The Genomic Basis for Short-Term Evolution of Environmental Adaptation in Maize

Randall J. Wisser,*,[1] Zhou Fang,[†,2] James B. Holland,[†,‡] Juliana E. C. Teixeira,*,[3] John Dougherty,*,[§] Teclemariam Weldekidan,* Natalia de Leon, Sherry Flint-Garcia,[††,‡‡] Nick Lauter,[§§,***] Seth C. Murray,[†††] Wenwei Xu,[‡‡‡] and Arnel Hallauer[§§§]**

*Department of Plant and Soil Sciences, University of Delaware, Newark, Delaware 19716, [†]Department of Crop and Soil Sciences, North Carolina State University, Raleigh, North Carolina 27695, [‡]US Department of Agriculture-Agricultural Research Service, Raleigh, North Carolina 27695, [§]Center for Bioinformatics and Computational Biology, University of Delaware, Newark, Delaware 19714, **Department of Agronomy, University of Wisconsin, Madison, Wisconsin 53706, [††]US Department of Agriculture-Agricultural Research Service, Columbia, Missouri 65211, [‡‡]Division of Plant Sciences, University of Missouri, Columbia, Missouri 65211, [§§]US Department of Agriculture-Agricultural Research Service, Ames, Iowa 50011, ***Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa 50011, [†††]Department of Soil and Crop Sciences, Texas A&M University, College Station, Texas 77843, [‡‡‡]Agricultural Research and Extension Center, Texas A&M AgriLife Research, Lubbock, Texas 79403, [§§§]Department of Agronomy, Iowa State University, Ames, Iowa 50011

ORCID IDs: 0000-0003-1075-0115 (R.J.W.); 0000-0002-4341-9675 (J.B.H.); 0000-0002-6862-3031 (J.E.C.T.); 0000-0001-7867-9058 (N.d.L.); 0000-0003-4156-5318 (S.F.-G.); 0000-0002-3127-8609 (N.L.); 0000-0002-2960-8226 (S.C.M.)

ABSTRACT Understanding the evolutionary capacity of populations to adapt to novel environments is one of the major pursuits in genetics. Moreover, for plant breeding, maladaptation is the foremost barrier to capitalizing on intraspecific variation in order to develop new breeds for future climate scenarios in agriculture. Using a unique study design, we simultaneously dissected the population and quantitative genomic basis of short-term evolution in a tropical landrace of maize that was translocated to a temperate environment and phenotypically selected for adaptation in flowering time phenology. Underlying 10 generations of directional selection, which resulted in a 26-day mean decrease in female-flowering time, 60% of the heritable variation mapped to 14% of the genome, where, overall, alleles shifted in frequency beyond the boundaries of genetic drift in the expected direction given their flowering time effects. However, clustering these non-neutral alleles based on their profiles of frequency change revealed transient shifts underpinning a transition in genotype–phenotype relationships across generations. This was distinguished by initial reductions in the frequencies of few relatively large positive effect alleles and subsequent enrichment of many rare negative effect alleles, some of which appear to represent allelic series. With these genomic shifts, the population reached an adapted state while retaining 99% of the standing molecular marker variation in the founding population. Robust selection and association mapping tests highlighted several key genes driving the phenotypic response to selection. Our results reveal the evolutionary dynamics of a finite polygenic architecture conditioning a capacity for rapid environmental adaptation in maize.

KEYWORDS recurrent selection; flowering time; genetic diversity; plant breeding; agriculture; climate change

AFTER ~150 years of progress toward understanding evolution—since *The Origin of Species* (Darwin 1859)—burgeoning experimental results fueled by advances in genomic technology are shedding light on still unresolved questions about the nature of phenotypic change, including: the impact of mutation (*e.g.*, Levy *et al.* 2015) and standing variation (*e.g.*, Burke *et al.* 2010; Jones *et al.* 2012); the role of epistasis (*e.g.*, Tenaillon *et al.* 2012); and the relationship between natural and artificial selection (*e.g.*, Chan *et al.* 2012). A key question, especially in the face of biological invasions and climate change, is how genomes confer and constrain the capacity for organisms to adapt to new environments (Orr 2005). Genetic dissection of experimentally evolved populations is a tractable framework for elucidating adaptive evolution

doi: https://doi.org/10.1534/genetics.119.302780
Manuscript received December 4, 2018; accepted for publication October 4, 2019; published Early Online October 15, 2019.
Available freely online through the author-supported open access option.
Supplemental material available at figshare: https://doi.org/10.25386/genetics.9936284.
[1]Corresponding author: Department of Plant and Soil Sciences, University of Delaware, 531 S. College Ave., 152 Townsend Hall, Newark, DE 19716. E-mail: rjw@udel.edu
[2]Present address: Syngenta, Research Triangle Park, NC 27709.
[3]Present address: Embrapa Agricultural Informatics, Campinas, SP 13083-886, Brazil.

since the experimenter can control selection and mating in particular environmental settings (Barrick and Lenski 2013; Schlötterer *et al.* 2015). As a new extension to this framework, we implemented an efficient study design for dual inference about the population and quantitative genomic basis of phenotypic evolution (Wisser *et al.* 2011). This was used to investigate the response to a decade of directional phenotypic selection for tropical-to-temperate adaptation in maize—a model species for plant genetics and a crop of global importance.

Genomic Basis of Response to Phenotypic Selection

The rate and history of mutations, the numbers and positions of functional variants, the distribution of allele effects, and the modes of gene action are among the genetic factors that shape the response to selection and influence the maintenance of phenotypic and genetic variability (Barton and Keightley 2002).

Considering theoretical population and quantitative genetic expectations for the response to directional selection, alleles at one or few loci with large effects on a selected trait should rapidly change in frequency, resulting in a corresponding phenotypic response (Falconer and Mackay 1996). As these alleles approach fixation, genetic variance is reduced and the response diminishes. Sustained responses may be attributed to standing polygenic variation, new mutations, epistatic interactions, or heritable epigenetic effects. For polygenic traits controlled by numerous loci of small effects, modeled at the extreme of infinite loci (Barton *et al.* 2017; Fisher 1918), responses to selection can arise from subtle changes in allele frequencies across many loci. Consequently, allelic variation is retained and the causal-genic variance is expected to undergo negligible change. However, directional selection also creates negative disequilibrium covariance between allele effects across loci, resulting in temporary reductions in genetic variance for the trait under selection, a phenomenon referred to as the Bulmer effect (Bulmer 1971; Walsh and Lynch 2018). Qualitatively similar expectations arise under a so-called finite polygenic architecture where tens or more loci with allele effects of varying magnitudes are at play (Chevalet 1994; Fernando *et al.* 1994; Turelli and Barton 1994).

An empirical understanding about the genetics of adaptation has been advanced through experimental population and quantitative genetic approaches (Savolainen *et al.* 2013). The relative importance of genes with major and minor effects varies among traits, populations and species. At one extreme, relatively rapid or dramatic phenotypic changes have resulted from a few alleles at loci with large effects on traits such as flowering behavior (Lowry and Willis 2010) and toxin resistance (Baxter *et al.* 2011). In contrast, other dramatic shifts in adaptive phenotypes have been ascribed to a polygenic architecture (Burke *et al.* 2010; Berg and Coop 2014). Drawing a clear line of distinction between the two is not straightforward and is partially confounded by differences in experimental systems and their statistical power, but adaptation from a mixture of genes with major and minor effects have been reported (Levy *et al.* 2015). Moreover, "evolve-and-resequence" studies have exposed unforeseen outcomes in the genomic changes underlying phenotypic evolution, including unique patterns of allele frequency change and the maintenance of molecular genetic diversity (Burke and Long 2012), the direct causes of which are unresolved.

Phenological Adaptation in Maize

Pivotal to adaptation and productivity in crop species is synchrony between the growing season and flowering time (Jung and Müller 2009). Numerous studies have investigated the genetic architecture of natural variation in flowering time for maize using a variety of methods, including genic analysis, linkage and association mapping, ecogeographical genetics and historical genetic analysis. Emerging from this body of literature is a consensus that allele effects dispersed across a finite polygenic architecture capture the major proportion of genotypic variation in flowering time (*e.g.*, Chardon *et al.* 2004; Buckler *et al.* 2009; Li *et al.* 2016), and that certain flowering time genes—*Vgt1* (Salvi *et al.* 2007; Ducrocq *et al.* 2008), *ZmCCT10* (Hung *et al.* 2012; Yang *et al.* 2013), *ZmCCT9* (Huang *et al.* 2018), and *ZCN8* (Guo *et al.* 2018)—appear to have been instrumental to the postdomestication spread of maize from its tropical origin to many different environments. However, with one exception (Durand *et al.* 2015), to our knowledge, the genomic basis of this adaptive trait has not been investigated in experimentally evolved populations. This could fill gaps in knowledge about the evolution of adaptation and lay a foundation to innovate breeding methods for rapidly adapting populations to new environments.

In this study, we investigated the genomic basis of adaptation from a distinct vantage point, where the entire period of evolution to an adapted state was captured in a single, multigenerational population—"Hallauer's Tusón" (Teixeira *et al.* 2015). Selection was initiated within an admixed founder population formed by intermating separate seed bank populations of Tusón, a landrace historically adapted to lowland tropical environments (Goodman and Brown 1988). Remarkably, 10 generations of directional phenotypic selection for early female-flowering time, with secondary selection for other traits in a temperate U.S. environment (Ames, IA; 42.03° N latitude), recapitulated the temperate-adapted state of maize achieved by early farmers; this is thought to have occurred over the course of thousands of years (Swarts *et al.* 2017), albeit by less deliberate breeding methodology. A study design, in which families derived from genotyped individuals sampled across generations were phenotypically evaluated, allowed us to employ methods for dissecting both the population and quantitative genomic basis of phenotypic evolution (Figure 1). Because tropical maize is a rich resource of potentially useful genetic diversity that has largely been under used by maize breeders in temperate environments

(Goodman 1998), findings from this study can guide future maize breeding for climate change and address fundamental questions about the genomic basis of environmental adaptation in plants.

Materials and Methods

Front matter

Unless otherwise noted, data analysis was performed using *R* (R Core Team 2016); *R* packages are cited accordingly. The following abbreviations are used: AFPC (allele frequency profile cluster); BLUEs (best linear unbiased estimates); FDR (false discovery rate); FITR (frequency increment test with reference loci); GWA (genome-wide association); LD (linkage disequilibrium); SIM (simulation test statistic).

Plant material

The subject of this study was Hallauer's Tusón, a multigenerational population of maize derived from a landrace of tropical origins that was subjected to 10 generations of phenotypic truncation selection for early female-flowering time in a temperate environment (Ames, IA; 42.03° N latitude) (Teixeira *et al.* 2015; Hallauer and Carena 2016). Figure 1 depicts the breeding scheme for Hallauer's Tusón and our study design.

As described by Hallauer and Carena (2016), the base population (g_0) used to initiate selection was produced in Iowa by isolated, open pollination (no intentional selection) among multiple seed-bank accessions of the Tusón landrace sampled from different countries; however, the origins of some of these accessions remain unclear. It should be noted that maize is a monoecious species with female and male organs on separate parts of the same plant, such that open pollination includes the possibility of self-fertilization. Using the base population, selection ensued where $\approx 8000 - 10,000$ plants were grown in isolation and allowed to intermate at random, which again included the possibility of selfing. During flowering, $300 - 500$ of the earliest female-flowering individuals (based on their silk-emergence phenotype), secondarily selected for other traits, were tagged and later harvested. An equal number of seeds per ear were mixed to form the subsequent generation of $\approx 8000 - 10,000$ individuals for selection. This recurrent selection scheme was applied for 10 generations (from 1995 to 2004) until the population was deemed phenologically adapted by comparison with other temperate adapted maize.

Phenotype data

Previously, Teixeira *et al.* (2015) phenotypically evaluated 297 self-pollinated ($S_{0:1}$) families derived from the even numbered generations of Hallauer's Tusón ($g_0 : n = 18; g_2 : n = 56; g_4 : n = 56; g_6 : n = 56; g_8 : n = 56$ and $g_{10} : n = 55$) for two years at multiple locations in North America, including the Iowa location where Hallauer's Tusón was originally selected. Under the design depicted in Figure 1, the present study combined the available phenotypic data for female-flowering time measured in the selection environment (Ames, IA) and a highly correlated environment (Newark, DE) with new genotype data.

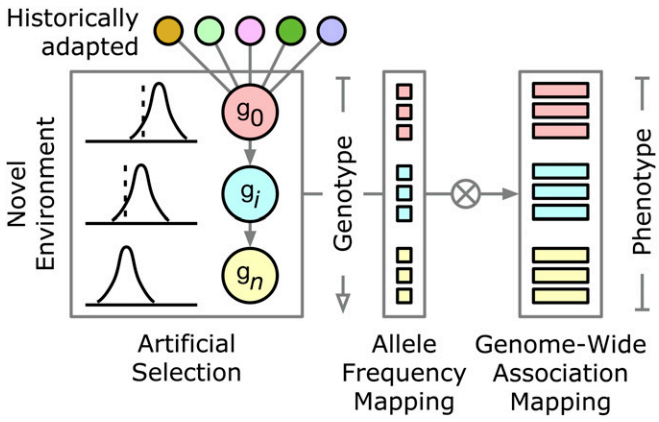

Figure 1 Study design adapted from Wisser *et al.* (2011). Filled circles represent populations, filled squares represent individuals, and filled rectangles represent families. Multiple populations of a historically adapted landrace of maize, sampled from different tropical regions, were randomly intermated to form g_0. Artificial truncation selection for early female-flowering time was performed in a novel temperate environment to which the source populations had not previously been exposed. Individuals were sampled from among the generations of selection, and genotyped. Allele frequency mapping was applied along the generational axis to characterize the population genetic basis of the phenotypic response to selection. The genotyped individuals were self-pollinated to generate families that were evaluated in a common garden experiment. Using the parental genotype data and corresponding family mean phenotype data, association mapping was applied to characterize the quantitative genetic basis of variation in the selected trait.

Genotype data

DNA was isolated from the 297 parents (noninbred) of the $S_{0:1}$ families that were evaluated phenotypically, plus an additional 90 random individuals from g_0 that were not phenotyped (this was done to provide a larger sample size of g_0 for more reliable genomic inference). Lyophilized leaf tissue was pulverized using a Geno/Grinder 2000 and extracted with the DNeasy 96 Plant Kit (Qiagen). Genotyping with Illumina's MaizeSNP50 Beadchip (Ganal *et al.* 2011) was performed by DNA LandMarks (Québec, Canada), producing genotype data at 56,110 SNP sites with an average of 2.8% missing data per sample (min: 0.2%; max: 20.9%). Additional genotype data were produced for variant sites upstream of *ZmCCT10* (Zm00001d024909) (Supplemental Material, Supplemental Methods), including a presence–absence causal variant for photoperiod sensitivity (Hung *et al.* 2012; Yang *et al.* 2013). See Supplemental Methods for genotypic data quality control (Table S1) and projection of markers onto the consensus linkage map for a maize nested association mapping population (McMullen *et al.* 2009).

Analysis of genetic diversity

Specific subsets of markers were used in examining different aspects of genetic diversity (File S1 and Table S2). Some subsets use marker names prefixed with "PZA" or "PZE,"

which have the lowest ascertainment bias among the MaizeSNP50 SNPs (Frascaroli *et al.* 2013).

Summary statistics of genetic diversity were computed with *hierfstat* v. 0.04–14 (Goudet 2005) to calculate H_O (average observed heterozygosity within generations), $\hat{H}_S$ (average expected heterozygosity within generations) and $\hat{H}_{T'}$ (average expected heterozygosity for the total population), as well as $\hat{F}_{IS}$ (average inbreeding coefficient within generations), $\hat{F}_{ST'}$ (average differentiation between generations), and $\hat{F}_{IT'}$ (average inbreeding coefficient for the total population) according to Nei (1986)—this corresponds to the sample-level scope of inference. *HardyWeinberg* v. 1.5.5 (Graffelman 2015) was used to perform exact tests for Hardy–Weinberg equilibrium.

The relationship of Hallauer's Tusón with maize more broadly was assessed using 934 samples representative of global maize germplasm for which genotype data for the same markers was available (Ganal *et al.* 2011). A two-dimensional projection of relationships among samples was computed using *PHATE* (Potential of Heat-diffusion for Affinity-based Trajectory Embedding) (Moon *et al.* 2017) implemented in *phateR* v. 0.2.7, using default settings with a precomputed simple matching distance matrix as input for the knn.dist.method.

Results from *PHATE* using only Hallauer's Tusón suggested some structure was present among samples from g_0 but not other generations. Therefore, *STRUCTURE* (Pritchard *et al.* 2000) was used to examine subpopulations among g_0 samples assuming admixture and independent allele frequencies (Table S3; see Supplemental Methods for details). The ΔK method informed our selection of K (Evanno *et al.* 2005). Although the exact accessions of Tusón used to form g_0 remain unclear, several were used (Hallauer and Carena 2016). Therefore, we ignored the stronger signal of ΔK suggesting $K = 2$ and chose the next highest peak at $K = 6$ (Figure S1). To examine whether any one subpopulation was favored in the initial generations of selection, *STRUCTURE* was also used to ascertain the g_0 ancestry present in g_2 samples, assuming the aforementioned subpopulation classification for individuals in g_0. A paired t-test was used to test the difference between the proportion of individuals per subpopulation in g_0 (Table S3) and the average per individual admixture proportion estimated by *STRUCTURE* for g_2 (Table S4).

LD (Hill and Robertson 1968), measured as r^2, was computed with *genetics* v. 1.3.8.1 (Warnes *et al.* 2013). The structure of LD between chromosomes was characterized per generation using low-ascertainment biased markers with no missing data and a minimum allele count of 12 in the given generation (Table S2). The structure of LD within chromosomes was similarly characterized, but with a larger number of markers (Table S2) and examined at different intervals of genetic distance: (0,1],(1,5],(5,10],(10,50],>50. In addition, r^2 was computed among generations (all samples) between sequential pairs of markers per chromosome. Using these latter estimates of LD, *lokern* 1.1.8 (Herrmann 2016) was used to perform kernel regression of the r^2 values as a function of the midpoint basepair coordinate between each pair of markers. A plot of pairwise r^2 between markers flanking the *ZmCCT10* associated causal site for photoperiodism was made using *LDheatmap* v. 0.99.2 (Shin *et al.* 2006).

Population genetic analysis

Allele frequency mapping: Three statistical tests were used to detect markers with nonrandom patterns in allele frequency change across generations: (i) a customized whole genome SIM test for departures from genetic drift; (ii) the Bayenv test for robust correlations between allele frequencies and generations (Coop *et al.* 2010); and (iii) the FITR for robust departures from genetic drift (Nishino 2013).

A detailed description of the SIM test can be found in Supplemental Methods. Briefly, by customizing simuPOP (Peng and Kimmel 2005), a simulator was constructed to generate in silico genomes (genome-wide genotypes) for individuals constituting a population that undergoes breeding according to the design for Hallauer's Tusón (with random selection to model genetic drift) and sampling according to our study design (to account for sampling variance). The simulator used the fixed g_0 genotype matrix, fixed recombination rates estimated from projection onto the genetic map (Supplemental Methods) and the fixed STRUCTURE matrix (Table S3) to initially generate $10,000$ random in silico genomes (derived based on the structured g_0 sample data) from which simulated breeding ensued. At each marker across the genome, the probability of the observed sample allele frequency change was computed relative to the expected distribution created from $10,000$ replicates of simulation. Marker p-values were adjusted for multiple testing (Benjamini and Hochberg 1995), and those with a 1% FDR are referred to as SIM^+ markers, while the remaining markers are referred to as SIM^-.

Bayenv 2.0 was used to identify robust correlations between allele frequencies (response variable) and generations (explanatory variable). Bayenv was originally designed to test for correlations with an environmental variable; here, generation numbers (standardized) were used instead. The covariance matrix used to model the background expectation of allele frequency change was estimated from the low-ascertainment biased SNPs (Table S2). Markers with the top 1% Bayes factor values were considered robust outliers, which we refer to as $Bayen^+$ markers.

The FITR statistic is conditioned on the variance in allele frequency change estimated from neutral markers in the sample (Nishino 2013). For this, we used SIM^- markers declared at a 10% FDR. To minimize bias from using a single set of reference markers, we modified the test by bootstrap resampling SIM^- markers and computing the proportion of times in which each marker was significant (at a 1% FDR) among $10,000$ bootstrap samples. In each bootstrap sample, the reference data comprised 1% of the markers ($n = 444$) sampled at random and without replacement. The test is not

definable for fixed sites, yet the alleles at markers used for testing were not always observed in each generation. Therefore, alleles with an observed frequency of zero in a particular generation were set to $1/2S_g$, where S_g is the number of individuals in the g^{th} generation. Markers with a 1% FDR in at least 75% of the bootstrap sample tests are henceforth referred to as FITR$^+$.

Localizing footprints of selection: Chromosomal regions with a local footprint of selection (*i.e.*, deviation from genetic drift across a segment of the genome) were delimited using chromosome-specific kernel regression functions of the $-log10(q)$ values from the SIM test on the physical coordinates of markers. To obtain a definable $-log10(q)$ input value for markers where the SIM test p-value equaled 0 (where the observed data fell outside the limits of the null distribution of simulated drift), $p = 0$ was set to 0.00057, which was half the minimum p-value among all SIM tests. Regions along each chromosome where the kernel regression line surpassed a threshold of $-log10(q = 0.05)$ were considered local footprints of selection, and are henceforth referred to as $\text{SIM}^+_{\text{regions}}$.

Characterizing features of allele frequency change: Divisive analysis of hierarchical clustering (Kaufman and Rousseeuw 1990) was performed with *cluster* v. 2.0.7 (Maechler *et al.* 2018) in order to group SIM$^+$ markers with similar profiles of allele frequency change for the minor allele in g_0. The number of clusters was determined using *clusGap* based on the *Tibs2001SEmax* criterion (Tibshirani *et al.* 2001).

Additional data summaries, using the minor allele in g_0 as the reference allele, were used to compare features of allele frequency change between SIM$^-$ and SIM$^+$ markers, including: (a) the slope and intercept from regression of allele frequency change on generations; (b) the mean absolute change in allele frequency among the highest to lowest ranking changes in frequency per marker across sequential pairs of generations (the largest amount of change between a given pair of generations was assigned a ranking of 1; the least amount of change was assigned a ranking of 5); (c) the distribution of the longest run (*i.e.*, number of generations) of positive and negative monotonic change in allele frequency across sequential generations; and (d) the rank distribution for the amount of allele frequency change across generation pairs.

Quantitative genetic analyses

Genetic differentiation: Q_{ST}, a measure of the proportion of genetic variance distributed among populations for quantitative traits (Spitze 1993), was used to estimate genetic differentiation in female-flowering time between g_0 and each subsequent generation of Hallauer's Tusón, as well as to examine Q_{ST} relative to the distributions of F_{ST} for SIM$^-$ and SIM$^+$ markers. Knowing that flowering time was under selection, the $Q_{ST} - F_{ST}$ comparison was used to characterize the relationship between population genetic and quantitative trait divergence (Le Corre and Kremer 2012). Following from Spitze (1993):

$$\hat{Q}_{ST} = \hat{\sigma}^2_{GB}/(\hat{\sigma}^2_{GB} + 2\overline{\hat{\sigma}^2_{GB}}); \quad (1)$$

where $\hat{\sigma}^2_{GB}$ and $\overline{\hat{\sigma}^2_{GB}}$ are estimates of the among-generation and average within-generation additive genetic variances, respectively. See the next section and Supplemental Methods for details on the estimation of variance components. For $Q_{ST} - F_{ST}$ comparison, we used the Hudson estimator, F^H_{ST} (Bhatia *et al.* 2013), which is compatible with restricted maximum likelihood estimation of genetic variances used for Q_{ST}, since both estimates correspond to the population-scope of inference in the broad sense.

Partitioning of the genotypic variance: Using *ASReml* v. 3 (Gilmour *et al.* 2009), the following mixed linear model was used to partition the phenotypic variance and decompose the genotypic variance into additive, dominance, and residual genetic variance components:

$$\begin{aligned}\mathbf{y} = \mathbf{X}_m\boldsymbol{\beta} + \mathbf{Z}_E\mathbf{e} + \mathbf{Z}_{I(R*E)}\mathbf{i} + \mathbf{Z}_{F(G)}\mathbf{a} + \mathbf{Z}_{F(G)}\mathbf{d} \\ + \mathbf{Z}_{F(G)}\mathbf{r} + \mathbf{Z}_{F(G)*E}\mathbf{f}{*}\mathbf{e} + \boldsymbol{\varepsilon};\end{aligned} \quad (2)$$

where $\mathbf{y}$ corresponds to the vector of observations (female-flowering time), $\boldsymbol{\beta}$ is the fixed overall mean effect, and $\mathbf{e}, \mathbf{i}, \mathbf{a}, \mathbf{d}, \mathbf{r}, \mathbf{f}{*}\mathbf{e}$ and $\boldsymbol{\varepsilon}$ correspond to the vectors of random environment effects, incomplete block nested in replication within environment effects, additive genetic family effects, dominance genetic family effects, residual genetic family effects, family × environment interaction effects, and residuals, respectively. The effect of replication nested in environments was excluded as it was not significant according to a likelihood ratio test. The respective design matrices $[\mathbf{X}_m, \mathbf{Z}_E, \mathbf{Z}_{I(R*E)}, \mathbf{Z}_{F(G)}$ (this has same structure for $\mathbf{a}, \mathbf{d}$, and $\mathbf{i}), \mathbf{Z}_{F(G)*E}]$ relate observations to their corresponding vectors of effects. The additive, dominance, and residual genetic family effects were assumed to be distributed independently of one another, where: $\mathbf{a} \sim (0, \mathbf{G}\hat{\sigma}^2_a), \mathbf{d} \sim (0, \mathbf{D}\hat{\sigma}^2_d), \mathbf{i} \sim (0, \mathbf{I}\hat{\sigma}^2_r)$. The $\mathbf{G}$ matrix was computed according to VanRaden (2008) and the $\mathbf{D}$ matrix according to Su *et al.* (2012), while $\mathbf{I}$ is an identity matrix.

Variance component estimates from Equation 2 were used to compute heritability in the broad ($\boldsymbol{H}^2$) and narrow ($\boldsymbol{h}^2$) sense on an entry mean-basis according to Holland *et al.* (2003). In addition, extensions of Equation 2 were used to examine the amount of genetic variance in female-flowering time explained by each chromosome and for SIM$^-$ *vs.* SIM$^+$ markers (Supplemental Methods).

Genome-wide association mapping: Kinship-controlled GWA was performed using a standard two-step procedure (Yu *et al.* 2006). First, BLUEs of female-flowering time for the $S_{0:1}$ families were estimated using Equation 2, but replacing additive, dominance, and independent random genetic effects with a single term for families fit as a fixed

effect. Second, using BLUEs as the response variable, markers were tested for trait association using the mixed linear model in *TASSEL* (Bradbury *et al.* 2007) standalone v. 5.2.12, while controlling for the random polygenic background with the aforementioned **G** matrix.

To reduce false-positive associations due to rare genotypes co-occurring with outlier phenotypes, if the sample size of phenotyped lines for a given genotypic class at a marker was less than five, individuals with the corresponding genotypic state (typically the homozygous minor allele class) were set to missing for that marker. The QQ plot of GWA *p*-values is shown in Figure S2. A 10% FDR was used to declare significant trait-marker associations; henceforth, these are referred to as GWA^+ markers.

When estimating additive allele effects, some markers had only two genotypic classes (heterozygous and one homozygous class). The effects of minor variants at these loci were estimated as the difference between the heterozygous class and the homozygous class. For markers with three genotypic classes, the additive effect was uniformly reported as half the difference between the homozygous variant class corresponding to the minor variant in g_0 and the homozygous variant class of the alternative variant.

Using the closest-features program of *BEDOPS* (Neph *et al.* 2012) v. 2.4.15, the flowering time candidate gene [Dataset S8 in Hung *et al.* (2012)] nearest to each GWA^+ marker was determined.

Synthesis map

A graphical map of the maize genome integrating multiple results was created. The map included local linkage disequilibrium estimated by kernel regression, the difference in heterozygosity between generations 0 and 10 ($H_{O g_{10}} - H_{O g_0}$), SIM test results ($-log10$-transformed q-values), Bayenv test results ($log10$-transformed Bayes factor values), FITR test results (bootstrap values >75%), GWA test results ($-log10$-transformed q-values), previously mapped QTL associated with flowering time {flowering time *per se* [Table S3 in Buckler *et al.* (2009)] and photoperiod sensitivity [Dataset S3 in Hung *et al.* (2012)]}, and candidate genes for flowering time [Dataset S8 in Hung *et al.* (2012)].

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. Supplemental methods, tables, and figures are available at figshare: https://doi.org/10.25386/genetics.9936284. Supplemental Tables: Tables S1 and S2 details results from quality control filtering and lists subsets of the genotype data used for analysis; Tables S3 and S4 contain results from *STRUCTURE* analysis; Table S5 summarizes $SIM^+_{regions}$ identified by kernel regression; Table S6 shows chromosome-specific genetic variance component estimates; Table S7 lists the markers detected by GWA; and Table S8 shows the candidate gene for flowering time nearest to each GWA^+ marker. Supplemental Figures: Figure S1 shows results from *STRUCTURE* used to select K; Figure S2 is a QQ plot of observed *vs.* expected *p*-values for GWA tests; Figure S3 is a Venn diagram of SIM^+ markers detected when the SIM test was applied to sequential pairs of generations; Figure S4 summarizes various features of allele frequency change for SIM^- and SIM^+ markers; Figure S5 shows the allele frequency profile clusters and corresponding distributions of additive allele effects for SIM^+ markers; Figure S6 shows the structure of LD within and between chromosomes for pairwise combinations of SIM^- and SIM^+ markers; Figure S7 shows the synthesis map of multiple analysis results. Supplemental Files: File_S1.txt contains a list of the quality control markers, their map locations on B73 AGPv2 and AGPv4 reference assemblies, and the analysis-specific, Table S2 subset to which they belong. File_S2.txt contains summary statistics and test results for each marker, including: allele frequencies per generation for the corresponding minor allele in g_0; observed heterozygosity (H_O) per generation; F^H_{ST} between g_0 and g_{10}; *p*-values and *q*-values for the SIM test; Bayes factor values and correlation statistics for Bayenv, bootstrap values for FITR; *p*-values and *q*-values for GWA; and the estimated additive allele effect. Phenotype and genotype data are available at Dryad: phenotype - https://doi.org/10.5061/dryad.8f64f; genotype - https://doi.org/10.5061/dryad.q573n5tdt. Python code used for genome simulation is available via GitHub: https://github.com/maizeatlas/saegus.

Results

Artificial selection generated a tropical genome with a temperate-adapted phenome

Hallauer's Tusón population was founded by intermating multiple seed bank accessions of a maize landrace historically adapted to tropical environments (Figure 1). Teixeira *et al.* (2015) demonstrated the capacity of this population to become phenologically adapted to a temperate environment within 10 generations of artificial selection, based primarily on selection for early female-flowering time. Here, we found that the population was highly diverse; nearly the entire set (96%) of $\approx 50,000$ SNPs on the MaizeSNP50 chip (Ganal *et al.* 2011) segregated within or among generations (Table 1).

When compared to a global sample of maize, all of the individuals across generations of Hallauer's Tusón clustered with tropical germplasm (Figure 2a), notwithstanding the temperate-adapted phenome of individuals belonging to the later generations (Teixeira *et al.* 2015). Thus, to tackle challenges associated with crop vulnerability through plant breeding, Hallauer's Tusón highlights the adaptive potential of maize landrace populations, and provides a unique source of germplasm likely to contain novel alleles for temperate maize breeding programs.

Retrospective analysis reveals admixture during selection on a structured founder population

Genomic inference combined with knowledge about the population development scheme was used to illuminate the

Table 1 Summary statistics for molecular genetic diversity in Hallauer's Tusón

Generation	Sample size	Proportion polymorphic[a]	H_O[b]	$\hat{H}_S$[c]	$\hat{H}_{T'}$[d]	$\hat{F}_{IS}$[e]	HWD$^+$[f]	HWD$^+$: $H_O \leq \hat{H}_S$[g]
0	105	0.897	0.285	0.320	NA	0.111	0.256	0.784
2	56	0.893	0.323	0.315	NA	−0.026	0.071	0.400
4	55	0.902	0.325	0.314	NA	−0.037	0.067	0.373
6	54	0.914	0.315	0.320	NA	0.016	0.072	0.475
8	56	0.923	0.328	0.325	NA	−0.006	0.072	0.440
10	55	0.929	0.333	0.327	NA	−0.017	0.069	0.411
0 − 10	381	0.963	0.318	0.320	0.305	0.021	0.262	0.704

[a] Results in the column are based on 49,477 markers (Table S1), while all remaining results are based on 44,445 markers with a minor variant count ≥ 12 among all samples.
[b] Average observed heterozygosity within generations. For all samples, this corresponds to the average among generations.
[c] Average expected heterozygosity within generations. For all samples, this corresponds to the average among generations.
[d] Average expected heterozygosity for the total population.
[e] Average inbreeding coefficient within generations. For all samples, this is $\hat{F}_{IT'}$, the average inbreeding coefficient for the total population.
[f] The proportion of markers that significantly deviated from Hardy–Weinberg equilibrium determined at a 5% FDR.
[g] The proportion of HWD$^+$ markers where H_O was less than $\hat{H}_S$.

breeding history of Hallauer's Tusón. Although generations showed some SNP-based differentiation (mean $\hat{F}_{ST'} = 0.014$), as expected for progeny generated by random mating among selected individuals, overall inbreeding was minimal ($\hat{F}_{IT'} = 0.021$). However, on a per generation basis, $\hat{F}_{IS}$ was noticeably higher in g_0, where 26% of the markers significantly deviated from Hardy–Weinberg equilibrium, with 78% of those deviations being due to an excess of homozygotes (Table 1). This suggested a Wahlund effect (Hartl and Clark 2007) from sampling geographically separated accessions that remained after the intermating step to form g_0, but subsequent generations showed evidence of random mating during selection (Table 1).

STRUCTURE analysis also indicated that individuals in g_0 had formed largely by intermating within separate founder accessions and hybridization between specific pairs of accessions (Figure 2b and Table S3). These findings are congruent with the way in which g_0 was bred, whereby the original Tusón accessions were planted in adjoining blocks and allowed to open-pollinate, which would have favored mating within and between pairs of subpopulations. Although selection could potentially favor specific subpopulations under these conditions, the genomic ancestral composition for g_2 individuals showed admixture profiles that were proportional to that of the subpopulation sizes in g_0 (Table S4). Thus, randomized bulking and planting of seed between each generation of artificial selection minimized subsequent inbreeding and population structure during selection.

Differentiation across a fraction of the genome potentiated strong phenotypic change

A decade of directional phenotypic selection, resulting in an overall mean decrease of 26 days to female-flowering time, caused generations to become strongly differentiated phenotypically, as measured by Q_{ST} (Figure 3). Simulation of neutral allele frequency changes that would occur under the breeding scheme used for phenotypic selection allowed us to identify 6115 of 43,628 (14%) markers with non-neutral allele frequency changes (referred to as SIM$^+$ markers). These markers were widely dispersed across the genome (but with some clusters of linked SIM$^+$ markers, as described later), and were distinguished from SIM$^-$ markers by their increasing levels of $\hat{F}_{ST}^{H}$ across generations relative to g_0 (Figure 3). The identification of a sizeable fraction of genome-wide markers as SIM$^+$ suggested a finite polygenic architecture (*i.e.*, possibly tens to hundreds of loci affecting flowering time) could underlie the phenotypic response to selection. Similarly, based on $Q_{ST} - F_{ST}$ comparisons, the very large increases in $\hat{Q}_{ST}$ could be explained by much smaller levels of $\hat{F}_{ST}$ amplified across a large number of loci.

Population genetic analysis pinpoints shifts in the genetic architecture underlying response to selection

Although $\hat{F}_{ST}$ increased overall across generations at SIM$^+$ markers, changes in the frequencies of alleles at a locus varied mostly among generations. Using the simulator to test for non-neutral allele frequency changes between sequential pairs of generations showed that a majority of marker-specific departures were exclusive to one consecutive pair (Figure S3). This was coincident with the common observation of "bursts" in allele frequency change within a few generations, rather than monotonic changes across all generations (Figure S4, A–C).

Because transient changes in allele frequency were a prominent feature of the genomic response to selection, we used clustering to examine the temporal structure of allele frequencies among SIM$^+$ markers. A total of 15 allele frequency profile clusters (AFPCs) were resolved with a high degree of overall clustering structure (divisive coefficient = 0.98; Figure S5). Minor alleles in g_0 with negative frequency trajectories were captured in AFPCs 1 − 3 comprising ≈ 10% of the SIM$^+$ markers. The remaining g_0 minor alleles, however, were enriched from starting frequencies that spanned the minor-allele frequency spectrum. Across clusters, some notable transitions in allele frequency responses were observed: (i) alleles in AFPC1, which included the photoperiod sensitive allele (*ZmCCT10-s*), substantially reduced in frequency to become rare or removed within the first four generations; (ii) several AFPCs (6, 10, 11, 13, 14, and 15) showed clear increases in allele frequency within the first few generations that were limited thereafter; and (iii) for AFPC4, a dominant cluster comprising 43% of all SIM$^+$ markers, initially rare alleles appreciably increased after g_4.

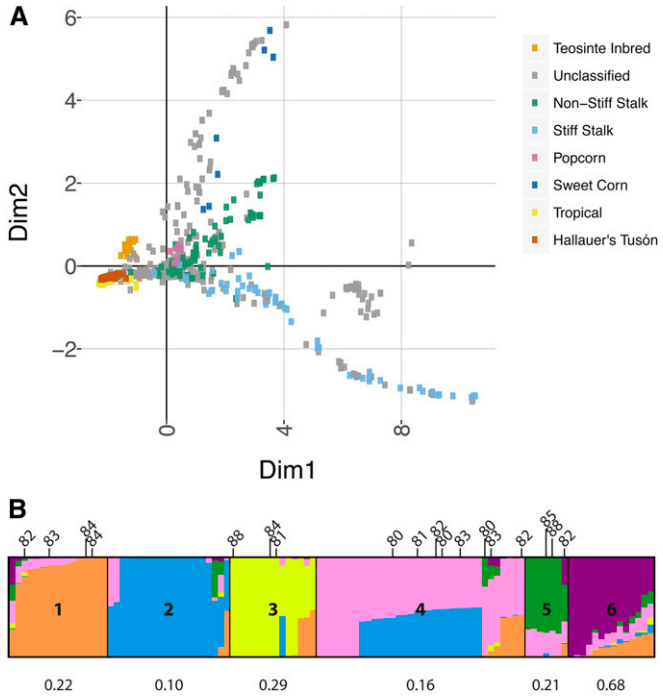

Figure 2 Genetic diversity and population structure in Hallauer's Tusón. (A) 2D *PHATE* plot showing relationships for Hallauer's Tusón and a broad range of maize germplasm. Aside from distinguishing samples from Hallauer's Tusón and teostine inbreds, population structure groups were assigned to samples based on Table S1 in Romay *et al.* (2013) (several samples were unclassified but retained in the analysis). (B) Population structure in g_0. Admixture profiles for six subpopulations are shown. Values at the top of the *STRUCTURE* plot correspond to days to female flowering time for individuals with phenotype data. Subpopulation frequencies of *ZmCCT10-s*, the deletion ("lack-of-insertion") variant associated with photoperiod sensitivity is shown.

There were few instances, like *ZmCCT10-s*, where alleles were purged from the population. The loss of SNP variants in Hallauer's Tusón was actually rare—1.0%($n = 439$ markers) of all SNPs in g_0 were purged during selection (10 of these were SIM$^+$), and this was 3.4 times less than the average proportion of SNPs purged across replicate simulations of neutral frequency change (range : 2.8% − 4.0%). Taken together, population genetic analysis suggested that selected genotype–phenotype relationships temporally shifted across a finite polygenic architecture with a predominant enrichment of initially minor alleles, resulting in the maintenance of genetic variation.

Quantitative genetic analysis contextualizes genome-wide population genetic dynamics

Our experimental design included phenotypic data on self-pollinated families of genotyped individuals sampled across generations evaluated in common environments (Figure 1), permitting estimation of quantitative genetic parameters and interpretation in the context of population genetic results. Underlying a 50-day range for female-flowering time in Hallauer's Tusón, high broad and narrow sense heritabilities ($H^2 = 0.96 \pm 0.01 : h^2 = 0.81 \pm 0.08$) indicated a large fraction of the phenotypic variance could be explained by genotypic effects, and the genotypic variance partitioned into 85% additive and 15% dominance variance with no residual genetic variance remaining. Including an additive-by-additive epistatic relationship matrix did not improve the model fit nor did it explain any of the genotypic variance.

The genetic variance explained by individual chromosomes varied widely (Table S6). For instance, chromosome 10, in which the *ZmCCT10*-associated causal variant is located, was an outlier that accounted for a large proportion (28%) of the additive variance, while chromosome 2 included no additive or dominance variance. This supports and extends the population genetic inference of a finite polygenic architecture, showing variability in the genetic effects across the genome available for selection.

Genome-wide additive allele effects (estimated among families across generations) correlated with average changes in allele frequency per generation ($r = -0.39, p < 2.2\text{e}-16$), whereby, as expected, alleles with negative effects on female-flowering time (contributing to early flowering) tended to have positive slopes in allele frequency change and vice versa (Figure 4; the Bayenv and GWA hits highlighted in the Figure are discussed in the next section). This relationship was largely driven by SIM$^+$ markers, which accounted for 60% of the additive variance (none of the dominance variance)—an excessive enrichment given these constituted ≈ 14% of the SNPs.

Nonlinear changes in the phenotypic mean and additive variance for female-flowering time across generations reflected some of the observed dynamics in allele frequency change (Figure 5). Changes in the mean could be modeled as a cubic function with significant ($\alpha = 0.05$) coefficients [$f(x) = -0.07gen^3 + 1.39gen^2 - 9.64gen + 100.27$], where female-flowering time decreased across all generations but to a larger degree between generations $0-4$ and $8-10$. These generations contained greater numbers of SIM$^+$ markers with larger magnitudes of allele frequency change (Figure S4d). The corresponding change in additive variance could be modeled as a quadratic function of generations with significant coefficients [$f(x) = -14.11gen^2 + 1.06gen + 31.58$], where initially the variance was greatly reduced but later increased across generations $6-10$; by g_{10}, the additive variance exceeded that of g_4. These changes were coincident with initial reductions in the frequencies of few relatively large positive effect alleles (AFPC1) and subsequent enrichment of many rare negative effect alleles (AFPC4) (Figure 5 and Figure S5).

Robust selection and association mapping identify associations with key flowering time genes

The simulation test clearly enriched for markers that differentiated generations (Figure 3), but not all of these are necessarily linked to causal variants underlying female-flowering time selection. For instance, chromosome 2 captured none of the genetic variance in female-flowering time; however, it contained SIM$^+$ markers (Table S6). The model used to simulate breeding events does not include components of potentially important sources of variation, such that

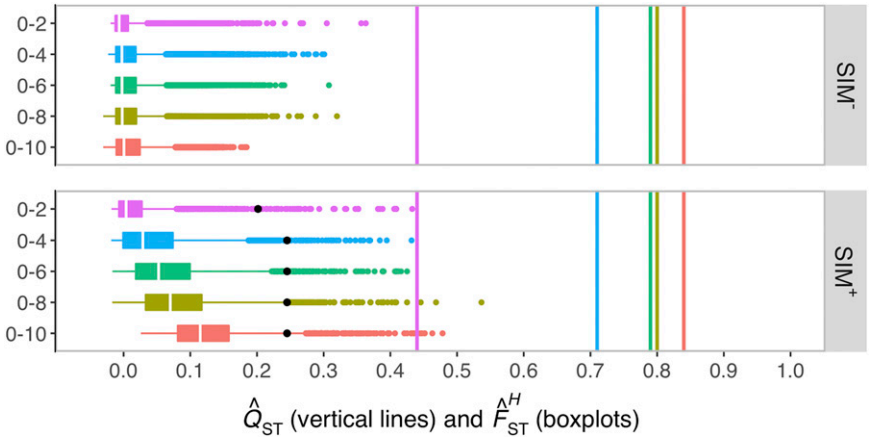

Figure 3 $\hat{Q}_{ST} - \hat{F}^H_{ST}$ comparisons. $\hat{Q}_{ST}$ (vertical lines) for female-flowering time is compared to the distributions of $\hat{F}^H_{ST}$ (boxplots) for SIM$^-$ and SIM$^+$ markers. Color coding is according to the pair of generations that were compared as indicated on the *y*-axis. Black points show $\hat{F}^H_{ST}$ of the ZmCCT10_CACTA marker which reached its maximum value of 0.25 by g_4, the generation by which *ZmCCT10-s* was purged from the population.

departures from the null distribution may also be due to factors such as individual differences in gametic fitness, or the secondary selection that was exerted for other traits (Hallauer and Carena 2016). On the other hand, covariance between causal and neutral allele frequency changes may generate false positives (Coop *et al.* 2010). Although we could not control for deviations due the former, selection and association mapping tests that control for genomic background effects were used to identify markers exhibiting robust changes in allele frequency across generations and robust associations with trait variation among generations, respectively.

Prior to applying these tests, we assessed the transgenerational structure of LD across the genome using subsets of low-ascertainment biased markers with a standardized minimum allele count per generation (Table S2). Linkage disequilibrium was examined for each combination of SIM$^-$ and SIM$^+$ markers. Based on all pairwise estimates of LD, median r^2 showed little variation across generations, and did not exceed 0.04 within chromosomes and 0.01 between chromosomes (data not shown). At increasingly higher percentiles of the r^2 distribution, LD between pairs of SIM$^+$ markers within chromosomes (but not between chromosomes) was elevated relative to other combinations of SIM$^-$ and SIM$^+$ markers (Figure S6). We found that the heightened LD between SIM$^+$ markers was restricted to local linkage blocks, and tended to increase across generations (Figure S6)—a hallmark footprint of selection.

Given that the structure of LD gave rise to nonindependent sets of linked SIM hits, kernel regression was used to delimit 29 SIM$^+_{\text{regions}}$ encompassing 1008 (16%) of the SIM$^+$ markers (Table S5). The Bayenv test, which controls for genome-wide covariance in sample allele frequencies, produced Bayes factor values that were correlated with the SIM test (Spearman's $|\rho| = 0.62$). The top 1% of Bayenv hits were located on all chromosomes and were present in most SIM$^+_{\text{regions}}$, with regions on chromosome 9 being heavily populated with Bayenv$^+$ markers (Figure 6 and Figure S7). Similarly, the FITR test, which is conditioned by variance in allele frequency change estimated from the sample, primarily implicated SIM$^+_{\text{regions}}$ on chromosome 9 as robust outliers.

GWA mapping performed on mean female-flowering time resulted in few genome-wide significant associations, notwithstanding the SNP-based polygenic model that explained essentially all of the genotypic variance. Between 2 and 12 GWA hits were detected across 1–10% FDR thresholds (Table S7). All but one of these showed the expected relationship between the sign for the additive allele effect and slope in frequency change (we note this one marker was on chromosome 2, which explained none of the genetic variance when the whole chromosome was modeled under a polygenic architecture). However, no GWA hits were detected on chromosome 9 nor within any of the SIM$^+_{\text{regions}}$. The top GWA hit was the presence–absence causal variant for *ZmCCT10*-regulated photoperiodism, which was also detected as hits by the SIM and Bayenv tests but not the FITR test. Otherwise, the strength of signals (slopes *vs.* effects) for top hits by selection and association mapping tests tended to differ (Figure 4).

Taken together, robust tests to dissect the genetic basis of the response to selection implicated a number of genes previously associated (causally or as a candidate gene) with variation in flowering time and photoperiodism in maize, several of which are highlighted in Figure 6.

Evidence for multiple local haplotypes underlying the phenotypic response to selection

The transition from selection on common to rare alleles occurred at some of the same regions of the genome. For instance, SIM$^+_{\text{regions}}$ on different chromosomes included SIM$^+$ markers in both AFPC1 and AFPC4, in which allele frequencies showed strong shifts during different periods of selection (Figure 5). Similarly, at the *ZmCCT10* locus, robust associations were detected for SNPs that responded to selection even after the elimination of *ZmCCT10-s*. Pairwise LD between significant markers at the *ZmCCT10* locus indicated two separate haplotypes were responsive to selection (Figure 7). These results reinforce the conclusion of a finite polygenic architecture underlying the response to selection, and extend

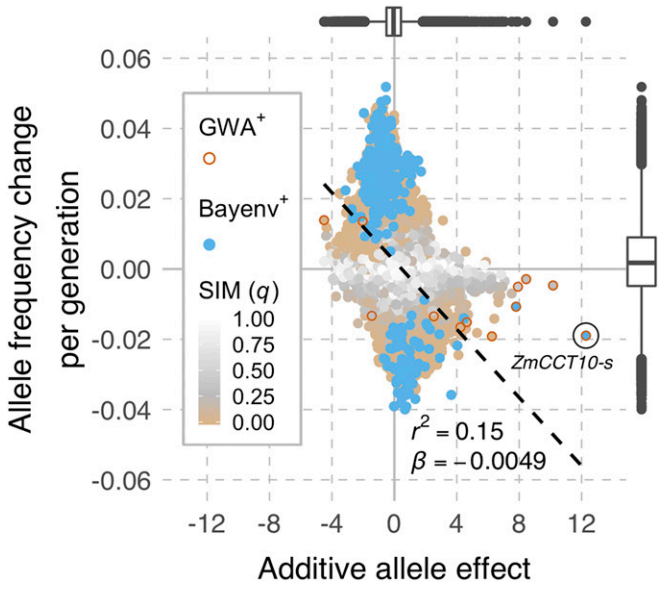

Figure 4 Genome-wide relationship between slopes in allele frequency change and additive allele effects. Each point corresponds to the minor allele in g_0 and indicates the slope in allele frequency change estimated across all generations (*y*-axis) *vs.* the additive allele effect estimated among genotypes from all generations (scaled in days; *x*-axis). Marker effects are estimated separately, and do not account for covariances among loci, such that their combined effects tend to be less than the sum of individual effects. Marginal histograms show their distributions. The coefficient of determination and slope of the relationship are shown. Color coding depicts test statistic results (10% FDR GWA hits; top 1% Bayenv hits; and FDR *q*-values for the SIM test). The *ZmCCT10-s* allele is marked by an open circle.

that to suggest selection on multiple local haplotypes was an important aspect of short-term evolution.

Discussion

Genetic analysis of adaptation in crop species provides a lens into evolution and generates relevant information for plant breeding. Although flowering time phenology has been widely studied in plants (Jung *et al.* 2017), we are aware of no study (in plants) that has dissected the transgenerational genomic basis of adaptive evolution (here, for flowering phenology) in a population translocated to a new environment. We help close this knowledge gap by investigating a tropical landrace of maize that was adapted to a temperate environment across a decade of artificial selection (Teixeira *et al.* 2015; Hallauer and Carena 2016). Using an efficient study design (Figure 1; Wisser *et al.* 2011), we simultaneously elucidated population and quantitative genetic components underlying the 10 generations of selection required for the population to reach a state of phenological adaptation similar to modern temperate maize lines.

The evolutionary capacity of the tropical Tusón landrace to become rapidly adapted to a temperate environment was attributed to a finite polygenic architecture, yet two genomic phases underlying the phenotypic response to selection could be discerned. The first phase, from generations $0-4$, was distinguished by an oligogenic-like architecture, where marked reductions of a relatively small number of moderate-frequency minor variants in g_0 (AFPCs 1 and 2), with relatively large positive effects on flowering time, contributed to an initial strong response to phenotypic selection and a large reduction in genetic variance (Figure 5). Afterward, the genomic basis of the response transitioned to become dominated by the enrichment of a large number of rare-minor variants in g_0 with smaller-sized negative effects on flowering time, leading to a genome-wide increase in heterozygosity (Table 1) and consequent increase in additive variance (Figure 5).

The observed changes in phenotypic mean and variance are similar to expected outcomes theorized for a finite polygenic architecture with additive allele effects (Chevalet 1994). Consistent with an additive genetic model, several AFPCs showed linear trends across all generations reflective of unconditionally (un)favorable alleles in Hallauer's Tusón (Figure S5). However, AFPCs with transient shifts in allele frequency were also detected, such as mid-to-late generational responses and plateaus in allele frequency change, highlighting a context-dependent component of the genetic architecture underlying the response to phenotypic selection. The same pattern of plateauing allele frequencies after an initially strong shift was found by temporal analysis of natural populations of *Drosophila melanogaster* adapted to a novel laboratory environment, which Orozco-terWengel *et al.* (2012) reasoned was due to overdominant or antagonistic pleiotropic effects. It has been demonstrated (mathematically) that the selection coefficient for an additive allele can vary across generations also as a result of changes in background polygenic variance (Chevin and Hospital 2008). The genotypic variance in Hallauer's Tusón partitioned into additive (primarily) and dominance genetic variance with no apparent epistatic genetic variance, but epistatic genetic effects will contribute to the additive genetic variance component in many cases, such that inferences about gene action should not be drawn from variance components estimates (Hill *et al.* 2008; Huang and Mackay 2016). Thus far, genetic studies on flowering time in maize have described an architecture with predominantly additive genetic variance (*e.g.*, Buckler *et al.* 2009; Coles *et al.* 2010; note that these studies use inbred lines, which precludes estimation of dominance variance), but reports of dominant, overdominant (Coles *et al.* 2011) and epsistatic (Blanc *et al.* 2006; Durand *et al.* 2012) allele effects on variation in flowering characteristics also exist. With a limited sample size for quantitative genetic dissection per generation, our study is unable to clarify the causes or relative contribution of context-dependent effects on the response to phenotypic selection.

Maize is highly diverse (Buckler *et al.* 2006), and landraces of maize are locally adapted to a wide range of environments (Committee on The Preservation of Indigenous Strains of Maize 1952–1963). Still, it was surprising that Hallauer's Tusón captured nearly all of the SNPs on the MaizeSNP50 chip (Ganal *et al.* 2011). This high level of molecular genetic

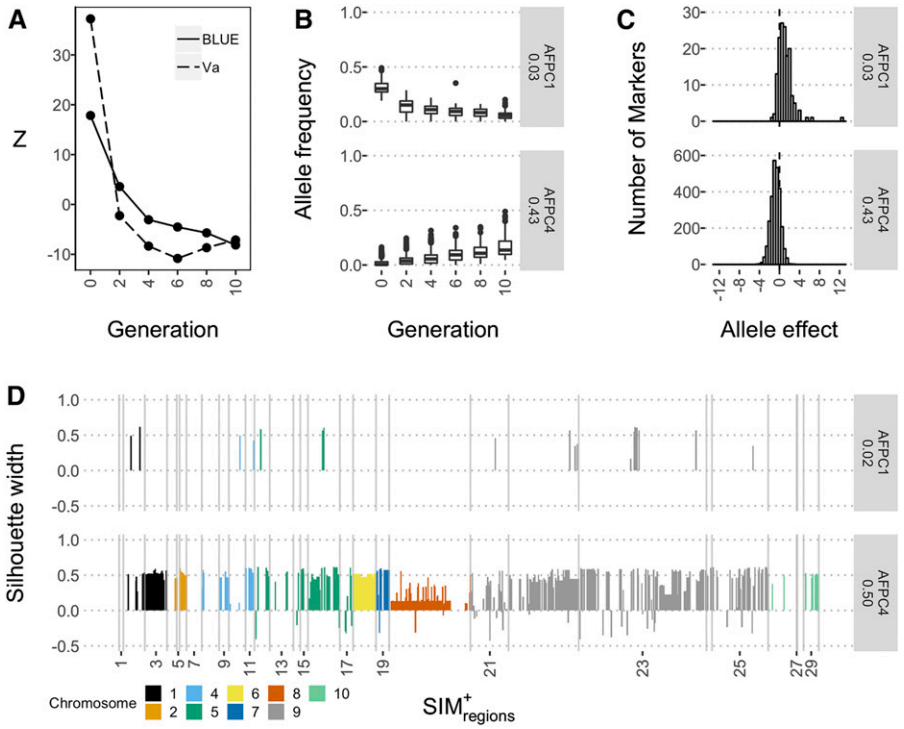

Figure 5 Quantitative and population genetic components of the response to selection. (A) Zero-centered (not standardized) "Z" values corresponding to the mean (BLUE) and additive variance (Va) for female-flowering time per generation. (B) Box plots of allele frequencies per generation for SIM^+ markers in AFPCs 1 and 4. (C) Histograms of additive allele effects for SIM^+ markers in AFPCs 1 and 4. (D) SIM hits (vertical lines colored by chromosome) within $SIM^+_{regions}$ (*x*-axis) that belonged to AFPCs 1 and 4. The silhouette width (*y*-axis) is a measure of a markers fit to the cluster (values > 0.5 can be considered a good fit). Facet labels indicate the AFPC identifier and the proportion of SIM^+ markers per cluster [in (B) and (C) this corresponds to the proportion among all SIM hits; in (D) this corresponds to the proportion among regional SIM hits].

variation, as well as the detection of an increasing proportion of polymorphic platform SNPs across generations (Table 1), led us to question whether migrant pollen had entered the population, particularly since it was open-pollinated during selection; although the population was bred in spatial and temporal isolation of other maize populations. Separate lines of evidence indicate the population could have very high diversity while remaining a closed system with no migration or pollen flow. First, similar to our finding, another study has found that individual populations of maize landraces can capture $>90\%$ of the SNPs on the same MaizeSNP50 platform (Arteaga *et al.* 2016). Because the base population of Hallauer's Tusón was admixed from multiple, geographically dispersed populations of the landrace Tusón, there is a greater likelihood for the level of diversity to be high. Second, the binomial sampling probabilities for our study limited detection of rare variants within generations despite their putative presence in the population. For instance, based on our sample sizes (which was larger for g_0), SNPs at a frequency of 1% have an 11% chance of being undetected in g_0 and a 33% chance of being undetected in the other generations, but subtle increases above 1% result in large increases in the probability of their detection. Therefore, variants that were not detected in one generation but detected in another may exist at low frequencies, and the transgenerational increase in polymorphic platform SNPs can be explained by selection of initially rare alleles. Finally, considering the most frequent migrant sources in Iowa where the population was selected would be of temperate origin, all of the genotyped individuals in Hallauer's Tusón clustered with other maize samples of tropical rather than temperate origin (Figure 2).

A fundamental question in genetics is how populations acquire and maintain variation that conditions them with the capacity to adapt to a novel environment. At the locations where the source populations were already adapted and grown, we presume stabilizing selection occurred on flowering time, as flowering time affects fitness in an environmentally dependent manner (Hall and Willis 2006; Mercer and Perales 2019). Stabilizing selection is expected to deplete genetic variation (Barton and Keightley 2002), such that the extensive functional variation for flowering time in Hallauer's Tusón suggests evolutionary forces beyond mutation and selection for a single optimal flowering time affected the founding populations. Teixeira *et al.* (2015) showed that flowering time variation in Hallauer's Tusón is under strong genetic and environmental control, with relatively little genotype-by-environment interaction (however, GxE effects were present across latitude and greater in the initial generations). Therefore, seasonal fluctuations that affect the relationship between flowering time and fitness (Giauffret *et al.* 2000) and multivariate constraints to evolution (Walsh and Blows 2009) likely contributed to the maintenance of substantial standing variation for this trait, and therefore its capacity for adaptation to a novel environment.

As the population was subjected to directional selection in a temperate environment, alleles contributing to earlier flowering tended to be enriched (Figure 4). Although SIM^+ alleles with negative effects on flowering time spanned the full allele frequency spectrum in g_0, including initially

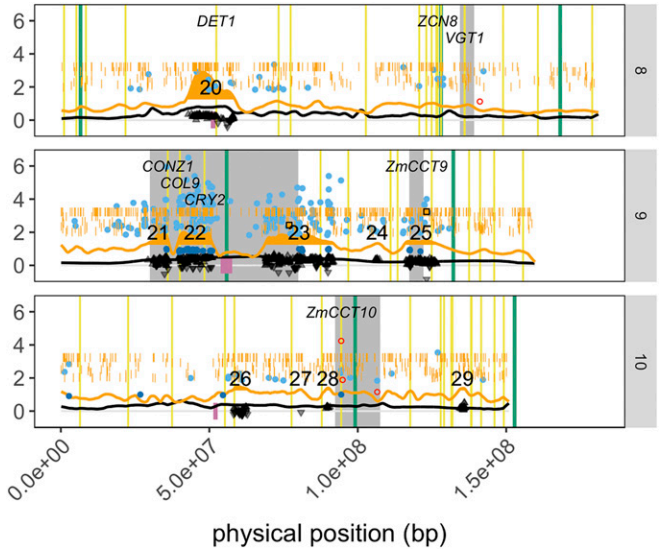

Figure 6 Synthesis map of chromosomes 8, 9, and 10. Multiple results are plotted on the physical map of each chromosome, with the *y*-axis corresponding to values for each of the following metrics: (i) kernel regression estimate of r^2 for LD between sequential pairs of markers (black line); (ii) kernel regression estimate of $-log10(q)$ for the SIM test (orange line: delimited $SIM^{+}_{regions}$ are enumerated and encompass the orange shaded areas); (iii) $-log10(q)$ value for SIM^{+} markers (orange vertical lines); (iv) $-log10(q)$ value for complete-sweep SIM^{+} markers (orange filled box); (v) difference in observed heterozygosity between g_0 and g_{10} for SIM^{+} markers in $SIM^{+}_{regions}$ (black-filled triangles: pointing up if the change was positive and down if the change was negative); (vi) *log*10(Bayes factor) values for $Bayenv^{+}$ markers (cyan-filled points); (vii) bootstrap values for $FITR^{+}$ markers (blue-filled points); (viii) $-log10(q)$ value for GWA^{+} markers (red-outlined points); (ix) QTL previously identified for photoperiod sensitivity (gray shaded areas corresponding to QTL intervals) and flowering time *per se* (green vertical lines corresponding to QTL peaks); (x) candidate genes for flowering time (yellow vertical lines and labels); and (xi) centromeres (lilac-colored boxes).

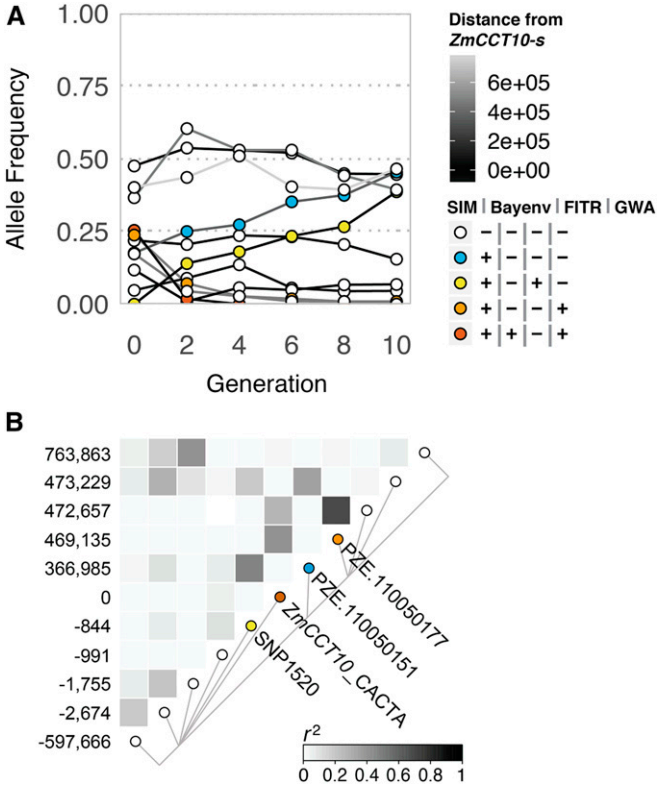

Figure 7 Allele frequency change and linkage disequilibrium at *ZmCCT10*. (A) Variant frequency change (*y*-axis) across generations (*x*-axis) for the nearest five markers flanking the ZmCCT10_CACTA causal site. Points are color coded according to the combination of test results (ZmCCT10_CACTA was positive for the SIM, Bayenv and GWA tests and is color coded red). (B) Pairwise LD between the markers in (A). Marker distances from the insertion site for ZmCCT10_CACTA (AGPv4 94,435,768) are indicated. Negative values are in the direction of the *ZmCCT10* transcription start site.

frequent variants such as the photoperiod insensitive allele *ZmCCT10-i* (present at 75% in g_0), most of these ($\approx$ 75%) existed in the minor frequency domain. This was in contrast to minor alleles in g_0 at SIM^{-} markers, for which $\approx$ 50% had negative effects on flowering time. Hence, in the base generation of Hallauer's Tusón, favorable alleles for temperate adaptation primarily exist in the minor frequency spectrum.

Due to the admixture of multiple Tusón populations to form g_0, however, the allele frequencies reflect those among (not within) the founder populations, such that native population allele frequencies of temperate-adaptive variants are confounded. To address this, subpopulation assignments for samples from g_0 (assumed to correspond to the founding populations) were used to compute subpopulation-specific allele frequencies for the corresponding minor allele in the whole g_0 sample (data not shown). Across all SIM^{+} markers, most alleles were shared among multiple subpopulations; only 8% of these markers included private alleles. This suggests that geographically separated populations of the landrace Tusón have retained shared alleles that enable latitudinal adaptation, a finding that is congruent with geographical association results from a diverse sample of maize landraces (Romero Navarro *et al.* 2017).

Our inference from allele frequency data, however, should be considered with caution. The inferred number of loci (finite polygenic architecture) and effect of selection on variants across the allele frequency spectrum is based on markers that are unlikely to be causal variants themselves, but are expected to tag local haplotype blocks containing causal variants (Nuzhdin and Turner 2013; Kelly and Hughes 2019). Moreover, any initially rare variants with late flowering time effects would likely have been purged or kept at low frequencies, leading to low power of detection for both selection and association mapping, despite the importance of such potential variants for environmental adaptation. This represents a bias to evolutionary inference in experimentally evolved populations and highlights the need for deeper sampling within generations and other approaches to elucidate the structure of natural variation for adaptation.

Our study design allows for the mapping and characterization of specific genomic loci underlying phenotypic evolution. Although the relatively small sample ($\approx$ 300 families) and low marker density (tens of thousands of SNPs) limited

the power of our study, we nevertheless detected robust associations with known genes involved in flowering time adaptation, raising confidence in our discovery of other unique loci (Figure S7). For instance, selection against photoperiod sensitivity was a major component of adaptation in Hallauer's Tusón (Teixeira *et al.* 2015). Regions encompassing *ZCN8*, *CONZ1*, *COL9*, *CRY2*, and *ZmCCT9* on chromosomes 8 and 9, and a causal regulatory variant of *ZmCCT10* on chromosome 10, were detected by selection or association mapping (Figure 6). These genes are regulators of photoperiodism in plants (Guo *et al.* 1989; Cheng and Wang 2005; Miller *et al.* 2008; Meng *et al.* 2011) that have contributed to latitudinal adaptation (Yang *et al.* 2013; Guo *et al.* 2018; Huang *et al.* 2018), all of which were identified in a large-scale mapping study on photoperiod sensitivity in maize (*cf*. Figure 6 and Figure 2 in Hung *et al.* 2012). Moreover, we found that different local haplotypes were responsive to selection at several of these same loci where Hung *et al.* (2012) detected allelic series [*e.g.*, $SIM^{+}_{regions}21-23$ and 25 (Figure 5) and the *ZmCCT10* locus (Figure 7)]. On chromosome 8, no robust association was detected for the *Vegetative to Generative Transition 1* (*VGT1*) gene (Salvi *et al.* 2007) involved in earliness *per se*, but a maize homolog of *Arabidopsis thaliana de-etiolated 1* (*DET1*) involved in photomorphogenesis (Pepper *et al.* 1994), not previously highlighted for maize adaptation, mapped to a conspicuous SIM^{+}_{region} across the centromere of the chromosome.

Development of next-generation crop varieties is crucial to ensuring ample production and quality of plant-based products for society. Reliance on limited pools of diversity for breeding and the creation of monocultures can lead to constrained and vulnerable production systems, but genetic diversity is ecologically structured whereby alleles that could be useful in a target production environment reside in germplasm that suffers from maladaptive syndromes (Teixeira *et al.* 2015). Our study shows the potential for maize landraces to be adapted to temperate environments by simple recurrent selection for early flowering time while resulting in minimal loss of molecular genetic variation to reach the adapted state. The lack of a linked footprint of selection (SIM^{+}_{region}) encompassing the *ZmCCT10* causal site, which had the strongest association with flowering time and underwent a rapid complete-sweep, suggests that some critical adaptive mutations in maize are embedded in regions of low LD, which would permit the maintenance of linked variation during directional selection. Moreover, the multidimensional nature of the genetic architecture underlying response to *phenotypic* selection, involving multiple loci and alleles with context-dependent effects, appears to enable a rapid shift toward an adapted state with limited loss in diversity. We anticipate that further characterization of these additional layers of the genetic architecture and dynamics of the genomic response to selection will lead to new advances in genomic prediction across generations. By bridging population and quantitative genetic inference, this study advances our understanding of short-term evolution, providing unique insights that aid in developing approaches to adapt crops to climate change.

Acknowledgments

We thank Michael Dumas (University of Delaware) who produced genotype data for the *ZmCCT10* causal site. R.J.W. thanks Keith Hopper (USDA-ARS) for meetings at Iron Hill where this work was discussed at length. This project was supported by Agriculture and Food Research Initiative Grant No. 2011-67003-30342 from the United States Department of Agriculture (USDA) National Institute of Food and Agriculture (Agriculture and Natural Resources Science for Climate Variability and Change Program).

Author contributions: A.H. created the original recurrently selected population. R.J.W., J.B.H., N.d.L., S.F.-G., N.L., S.C.M., and W.X. designed the study and conducted field experiments, while T.W. assisted with coordination. J.E.C.T. led the genotyping project and contributed intellectually. R.J.W. and J.D. developed the whole genome simulator. R.J.W., Z.F., and J.B.H. developed analysis procedures and analyzed the data. R.J.W. and J.B.H. interpreted the results and wrote the paper, with contributions from Z.F. All authors except T.W. and A.H. edited the manuscript.

Literature Cited

Arteaga, M. C., A. Moreno-Letelier, A. Mastretta-Yanes, A. Vázquez-Lobo, A. Breña-Ochoa *et al.*, 2016 Genomic variation in recently collected maize landraces from Mexico. Genom. Data 7: 38–45. https://doi.org/10.1016/j.gdata.2015.11.002

Barrick, J. E., and R. E. Lenski, 2013 Genome dynamics during experimental evolution. Nat. Rev. Genet. 14: 827–839. https://doi.org/10.1038/nrg3564

Barton, N. H., and P. D. Keightley, 2002 Understanding quantitative genetic variation. Nat. Rev. Genet. 3: 11–21. https://doi.org/10.1038/nrg700

Barton, N. H., A. M. Etheridge, and A. Véber, 2017 The infinitesimal model: definition, derivation, and implications. Theor. Popul. Biol. 118: 50–73. https://doi.org/10.1016/j.tpb.2017.06.001

Baxter, S. W., F. R. Badenes-Pérez, A. Morrison, H. Vogel, N. Crickmore *et al.*, 2011 Parallel evolution of *Bacillus thuringiensis* toxin resistance in Lepidoptera. Genetics 189: 675–679. https://doi.org/10.1534/genetics.111.130971

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57: 289–300.

Berg, J. J., and G. Coop, 2014 A population genetic signal of polygenic adaptation. PLoS Genet. 10: e1004412. https://doi.org/10.1371/journal.pgen.1004412

Bhatia, G., N. Patterson, S. Sankararaman, and A. Price, 2013 Estimating and interpreting FST: the impact of rare variants. Genome Res. 23: 1514–1521. https://doi.org/10.1101/gr.154831.113

Blanc, G., A. Charcosset, B. Mangin, A. Gallais, and L. Moreau, 2006 Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor. Appl. Genet. 113: 206–224. https://doi.org/10.1007/s00122-006-0287-1

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss *et al.*, 2007 TASSEL: software for association mapping of

complex traits in diverse samples. Bioinformatics 23: 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Buckler, E. S., B. S. Gaut, and M. D. McMullen, 2006 Molecular and functional diversity of maize. Curr. Opin. Plant Biol. 9: 172–176. https://doi.org/10.1016/j.pbi.2006.01.013

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown *et al.*, 2009 The genetic architecture of maize flowering time. Science 325: 714–718. https://doi.org/10.1126/science.1174276

Bulmer, M. G., 1971 The effect of selection on genetic variability. Am. Nat. 105: 201–211. https://doi.org/10.1086/282718

Burke, M. K., and A. D. Long, 2012 What paths do advantageous alleles take during short-term evolutionary change? Mol. Ecol. 21: 4913–4916. https://doi.org/10.1111/j.1365-294X.2012.05745.x

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose *et al.*, 2010 Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature 467: 587–590. https://doi.org/10.1038/nature09352

Chan, Y. F., F. C. Jones, E. McConnell, J. Bryk, L. Bünger *et al.*, 2012 Parallel selection mapping using artificially selected mice reveals body weight control loci. Curr. Biol. 22: 794–800. https://doi.org/10.1016/j.cub.2012.03.011

Chardon, F., L. Moreau, M. Falque, J. Joets, L. Decousset *et al.*, 2004 Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. Genetics 168: 2169–2185. https://doi.org/10.1534/genetics.104.032375

Cheng, X. F., and Z. Y. Wang, 2005 Overexpression of *COL9*, a *CONSTANS-LIKE* gene, delays flowering by reducing expression of *CO* and *FT* in *Arabidopsis thaliana*. Plant J. 43: 758–768. https://doi.org/10.1111/j.1365-313X.2005.02491.x

Chevalet, C., 1994 An approximate theory of selection assuming a finite number of quantitative trait loci. Genet. Sel. Evol. 26: 379–400. https://doi.org/10.1186/1297-9686-26-5-379

Chevin, L. M., and F. Hospital, 2008 Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics 180: 1645–1660. https://doi.org/10.1534/genetics.108.093351

Coles, N. D., M. D. McMullen, P. J. Balint-Kurti, R. C. Pratt, and J. B. Holland, 2010 Genetic control of photoperiod sensitivity in maize revealed by joint multiple population analysis. Genetics 184: 799–812. https://doi.org/10.1534/genetics.109.110304

Coles, N. D., C. T. Zila, and J. B. Holland, 2011 Allelic effect variation at key photoperiod response quantitative trait loci in maize. Crop Sci. 51: 1036–1049. https://doi.org/10.2135/cropsci2010.08.0488

Committee on The Preservation of Indigenous Strains of Maize monographs on Races of Maize (1952–1963). National Academy of Sciences—National Research, Washington, D.C. Pub. N 1136 (Venezuela); Pub. N 510 (Columbia); Pub. N 511 (Central America); Pub. N 593 (Brazil and other Eastern South American Countries); Pub. N 747 (Bolivia); Pub. N 792 (West Indies); Pub. N 847 (Chile); Pub. N 915 (Peru); Pub. N 975 (Ecuador); Pub. N 453 (Cuba); and Races of Maize in Mexico by the Bussey Institute, Harvard University, MA.

Coop, G., D. Witonsky, A. Di Rienzo, and J. K. Pritchard, 2010 Using environmental correlations to identify loci underlying local adaptation. Genetics 185: 1411–1423. https://doi.org/10.1534/genetics.110.114819

Darwin, C., 1859 *On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Ducrocq, S., D. Madur, J. Veyrieras, L. Camus-Kulandaivelu, M. Kloiber-Maitz *et al.*, 2008 Key impact of *Vgt1* on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. Genetics 178: 2433–2437. https://doi.org/10.1534/genetics.107.084830

Durand, E., S. Bouchet, P. Bertin, A. Ressayre, P. Jamin *et al.*, 2012 Flowering time in maize: linkage and epistasis at a major effect locus. Genetics 190: 1547–1562. https://doi.org/10.1534/genetics.111.136903

Durand, E., M. I. Tenaillon, X. Raffoux, S. Thépot, M. Falque *et al.*, 2015 Dearth of polymorphism associated with a sustained response to selection for flowering time in maize. BMC Evol. Biol. 15: 103. https://doi.org/10.1186/s12862-015-0382-5

Evanno, G., S. Regnaut, and J. Goudet, 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, Ed. 4. Pearson, Essex.

Fernando, R. L., C. Stricker, and R. C. Elston, 1994 The finite polygenic mixed model: an alternative formulation for the mixed model of inheritance. Theor. Appl. Genet. 88: 573–580. https://doi.org/10.1007/BF01240920

Fisher, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinb. 52: 399–433. https://doi.org/10.1017/S0080456800012163

Frascaroli, E., T. A. Schrag, and A. E. Melchinger, 2013 Genetic diversity analysis of elite European maize (Zea mays L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. Theor. Appl. Genet. 126: 133–141. https://doi.org/10.1007/s00122-012-1968-6

Ganal, M. W., G. Durstewitz, A. Polley, A. Bérard, E. S. Buckler *et al.*, 2011 A large maize (Zea mays L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6: e28334. https://doi.org/10.1371/journal.pone.0028334

Giauffret, C., J. Lothrop, D. Dorvillez, B. Gouesnard, and M. Derieux, 2000 Genotype x environment interactions in maize hybrids from temperate or highland tropical origin. Crop Sci. 40: 1004–1012. https://doi.org/10.2135/cropsci2000.4041004x

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson, 2009 ASReml user guide release 3.0. Hemel Hemptead: VSN International Ltd.

Goodman, M., 1998 Research policies thwart potential payoff of exotic germplasm. Diversity (Basel) 14: 30–35.

Goodman, M. M., and W. L. Brown, 1988 Races of corn, pp. 33–74 in *Corn and Corn Improvement*, Ed. 3, Chap. 2, edited by G. F. Sprague and J. W. Dudley. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI.

Goudet, J., 2005 HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Mol. Ecol. Notes 5: 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Graffelman, J., 2015 Exploring diallelic genetic markers: the HardyWeinberg package. J. Stat. Softw. 64: 1–23. https://doi.org/10.18637/jss.v064.i03

Guo, H., H. Yang, T. C. Mockler, and C. Lin, 1989 Regulation of flowering time by *Arabidopsis* photoreceptors. Science 279: 1360–1363. https://doi.org/10.1126/science.279.5355.1360

Guo, L., X. Wang, M. Zhao, C. Huang, C. Li *et al.*, 2018 Stepwise cis-regulatory changes in ZCN8 contribute to maize flowering-time adaptation. Curr. Biol. 28: 3005–3015.e4. https://doi.org/10.1016/j.cub.2018.07.029

Hall, M. C., and J. H. Willis, 2006 Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. Evolution 60: 2466–2477. https://doi.org/10.1111/j.0014-3820.2006.tb01882.x

Hallauer, A. R., and M. J. Carena, 2016 Registration of BS39 maize germplasm. J. Plant Regist. 10: 296–300. https://doi.org/10.3198/jpr2015.02.0008crg

Hartl, D. L., and A. G. Clark, 2007 *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA.

Herrmann, E., 2016 lokern: kernel regression smoothing with local or global plug-in bandwidth. Packaged for R and enhanced by Martin Maechler. R package version 1.1-6. Available at: https://CRAN.R-project.org/package=lokern
Hill, W. G., and A. Robertson, 1968 Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38: 226–231. https://doi.org/10.1007/BF01245622
Hill, W. G., M. E. Goddard, and P. M. Visscher, 2008 Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet. 4: e1000008. https://doi.org/10.1371/journal.pgen.1000008
Holland, J., W. Nyquist, and C. Cervantes-Martínez, 2003 Estimating and interpreting heritability for plant breeding: an update, pp. 9–112 in *Plant Breeding Reviews*, Vol. 22, Chap. 2, edited by J. Janick. John Wiley and Sons, Inc., Hoboken, NJ.
Huang, C., H. Sun, D. Xu, Q. Chen, Y. Liang *et al*., 2018 ZmCCT9 enhances maize adaptation to higher latitudes. Proc. Natl. Acad. Sci. USA 115: E334–E341.
Huang, W., and T. F. Mackay, 2016 The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLoS Genet. 12: e1006421. https://doi.org/10.1371/journal.pgen.1006421
Hung, H.-Y., L. M. Shannon, F. Tian, P. J. Bradbury, C. Chen *et al*., 2012 *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc. Natl. Acad. Sci. USA 109: E1913–E1921. https://doi.org/10.1073/pnas.1203189109
Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli *et al*., 2012 The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484: 55–61. https://doi.org/10.1038/nature10944
Jung, C., and A. E. Müller, 2009 Flowering time control and applications in plant breeding. Trends Plant Sci. 14: 563–573. https://doi.org/10.1016/j.tplants.2009.07.005
Jung, C., K. Pillen, D. Staiger, G. Coupland, and M. von Korff, 2017 Editorial: recent advances in flowering time control. Front. Plant Sci. 7: 2011. https://doi.org/10.3389/fpls.2016.02011
Kaufman, L., and P. J. Rousseeuw, 1990 *Finding Groups in Data: an Introduction to Cluster Analysis*. Wiley-Interscience, Hoboken, NJ. https://doi.org/10.1002/9780470316801
Kelly, J. K., and K. A. Hughes, 2019 Pervasive linked selection and intermediate-frequency alleles are implicated in an evolve-and-resequencing experiment of *Drosophila simulans*. Genetics 211: 943–961. https://doi.org/10.1534/genetics.118.301824
Le Corre, V., and A. Kremer, 2012 The genetic differentiation at quantitative trait loci under local adaptation. Mol. Ecol. 21: 1548–1566. https://doi.org/10.1111/j.1365-294X.2012.05479.x
Levy, S. F., J. R. Blundell, S. Venkataram, D. A. Petrov, D. S. Fisher *et al*., 2015 Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature 519: 181–186. https://doi.org/10.1038/nature14279
Li, Y. X., C. Li, P. J. Bradbury, X. Liu, F. Lu *et al*., 2016 Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. Plant J. 86: 391–402. https://doi.org/10.1111/tpj.13174
Lowry, D. B., and J. H. Willis, 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol. 8: pii: e1000500 [corrigenda: PLoS Biol. 10: 10 (2012)]. https://doi.org/10.1371/journal.pbio.1000500
Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik, 2018 *Cluster: cluster analysis basics and extensions*. R package version 2.1.0.
McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. H. Li *et al*., 2009 Genetic properties of the maize nested association mapping population. Science 325: 737–740. https://doi.org/10.1126/science.1174320
Meng, X., M. G. Muszynski, and O. N. Danilevskaya, 2011 The *FT*-Like *ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize. Plant Cell 23: 942–960. https://doi.org/10.1105/tpc.110.081406
Mercer, K. L., and H. Perales, 2019 Structure of local adaptation across the landscape: flowering time and fitness in Mexican maize (*Zea mays* L. subsp. *mays*) landraces. Genet. Resour. Crop Evol. 66: 27–45. https://doi.org/10.1007/s10722-018-0693-7
Miller, T. A., E. H. Muslin, and J. E. Dorweiler, 2008 A maize *CONSTANS*-like gene, *conz1*, exhibits distinct diurnal expression patterns in varied photoperiods. Planta 227: 1377–1388. https://doi.org/10.1007/s00425-008-0709-1
Moon, K. R., D. van Dijk, Z. Wang, D. Burkhardt, W. Chen *et al*., 2017 Visualizing transitions and structure for high dimensional data exploration. bioRxiv.
Nei, M., 1986 Definition and estimation of fixation indices. Evolution 40: 643–645. https://doi.org/10.1111/j.1558-5646.1986.tb00516.x
Neph, S., M. S. Kuehn, A. P. Reynolds, E. Haugen, R. E. Thurman *et al*., 2012 BEDOPS: high-performance genomic feature operations. Bioinformatics 28: 1919–1920. https://doi.org/10.1093/bioinformatics/bts277
Nishino, J., 2013 Detecting selection using time-series data of allele frequencies with multiple independent reference loci. G3 (Bethesda) 3: 2151–2161. https://doi.org/10.1534/g3.113.008276
Nuzhdin, S. V., and T. L. Turner, 2013 Promises and limitations of hitchhiking mapping. Curr. Opin. Genet. Dev. 23: 694–699. https://doi.org/10.1016/j.gde.2013.10.002
Orozco-terWengel, P., M. Kapun, V. Nolte, R. Kofler, T. Flatt *et al*., 2012 Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. Mol. Ecol. 21: 4931–4941. https://doi.org/10.1111/j.1365-294X.2012.05673.x
Orr, H. A., 2005 The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119–127. https://doi.org/10.1038/nrg1523
Peng, B., and M. Kimmel, 2005 simuPOP: a forward-time population genetics simulation environment. Bioinformatics 21: 3686–3687. https://doi.org/10.1093/bioinformatics/bti584
Pepper, A., T. Delaney, T. Washburnt, D. Poole, and J. Chory, 1994 *DET1*, a negative regulator of light-mediated development and gene expression in *Arabidopsis*, encodes a novel nuclear-localized protein. Cell 78: 109–116. https://doi.org/10.1016/0092-8674(94)90577-0
Pritchard, J. K., M. Stephens, and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. Genetics 155: 945–959.
R Core Team, 2016 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Romay, M. C., M. J. Millard, J. C. Glaubitz, J. Peiffer, K. L. Swarts *et al*., 2013 Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14: R55. https://doi.org/10.1186/gb-2013-14-6-r55
Romero Navarro, J. A. R., M. Willcox, J. Burgueño, C. Romay, K. Swarts *et al*., 2017 A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nat. Genet. 49: 476–480 (erratum: Nat. Genet. 49: 970). https://doi.org/10.1038/ng.3784
Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu *et al*., 2007 Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc. Natl. Acad. Sci. USA 104: 11376–11381. https://doi.org/10.1073/pnas.0704145104
Savolainen, O., M. Lascoux, and J. Merilä, 2013 Ecological genomics of local adaptation. Nat. Rev. Genet. 14: 807–820. https://doi.org/10.1038/nrg3522
Schlötterer, C., R. Kofler, E. Versace, R. Tobler, and S. U. Franssen, 2015 Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. Heredity (Edinb) 114: 431–440 [corrigenda:

Heredity (Edinb) 116: 248 (2016)]. https://doi.org/10.1038/hdy.2014.86

Shin, J.-H., S. Blay, B. McNeney, and J. Graham, 2006 Ldheatmap: an r function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. J. Stat. Softw. 16: 1–9. https://doi.org/10.18637/jss.v016.c03

Spitze, K., 1993 Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135: 367–374.

Su, G., O. F. Christensen, T. Ostersen, M. Henryon, and M. S. Lund, 2012 Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. PLoS One 7: e45293. https://doi.org/10.1371/journal.pone.0045293

Swarts, K., R. M. Gutaker, B. Benz, M. Blake, R. Bukowski *et al.*, 2017 Genomic estimation of complex traits reveals ancient maize adaptation to temperate North. Am. Sci. 357: 512–515.

Teixeira, J. E. C., T. Weldekidan, N. de Leon, S. Flint-Garcia, J. B. Holland *et al.*, 2015 Hallauer's Tusón: a decade of selection for tropical-to-temperate phenological adaptation in maize. Heredity 114: 229–240. https://doi.org/10.1038/hdy.2014.90

Tenaillon, O., A. Rodriguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett *et al.*, 2012 The molecular diversity of adaptive convergence. Science 335: 457–461. https://doi.org/10.1126/science.1212986

Tibshirani, R., G. Walther, and T. Hastie, 2001 Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. B 63: 411–423. https://doi.org/10.1111/1467-9868.00293

Turelli, M., and N. Barton, 1994 Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138: 913–941.

VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. J. Dairy Sci. 91: 4414–4423. https://doi.org/10.3168/jds.2007-0980

Walsh, B., and M. W. Blows, 2009 Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40: 41–59. https://doi.org/10.1146/annurev.ecolsys.110308.120232

Walsh, B., and M. Lynch, 2018 Short-term changes in the variance, in *Evolution and Selection of Quantitative Traits*, Chap. 16. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780198830870.001.0001

Warnes, G., G. Gorjanc, F. Leisch, and M. Man, 2013 genetics: population genetics. R package version 1.3.8.1. Available at: https://CRAN.R-project.org/package=genetics

Wisser, R. J., P. J. Balint-Kurti, and J. B. Holland, 2011 A novel genetic framework for studying response to artificial selection. Plant Genet. Resour. 9: 281–283. https://doi.org/10.1017/S1479262111000359

Yang, Q., Z. Li, W. Li, L. Ku, C. Wang *et al.*, 2013 CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proc. Natl. Acad. Sci. USA 110: 16969–16974. https://doi.org/10.1073/pnas.1310949110

Yu, J. M., G. Pressoir, W. H. Briggs, I. V. Bi, M. Yamasaki *et al.*, 2006 A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203–208. https://doi.org/10.1038/ng1702

Communicating editor: S. Wright