

The dynamics of selection and adaptive diversity in modern maize inbred lines

Kate Crosby^{*,1}, Justin Gerke[§], Oscar “Howie” Smith[†] and Jeffrey Ross-Ibarra^{‡,1}

^{*}Author one affiliation, [§]Author two affiliation, [†]Author three affiliation, [‡]Author four affiliation

ABSTRACT

blah blah

KEYWORDS maize; inbreds; selection; adaptive diversity

Introduction

Maize is a staple crop in the United States making up X agriculture acres and producing X weight in animal feed. Maize yields (a fitness proxy) in the United States have increased in terms of total numbers since the late 1800s, but have largely slowed in terms of growth rate since the influx of new adaptive material in the mid 1950s to early 1970s. As growth rate in fitness is generally correlated with the amount of available adaptive genetic diversity Total decrease in heterozygosity amongst and within heterotic breeding pools (cite Gerke - maybe give more of an example here) of modern maize, suggests that genetic diversity is lower than it once was. If much of the diversity that has been lost is adaptive, then present day lines would have less variation available to respond to various selection regimes imposed by breeders. If diversity has indeed been lost over time, and has not been recouped by the introgression of new material or mutational input, then what is the composition of this lost genetic diversity? Previous studies suggest that *find lit. on this ... note in gerke paper that haplotypes fixed (presumably under selection) not found in lines derived from those cycles, suggesting perhaps that adaptive haps have been lost.* If the loss is largely composed of deleterious or neutral alleles, then the loss is not of concern to farmers and breeders. If this loss contains even a small amount of beneficial/adaptive alleles, this would mean that the efficacy of selection on modern maize lines would suffer as time went on. This is because, save for the inputs of novel variation in the 1950s when heterotic breeding pools were originally formed - there has not been much in terms of inputs from novel maize

material save for... *cite tropical introgression for blight stuff*.

The “inbred-hybrid” selection method for propagating and distributing maize hybrids has been in use for over 100 years in many breeding platforms and programs in the United States. This method selects on inbred “combining ability” (under a given selection program) to produce a vigorous and viable hybrid, which is then grown by farmers. The loss of heterosis or “hybrid breakdown” then occurs in the next generation, with a loss in fitness/yield. Phenotypic selection by breeders using mapping populations was a common practice prior to the advent of molecular markers. While past studies have found signatures of selection in modern maize: list (cite,cite,cite) these studies have largely fallen short in finding signatures of soft selection and the polygenic basis of likely many traits under the influence of selection. Large effect quantitative trait loci (QTL) and SNPs at these loci are generally markers that are consistently observed to be associated with traits across or within environments. However, the genetic basis of most adaptive traits is likely to be polygenic in nature (cite examples Doebley, etc Rockman review).

We focus narrowly on an inbred line’s year release date to more accurately pinpoint these signatures of selection through time (compare and contrast with Joost) within a heterotic group *explain the concept of heterotic group or breeding pool above. One part of our study builds on the*

Materials and Methods

We obtained the release dates of over 3000 public maize inbred lines, using the USDA database, Gerdes and Tracy, original line release sheets, or communication with maize breeding experts.

Statistical Analysis

To isolate signals of selection and adaptation within each heterotic group of maize inbreds, we used BayEnv 2.0, which estimates Bayes factors. β broadly approximates the parameter Ns according to a linear model according to Bayes Factors *provide a*

brief description of Bayes factors and what high vs. low values indicate . We estimated the neutral covariance matrix of a group of 10,000 (presumably unlinked) markers by choosing every 25th GBS marker by chromosome. Given that there was approximately *give centimorgan distance here per physical map distance* these markers are at the very least not cis-linked, which is key to estimating the covariance matrix. The neutral covariance matrix is used to estimate the null model of evolution that can be expected across environments, in this case temporal environments. Bayes Factors that deviated or fell outside the n

First we grouped inbred lines according to their broad heterotic group using principal components analysis (PCA) using the approach described in PCA using flashpca *cite plos one G Abraham* . Maize inbred lines within a broad heterotic group were then dated by year of release, any inbred line without a year of release affiliated with it was discarded from further analysis.

Our second approach is described in *Berg and Coop 2014*

Our final approach used a more recent version of *Nielsen* of finding the signature of positive selection using SWeeD

Data Availability

Data for GBS 2.7 are available on PANZEA. *add linke* All scripts are available on github (give directory).

Results and Discussion

A. Identification of heterotic groups

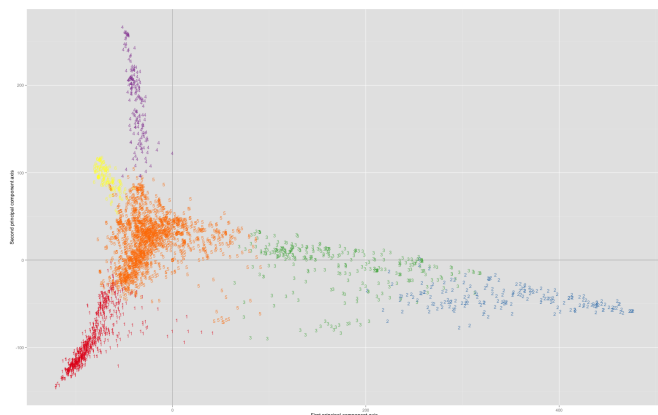


Figure 1 Six clusters representing the broad heterotic groups of modern maize from the Ames panel using GBS 2.7 aligned to version 3 of the B73 genome

Acknowledgements

Financial support for this work came from NSF grant IOS-1238014, USDA grant 2009-65300-05668, and DuPont Pioneer. We'd like to thank DuPont Pioneer as well for their kind support gathering information about years of release. We thank Graham Coop, James Holland, Edward Buckler, X and X reviewers for helpful discussion.