources. *In Genetics and Exploitation of Heterosis in Crops*, ed. JG Coors, S Pandey, pp. 19–30. Madison, WI: Am. Soc. Agron., Crop Sci. Soc. Am., and Soil Sci. Soc. Am
2. Hansey CN, Vaillancourt B, Sekhon RS, de Leon N, Kaeppler SM, et al. 2012. Maize (*Zea mays* L.) genome diversity as revealed by RNA-sequencing. *PLoS ONE* 7:e33071
3. Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Penagaricano F, Lindquist E, Pedraza MA, Barry K, et al. 2014. Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* 26:121-135.
4. Lai J, Li R, Xu X, Jin W, Xu M, et al. 2010. Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat. Genet.* 42:1027–30
5. Springer NM, Ying K, Fu Y, Ji T, Yeh CT, et al. 2009. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet.* 5:e1000734
6. Troyer AF. Adaptedness and Heterosis in Corn and Mule Hybrids. *Crop Science* 2006, 46:528-543.