Published online March 15, 2018

ORIGINAL RESEARCH

Association Mapping of Flowering and Height Traits in Germplasm Enhancement of Maize Doubled Haploid (GEM-DH) Lines

Adam Vanous, Candice Gardner,* Michael Blanco, Adam Martin-Schwarze, Alexander E. Lipka, Sherry Flint-Garcia, Martin Bohn, Jode Edwards, and Thomas Lübberstedt

ABSTRACT

Flowering and height related traits are extensively studied in maize for three main reasons: 1) easily obtained phenotypic measurements, 2) highly heritable, and 3) importance of these traits to adaptation and grain yield. However, variation in flowering and height traits is extensive and findings from previous studies are genotype specific. Herein, a diverse panel of exotic derived doubled haploid lines, in conjunction with genome-wide association analysis, is used to further explore adaptation related trait variation of exotic germplasm for potential use in adapting exotic germplasm to the U.S. Corn-Belt. Phenotypes for the association panel were obtained from six locations across the central-U.S. and genotyping was performed using the genotyping-by-sequencing method. Nineteen flowering time candidate genes were found for three flowering traits. Eighteen candidate genes were found for four height related traits, with the majority of the candidate genes relating to plant hormones auxin and gibberellin. A single gene was discovered for ear height that also had effects on *FT*-like flowering gene expression levels. Findings will be used to inform future research efforts of the USDA Germplasm Enhancement of Maize project and eventually aid in the rapid adaptation of exotic germplasm to temperate U.S. environments.

Core Ideas

- Genome-wide association mapping in exotic derived double haploid maize lines.
- Genomic areas collocate with previously identified QTLs and candidate genes for flowering and height.
- Novel regions identified allow for future research of adapting exotic maize to the central-U.S. Corn-Belt.

A. Vanous, T. Lübberstedt, Dep. of Agronomy, Iowa State Univ., Ames, IA 50011; M. Blanco, J. Edwards, C. Gardner, USDA-ARS and Dep. of Agronomy, Iowa State Univ., Ames, IA 50011; A. Martin-Schwarze, Dep. of Statistics, Iowa State Univ., Ames, IA 50011; S. Flint-Garcia, Division of Plant Sciences, Univ. of Missouri, Columbia, MO 65211; M. Bohn, Dep. of Crop Sciences and the Illinois Plant Breeding Center, Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801; A. Lipka, Dep. of Crop Sciences, Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801. Received January 30 Sep. 2017. Accepted 4 Jan. 2018. *Corresponding author (Candice.Gardner@ars.usda.gov).

Abbreviations: AEPH, above ear plant height; ASI, anthesis to silking interval; ARF, auxin response factor; BLUPs, BGEM, Iowa State University and Germplasm Enhancement of Maize doubled haploid; best linear unbiased predictions; DH, doubled haploid; EH, ear height; GBS, genotyping-by-sequencing; GDUs, growing degree units; GDUSHD, growing degree units to anthesis; GDUSLK, growing degree units to silk emergence; GEM, Germplasm Enhancement of Maize; GEM-DH, germplasm enhancement of maize doubled haploid; GLM, general linear model; GWAS, genome-wide association study; IBM, intermated B73 × Mo17; LD, linkage disequilibrium; LOESS, local regression; MLM, mixed linear model; NAM, nested association mapping; NSS, non-stiff stalk; PCA, principal component analysis; PH, plant height; PVP, plant variety protection; QQ, quantile-quantile; QTL, quantitative trait loci; REHPH, ratio of ear height to plant height; RILs, recombinant inbred lines; RSHDSLK, ratio of growing degree units to anthesis to growing degree units to silk emergence; SAM, shoot apical meristem; SAUR, small auxin upregulated RNA; SNP, single nucleotide polymorphism; SS, stiff stalk.

Plant Genome 11:170083
doi: 10.3835/plantgenome2017.09.0083

Improvement of maize (Zea mays L.) is important to deal with challenges in production due to changes in the environment, society, management practices, and resource availability (Cohen, 2003; Ranum et al., 2014; Halford and Foyer, 2015). Increasing genetic diversity is essential when considering the current and predicted changes in climate and the world's ever increasing population size. Exotic maize sources can be used to increase diversity and provide genetic variation needed by plant breeders to address these issues (Hallauer and Carena, 2014). Lopes et al. (2015) discuss the successes when using landraces for increasing genetic diversity in wheat due to adaptation and climate change. Tropical maize germplasm was used by Menkir et al. (2015) to increase the provitamin-A content to improve nutrition for human consumption of maize. However, increasing genetic diversity in maize does not come without problems. Flowering time in maize is one of the most important adaptive traits that has led to the domestication and spread of maize (Li et al., 2016b). Flowering time is important in determining local adaptation, and is one of the largest issues that must be overcome when locally adapting new maize germplasm.

Flowering time has been studied in a number of plant species (Izawa et al., 2003; Nemoto et al., 2003; Kong et al., 2010; Blackman et al., 2011; Murphy et al., 2011). The model organism Arabidopsis thaliana has been used to identify hundreds of flowering time genes and has aided in explaining the overall regulatory system controlling flowering time (Gazzani et al., 2003; Izawa et al., 2003; Jung and Müller, 2009; Brachi et al., 2010; Méndez-Vigo et al., 2013), with many of these findings contributing to the understanding of maize flowering time. Identified genes from Arabidopsis thaliana include the FLOWERING LOCUS T (FT), CONSTANS (CO), TERMINAL FLOWER1 (TFL1), and SUPPRESOR OF OVEREXPRESSION OF CONSTANS (SOC1). Many important maize flowering time studies have been conducted (Buckler et al., 2009; Romay et al., 2013; Li et al., 2016b; Wallace et al., 2016); however, flowering time is a highly quantitative trait (Buckler et al., 2009), and maize is highly diverse (Wright et al., 2005), adding to the difficulty of explaining flowering time in maize. Buckler et al. (2009) used the 5000 recombinant inbred lines (RILs) nested association mapping (NAM) population to find 36, 39, and 29 quantitative trait loci (QTL) for days to anthesis, days to silking, and anthesis to silking interval (ASI), respectively. Buckler et al. (2009) suggests a simple additive model explaining flowering time in maize, showing that each QTL explained only a small portion of flowering time variation. In a follow up study, Li et al. (2016b) used the NAM population, 1745 lines from the Ames panel (Romay et al., 2013), and 2000 RILs from the Chinese nested association mapping population to identify 90 flowering time regions and 220 candidate genes. From these studies, numerous genes have been identified and include FT-like genes (ZCN8, ZCN13, ZCN16, ZCN20, ZCN24, and ZCN26), TFL-like genes (ZCN5 and ZCN6), CO-like genes (ZmCCT, CONZ1, COL3, COL6, and COL7), and the SOC-like gene (ZMM4). The QTLs, such as Vgt1 (Vegetative to generative 1; Salvi et al., 2007), dlf1 (delayed flowering1; Muszynski et al., 2006), and ZmCCT (Hung et al., 2012; Yang et al., 2013) have been positionally cloned; however, the effects of these positionally cloned QTL on adaptation of maize germplasm is unknown.

Plant height (PH) and ear height (EH) are also traits of interest when adapting germplasm as they are closely associated with flowering time, biomass, lodging resistance, and yield (Durand et al., 2012; Teng et al., 2013). For grain production, a short statured plant is preferable; however, unadapted tropical germplasm tends to be late flowering and tall. Peiffer et al. (2014) used the NAM population (Buckler et al., 2009) and the intermated B73 × Mo17 (IBM) (Lee et al., 2002) population to investigate PH and EH in maize. They identified 89 and 92 QTL for PH and EH, respectively. They also noted that PH and EH were highly polygenic, and that effects of these genes could explain a population's variation in height; however, due to changing effect sizes across populations, predicting height could be problematic.

Plant height, EH, and flowering traits are closely correlated. Gibberellin is a plant hormone that has close ties to height traits in maize but can also affect flowering time (Song et al., 2012; Li et al., 2016a). Song et al. (2012) found that gibberellins can affect the expression levels of *FT*-like flowering genes when under short-day or long-day conditions. Li et al. (2016a) identified eight genes in a flowering time study that are known to affect gibberellin levels. Most importantly, three of these genes are known to positively regulate the *GA20ox1* gene, in *Arabidopsis thaliana*, which is also known to greatly affect plant stature in maize (Wang et al., 2013; Voorend et al., 2016).

Flowering and height traits are two of the most studied traits in maize. In this study, we used a diverse panel of exotic derived doubled haploid (DH) lines, representing 52 exotic maize races, to investigate flowering and height traits to aid in the understanding of adapting exotic germplasm to the U.S. Corn Belt, with the underlying intention to implement useful findings in future work that would aid the USDA GEM adaptation program. The overall objective was to characterize the genetic architecture of adaptation traits flowering time and height. Flowering traits, anthesis and silking, were recorded and ASI was calculated. Plant height and EH measurements were also obtained. Trait correlations were explored to assist in explaining relationships among these adaptation traits. A genome-wide association study was used to explore the underlying genetic basis for variation within traits and identified significant associations were contrasted against prior published QTL regions to explore novelty of associations. Candidate genes were also identified for significant associations.

MATERIALS AND METHODS

Germplasm

A diverse panel consisting of 252 exotic derived DH lines was created by crossing 54 exotic maize accessions, representing 52 exotic maize races, to expired Plant Variety

Protection (PVP) line PHZ51, PHB47, or both PHZ51 and PHB47, and backcrossing each F_1 to the respective PVP parent of the F_1, resulting in 71 unique BC_1F_1 populations. These PVP lines represent popular heterotic breeding pools commonly used in maize breeding. Inbred PHZ51 represents the Lancaster or non-Stiff-Stalk (NSS) heterotic pattern, while PHB47 represents the Stiff-Stalk (SS) heterotic pattern. These lines were developed and released in a joint collaboration between the Iowa State University Doubled Haploid Facility and the USDA-ARS Germplasm Enhancement of Maize (GEM) project. Released lines are known as BGEM lines, B indicating Iowa State University inbred line and GEM indicating the Germplasm Enhancement of Maize project. Supplemental Table S1 lists the 252 BGEM lines, parental race, accession number, and the country of origin and elevation of the parental landraces.

The BC_1F_1 plants were grown and crossed to the maternal haploid inducer line, RWS/RWK-76 (Röber et al., 2005). Seed produced from these crosses was screened and haploid kernels were identified. Haploid kernels were planted in the greenhouse and seedlings underwent artificial chromosome doubling protocols used by the Iowa State Doubled Haploid Facility and the GEM project (Brenner et al., 2012). Haploid plants were then transplanted to the field, and those producing fertile pollen were self-pollinated to produce D_0 generation seed. Seed was increased of these D_1 generation DH lines in subsequent generations. During subsequent generations, lines were screened for uniformity and discarded if found to be contaminants, or if general agronomic traits strongly impaired maintenance.

Experimental Design and Data Collection

The diverse panel of 252 exotic derived DH lines was planted in the field in 2013 in an α incomplete block design across multiple environments. Row lengths were nine meters and 15 seeds were planted per row. Environments were: 1) Crop Sciences Research and Education Center, Champaign, Illinois, two replications (University of Illinois), 2) North Central Regional Plant Introduction Station, Ames, Iowa, three replications (Iowa State University, planting date 1-May 16, 2013), 3) Bradford Farm, Columbia, Missouri, two replications (University of Missouri), 4) Genetics Farm, Columbia, Missouri, one replication (University of Missouri), 5) North Central Regional Plant Introduction Station, Ames, IA, two replications (Iowa State University, planting date 2-June 3, 2013), and 6) Burkey Farm, Ames, Iowa, three replications (Iowa State University). Within each replication, B73 (Russell, 1972), Mo17 (Zuber, 1973), PHZ51, and PHB47 were used as inbred checks. The checks were included a minimum of six times per replication and as many as nine times, totaling 24 to 36 checks per replication. The R (R Core Team, 2014) package agricolae (de Mendiburu, 2015) was used to design the α incomplete block design.

During flowering time, dates were recorded when 50% of the plants in a row were shedding pollen on 50% of the main tassel branch. Silking dates were recorded when 50% of the plants in a row had visible silks emerging from the ear shoot. Days to shedding and silking were then converted to growing degree units (GDUs), calculated as $(T^{max} + T^{min}) / 2 - T^{base}$, where T^{max} is the maximum daily temperature up to 86°, at which point, T^{max} is set to 86°, T^{min} is the minimum daily temperature down to 50° degrees, at which point it is set to 50°, and T^{base} is 50°. All temperatures are in degrees Fahrenheit. The ASI is the difference, in GDUs, between anthesis GDUs (GDUSHD) and silking GDUs (GDUSLK), and is abbreviated GDUASI. The ratio of GDUSHD to GDUSLK (RSHDSLK) was calculated by dividing GDUSHD by GDUSLK and is an additional measure of ASI.

After pollination, PH and EH data were recorded. Plant height is recorded as the distance from the base of the ground to the ligule of the flag leaf. Ear height is measured as the distance from the base of the ground to the stalk node at which the highest ear has emerged. Three uniform, representative plants from the middle of each row were used in all locations, except Illinois, where a single representative plant was used. Rows lacking uniformity were noted and removed from further analysis. Both measurements were taken in centimeters (cm). Measurements from the three plants were averaged for further analysis. Above ear plant height (AEPH) was calculated by subtracting EH from PH. The ratio of ear height to plant height (REHPH) was also calculated by dividing EH by PH.

Statistical Analysis

Phenotypic analyses were conducted using a mixed linear model (MLM) in SAS 9.2 (SAS Inst. Inc., Cary, NC) as described by Wolfinger et al. (1997). Using the analysis described by Wolfinger et al. (1997) allows for the recovery of incomplete blocking and inter-variety information when the blocking and varieties are random effects. Model parameters were estimated using Henderson's mixed model equations and variances components were estimated using the restricted maximum likelihood method.

First, a MLM was fit to the data from each location. Replications and check inbreds were considered fixed effects. Random effects included the BGEM lines and incomplete blocks. Outliers were then identified and removed. Outliers were identified using the studentized conditional residual and determining the 95th quantile, with Bonferroni correction, based on a *t*-distribution. Data falling outside the defined regions were removed from further analysis. Data were assembled to include all locations once outliers were removed. These data were fit to a MLM containing location, replications, incomplete blocks, and BGEM lines as random effects. Replications were nested within location and incomplete blocks were nested within replications. Data files for the estimates of random and fixed effects were obtained from SAS 9.2 (SAS Inst. Inc.) and were used in R (R Core Team, 2014) to calculate best linear unbiased predictions (BLUPs) using a custom script. The custom script added the average BLUP of each germplasm enhancement of maize

doubled haploid (GEM-DH) line in each location and the BLUP of each GEM-DH line to the least squares mean estimate of all GEM-DH lines. Phenotypic trait correlations were calculated using Pearson's product-moment correlation coefficient (*r*) in package Hmisc (Harrell et al., 2016) in R (R Core Team, 2014) on trait BLUP values.

Broad-sense heritability, on an entry-mean basis, was calculated from the equation:

$$h^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \frac{\hat{\sigma}_{gxe}^2}{e} + \frac{\hat{\sigma}_e^2}{re}},$$

where $\hat{\sigma}_g^2$ is the genotypic variance estimate, $\hat{\sigma}_{gxe}^2$ is the interaction of genotype by environment, $\hat{\sigma}_e^2$ is the error variance estimation, *r* is the number of replications, and *e* is the number of environments (Hallauer et al., 1988). Standard errors were estimated using Dickerson's approximation (Dickerson, 1969), as explained in Hallauer et al. (1988).

Genotyping

Genotyping of the exotic derived BGEM lines was performed using genotyping-by-sequencing (GBS) (Elshire et al., 2011) by the Cornell University Genomic Diversity Facility, and the Buckler Lab for Maize Genetics and Diversity performed the assembly process. Files obtained were then filtered to remove single nucleotide polymorphisms (SNPs) with greater than 25% missing data and those with minor allele frequencies less than 2.5%. The remaining SNPs were filtered further by removing SNPs that were in the same position based on their genetic map position. Map positions were based on the IBM (Lee et al., 2002) genetic map, described by Wei et al. (2007). Within these regions, a single random SNP was chosen for the final genotypic file, which resulted in 62,077 SNPs being retained.

Sanchez et al. (unpublished data, 2017) observed that the proportion of recurrent parent within the BGEM lines was found to be higher than the expected 75%, based on SNP data, and that recombination rates were also higher than expected. To correct these issues, Sanchez et al. (unpublished data, 2017) used a method called Monomorphic Marker Correction (MMC), as described by Rice et al. (unpublished data, 2017) to correct for monomorphic markers within large donor parent segments. If we view a single SNP marker from an individual of the BGEM population and the SNP marker allele present is from the recurrent parent, but is flanked by donor parent alleles, the resulting sequence of these SNP markers would be D-R-D (Donor-Recurrent-Donor) for the individual when considering which parent the allele was inherited from. To determine if the middle SNP marker allele is from the donor parent and that is monomorphic with the allele from the recurrent parent, Bayes' theorem is used to calculate the probability, using the equation:

$$P\{M|DRD\} = \frac{P\{DRD|M\}P\{M\}}{P\{DRD|M\}P\{M\} + P\{DRD|\bar{M}\}P\{\bar{M}\}},$$

where $P\{DRD \mid M\}$ is the event M that the middle marker is a donor parent allele that is monomorphic with the recurrent parent allele, thus lending no information to the number of crossover events and is calculated as:

$$P\{DRD|M\} = \left(\frac{1-r}{2}\right)^2,$$

where r is the probability of an odd number of recombination events occurring between the first and third markers. $P\{M\}$ is the probability that the middle marker has a donor marker allele that is monomorphic to with the recurrent parent and is calculated as:

$$P\{M\} = \max\left\{\frac{\hat{p} - 0.75}{0.25}, 0\right\},$$

where $\hat{p}$ is the observed proportion of alleles that are recurrent parent at this marker. If this number exceeds the theoretical 0.75 proportion, the numerator of the fraction denotes the proportion of donor alleles that are monomorphic with the recurrent parent. If $\hat{p}$ is less than or equal to 0.75, then none of the donor alleles are monomorphic to the recurrent parent, thus $P\{M\} = 0$. $P\{\bar{M}\}$ is calculated as:

$$P\{\bar{M}\} = 1 - P\{M\}.$$

Finally, $P\{DRD \mid \bar{M}\}$ is calculated as:

$$P\{DRD|\bar{M}\} = \frac{1}{2} r_1 r_2 (1-r_1)(1-r_2),$$

where r_1*is* the probability of an odd number of crossover events between the first and second markers and r_2 is the probability of an odd number of crossovers between the second and third markers.

The equation P{*M*|*DRD*} can be expanded to allow for sequence segments that contain additional markers than the given example.

After this correction, recurrent parent portions were found to be closer to the expected 75% and recombination rates were greatly reduced. This corrected dataset was used for genome-wide association analysis (GWAS) for the flowering and height-related traits.

Genome-wide Association Studies

Using the genotypic dataset obtained from Sanchez et al. (unpublished data, 2017), all 62,077 SNP markers were used to calculate linkage disequilibrium (LD) between SNP markers using TASSEL 5.0 (Bradbury et al., 2007). R (R Core Team, 2014) was used to fit a smooth line to the data using local regression (LOESS). R package (R Core Team, 2014) GAPIT (Lipka et al., 2012) was used to estimate population structure based on all 62,077 SNPs using principal component analysis (PCA). Scree plots were

used to visualize the amount of variance explained by each principal component. The point at which the curve in the scree plot flattens was deemed the proper number of principal components to retain (Lipka et al., 2012).

A GWAS was performed using trait BLUPs for 232 BGEM lines. Twenty lines were discarded due to segregation noted in field trials, missing genotypic data or heterozygosity noted in the genotypic data. Ninety-eight BGEM lines represented the PHZ51 NSS group, while 134 represented the PHB47 SS group. Three analytical softwares were used in the GWAS analysis: 1) TASSEL 5.0 (Bradbury et al., 2007), 2) GAPIT (Lipka et al., 2012), and 3) FarmCPU (Liu et al., 2016). TASSEL 5.0 (Bradbury et al., 2007) was used to conduct a general linear model (GLM) approach, that also included PCA results from GAPIT as a fixed effect covariate (Lipka et al., 2012) to account for population structure. A MLM (Yu et al., 2006), that included the PCA results as fixed effects to control for population structure and an additive genetic relatedness matrix (VanRaden, 2008) was used to estimate the variance-covariance between individuals, that was included in the model as a random effect, was conducted in GAPIT (Lipka et al., 2012). Finally, FarmCPU (Liu et al., 2016) included PCA results for population structure as a covariate, kinship to account for relatedness among individuals as an additional covariate (VanRaden, 2008), and additional algorithms that aid in solving the confounding problem between testing markers and covariates. The purpose of using three different models was to reduce the chances of committing type 1 and type 2 errors. In each of the three models, a familywise error rate obtained from simpleM (Gao et al., 2010) was obtained using R (R Core Team, 2014). SimpleM calculates the effective number of independent tests, M_{eff} (Cheverud, 2001; Gao et al., 2008; Li and Ji, 2005; Moskvina and Schmidt 2008; Nyholt, 2004), by using composite LD among SNPs to capture the correlation and derives M_{eff} using the number of principal components that contribute to 99.5% of variation (Gao et al., 2010). Bonferroni correction is then conducted using the M_{eff} estimate. The resulting significance threshold was set at 3.17×10^{-6}. Quantile-quantile (QQ) plots were obtained using package qqman (Turner, 2014) in R (R Core Team, 2014). Comparisons were made between QQ plots to identify which GWAS method best fit each of the flowering and height related traits.

Significant SNPs identified in GWAS were compared to previously identified QTL regions. The QTL regions were defined by determining the placement of the markers in the reported QTL interval on the B73 RefGen_v2 (Schnable et al., 2009) map. The SNPs from the GWAS dataset were scanned and SNPs falling within QTL regions were identified. The LD was computed for these SNPs and the SNP identified through GWAS. The LD was calculated in TASSEL 5.0 (Bradbury et al., 2007). A threshold of $r^2 = 0.2$ was used to determine if the SNP identified through GWAS was in a novel region or overlapped with existing QTL regions. Candidate genes were identified using MaizeGDB (Andorf et al., 2015). Candidate genes were considered if significantly associated SNP fell within regions of the candidate gene as defined by B73 RefGen_v2 (Schnable et al., 2009) or were within 1 Mb of the identified significant SNP.

RESULTS

Phenotypic Analysis and Trait Correlations of BGEM Lines

The BGEM panel used in this study displayed significant variation for all studied flowering and height traits. Estimates for genotype and the interaction between genotype and location were found to be significant for all traits ($P < 0.0001$) based on MLM results. Details of the results for the MLM parameter fit to the data are shown in Supplemental Table S2. The BLUPs had large ranges for flowering traits that equated to approximately a 10-d span for GDUSHD and GDUSLK and approximately a four-d span for ASI, assuming 36 GDUs per day. PH and EH BLUPs ranged 111.7 cm and 67.5 cm, respectively. Heritability estimates were found to be high across all traits. The lowest heritability of 0.88 was found for REHPH and the highest of 0.94 was found for GDUSLK. Mean, minimum, maximum, range, standard deviation, and heritability estimates are listed in Table 1. Supplemental Table S4 shows BLUP values for 232 BGEM lines, as well as the two recurrent parents used for DH line development.

Table 2 shows all Pearson's product-moment correlation coefficients (r) between all flowering and height traits. Traits RSHDSLK and ASI were found to have the closest correlation value with $r = 0.994$. Traits GDUSHD and GDUSLK were found to be significantly ($P < 0.001$) and positively correlated with $r = 0.866$. Trait GDUSLK was found to be significantly correlated with all other studied traits. Trait GDUSHD was not significantly correlated with ASI or RSHDSLK. All height traits were significantly ($P < 0.001$) correlated with other height traits, with correlation between REHPH and AEPH being the only negative correlation. Plant height and EH were significantly ($P < 0.001$) correlated with GDUSHD and GDUSLK.

Population Structure

Two sub-populations were found in the BGEM lines, which is consistent with having two recurrent parents. However, some BGEM lines did not correspond to the correct sub-population. The sub-population corresponding to PHZ51, or NSS BGEMs, contained 112 BGEM lines with 23 having PHB47 as the recurrent parent. These lines included BGEM-0007-S, BGEM-0017-S, BGEM-0052-S, BGEM-0053-S, BGEM-0076-S, BGEM-0078-S, BGEM-0092-S, BGEM-0094-S, BGEM-0098-S, BGEM-0111-S, BGEM-0112-S, BGEM-0113-S, BGEM-0114-S, BGEM-0115-S, BGEM-0116-S, BGEM-0117-S, BGEM-0118-S, BGEM-0165-S, BGEM-0166-S, BGEM-0171-S, BGEM-0175-S, BGEM-0189-S, and BGEM-0220-S. The other sub-population that corresponded to PHB47, or SS BGEMs, contained 120 BGEM lines with nine having PHZ51 as recurrent parent. The nine lines were BGEM-0005-N,

Table 1. Summary statistics of flowering and plant architecture traits for PHZ51 derived, PHB47 derived, and combined PHZ51/PHB47 derived BGEM lines.

Group†	Trait	RP‡§	Mean¶	Minimum¶	Maximum¶	Range¶	SD‡¶	h^2 (SE)
PHZ51 derived BGEM Lines	GDUSHD	1481.6	1545.5	1407.6	1712.1	304.4	60.2	–
	GDUSLK	1507.7	1591.8	1426.2	1764.3	338.1	74.5	–
	ASI	-28.3	–46.3	-121.1	21.8	142.8	33.7	–
	RSHDSLK	1.0	1.0	0.9	1.0	0.1	0.02	–
	PH	181.8	182.0	143.2	246.6	103.4	16.0	–
	EH	66.4	72.1	50.3	110.9	60.6	10.2	–
	AEPH	115.6	110.0	89.8	146.3	56.5	10.1	–
	REHPH	0.4	0.4	0.3	0.5	0.2	0.03	–
PHB47 derived BGEM Lines	GDUSHD	1473.7	1528.9	1368.1	1718.7	350.6	63.6	–
	GDUSLK	1488.4	1557.8	1409.0	1719.6	310.6	61.9	–
	ASI	-14.6	-29.9	-120.4	21.1	141.5	26.4	–
	RSHDSLK	1.0	1.0	0.9	1.0	0.1	0.02	–
	PH	179.9	179.3	134.9	221.9	87.0	14.8	–
	EH	63.0	68.9	43.4	96.1	52.8	9.3	–
	AEPH	117.1	110.7	90.9	137.7	46.7	8.9	–
	REHPH	0.4	0.4	0.3	0.5	0.2	0.03	–
Combined	GDUSHD	–	1535.9	1368.1	1718.7	350.6	62.6	0.93 (0.09)
	GDUSLK	–	1572.2	1409.0	1764.3	355.3	69.4	0.94 (0.09)
	ASI	–	-36.8	-121.1	21.8	142.8	30.7	0.90 (0.09)
	RSHDSLK	–	1.0	0.9	1.0	0.1	0.02	0.90 (0.09)
	PH	–	180.4	134.9	246.6	111.7	15.3	0.90 (0.09)
	EH	–	70.2	43.4	110.9	67.5	9.8	0.88 (0.09)
	AEPH	–	110.4	89.8	146.3	56.5	9.4	0.87 (0.09)
	REHPH	–	0.4	0.3	0.5	0.2	0.03	0.83 (0.09)

† Group, PHZ51 derived BGEM Lines, corresponds to the 98 BEM lines with PHZ51 as the recurrent parent. PHB47 derived BGEM Lines corresponds to the 134 BGEM lines with PHB47 as the recurrent parent. Combined, 232 BGEM lines were used in GWAS.

‡ RP, recurrent parent; SD, standard deviation.

§ RP corresponds to groups' respective recurrent parent. Values listed are trait least-square means of recurrent parent.

¶ Values are estimated from trait BLUPs of n lines within each group. n = 98, PHZ51; n = 134, PHB47.

Table 2. Pearson's product-moment correlation coefficient (*r*) of flowering and plant architecture traits from 232 BGEM lines.

Trait	GDUSHD	GDUSLK	ASI	RSHDSLK	PH	EH	AEPH	REHPH
GDUSHD								
GDUSLK	0.87**							
ASI	0.03	-0.43**						
RSHDSLK	0.09	-0.38**	0.99**					
PH	0.24**	0.25**	-0.07	-0.05				
EH	0.30**	0.28**	0.00	0.03	0.79**			
AEPH	0.07	0.11**	-0.11	-0.10	0.76**	0.19*		
REHPH	0.24**	0.20**	0.06	0.08	0.28**	0.81**	-0.41**	

*Significant at $P < 0.01$

**Significant at $P < 0.001$.

BGEM-0085-N, BGEM-0107-N, BGEM-0129-N, BGEM-0132-N, BGEM-0215-N, BGEM-0227-N, BGEM-0232-N, and BGEM-0248-N. The misclassification within sub-populations is in accordance with the findings of Hu (2016) and is most likely due to the exotic germplasm contribution of the BGEM lines.

Investigation of QQ Plots

The QQ plots obtained from the analyses used in GWAS (Fig. 1; Supplemental Figures S1 through S7) showed substantial differences in the performance of the three methods. Among the flowering and height traits, each analysis method performed differently. For example, observed $-\log_{10}(p)$ values were overestimated when compared to the corresponding expected values for GLM+PCA approach for GDUSHD (Fig. 1A). Mixed linear model underestimated $-\log_{10}(p)$ values for GDUSHD (Fig. 1B), but the underestimation was less extensive than the overestimation by GLM+PCA. FarmCPU overestimated $-\log_{10}(p)$ values for GDUSLK, while underestimating for GDUSHD (Fig. 1C) and all other traits. The QQ plots are shown for all traits in Fig. 1, Supplemental Figures S1 through S7.

GWAS Summary for Flowering and Height Traits

The LD decayed ($r^2 = 0.2$ threshold) over a distance greater than 500 kb for all chromosomes. However, LD within the BGEM population is two-fold. The BGEM lines within either of the recurrent parent backgrounds are 75% related, resulting in the high levels of linkage within lines sharing a common recurrent parent, causing

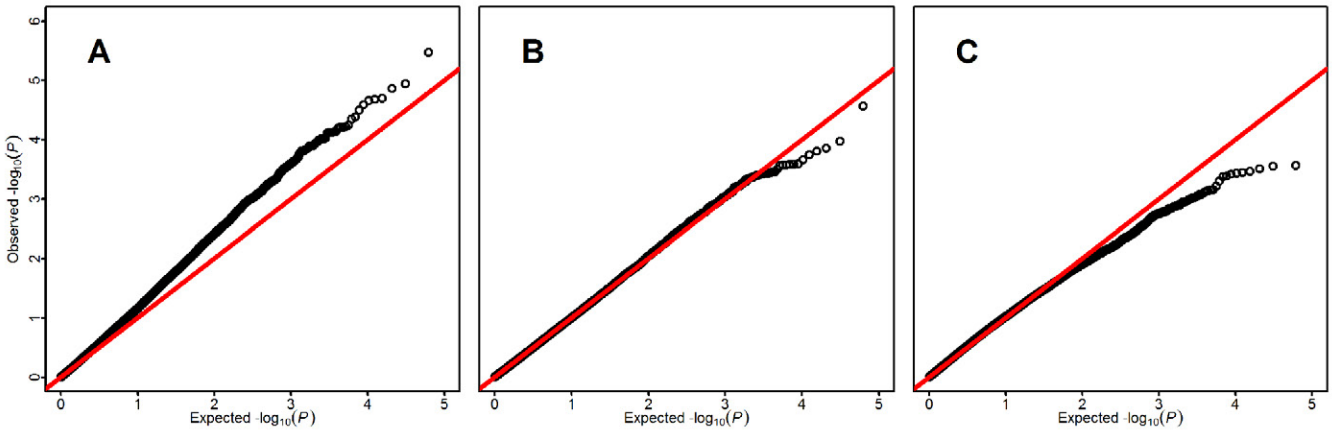

Fig. 1. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUSHD.

the large decay distances. It is expected, however, that genome regions containing exotic, donor parent genome to decay at distances around 1kb, as previously described for exotic landraces by Romay et al. (2013).

Significant marker-trait associations ($p = 3.17 \times 10^{-6}$ threshold calculated with simpleM [Gao et al., 2010]) were found in flowering and height traits. The GLM + PCA identified 2 SNPs for RSHDSLK, 3 SNPS for ASI, 12 SNPs for PH, 30 SNPs for EH, 6 SNPs for AEPH, and 3 SNPs for REHPH. FarmCPU identified 7 RSHDSLK and 6 GDUSLK SNP associations. No associations were found for GDUSHD across the three methods tested and the MLM method found no associations for any trait. Trait GDUSLK associations were found on chromosomes 1, 2, 4, 7, and 10, with two associations on chromosome 7. The ASI and RSHDSLK associations were found on chromosomes 1 and 2 and chromosomes 1, 2, 3, 7, and 9, respectively. Height associations were found on all chromosomes, except chromosomes 8 and 9. Chromosomes 2 and 3 contained the highest number of associations for height traits.

The SNP *S1_15995431* and *S2_60273238* were identified as significant associations for ASI and RSHDSLK. Two SNPs, *S2_236291252* and *S2_236484950*, were both found to be associated for traits EH and REHPH. *S2_7657546*, *S3_135695490*, *S3_138773405*, *S3_160516096*, and *S4_235711113* were identified for traits PH and EH as significant associations. No SNPs were found in common between flowering and height traits. All significant SNP markers, with effect estimates, are listed in Supplemental Table S3. Manhattan plots for all traits are shown in Supplemental Figures S8 through 15.

Previously identified QTL regions were compared to significantly associated SNPs. Twenty-six regions found by Li et al. (2016b) were compared to flowering associations. Nine of these regions contained SNPs within the QTL region. LD between the remaining 17 regions and nearby associated SNP was found to be from $r^2 = 0.05$ to 0.30 (Table 3). Sixteen QTL regions from Peiffer et al. (2014) were used for comparisons with identified height trait associations. Seven identified SNPs fell outside the QTL regions and LD was found to be from $r^2 = 0.19$ to 0.33. Results of height comparisons to QTL are shown in Table 4.

Fifteen flowering time associated SNPs were found to fall within gene regions. Height trait associations were found within regions for 29 genes across all height related traits. Flowering and height candidate genes are listed in Table 5 and Table 6, respectively. Significantly associated SNPs were also compared to previously published candidate genes. Nineteen candidate genes were found to be within 1 Mb of associated SNP for flowering traits, as well as 18 candidate genes for height related traits. Flowering candidate genes identified include four *APETALA2*-like genes and height candidate genes include various auxin and gibberellin pathway candidate genes.

DISCUSSION

Phenotypic data used in our study were found to have high heritabilities across all studied traits ranging from 0.83 to 0.94. These findings are in agreement with reports from previous studies (Buckler et al., 2009; Romay et al., 2013; Li et al., 2016b). Traits GDUSHD and GDUSLK had ranges of 350 GDUs, or 11 to 12 d; however, all lines flowered within a normal timeframe expected for the central U.S. Corn Belt. Therefore, the backcrossing method used to adapt these lines was effective, as indicated by flowering traits. Many correlations between flowering and height traits were found to be significant, indicating that adaptation based on flowering time is a viable option in altering other important adaptation related traits.

LD between Identified SNPs with Previous QTL Studies

Tables 3 and 4 show the 26 previously identified QTLs from Li et al. (2016b) for flowering traits and 19 QTL from Peiffer et al. (2014) that were compared to SNPs identified in the current study, respectively. Three classes of outcomes were noted: 1) GWAS identified SNP was located within QTL region, 2) QTL region did not contain GWAS identified SNP but was in LD with GWAS identified SNP, and 3) GWAS identified SNP was located

Table 3. Linkage disequilibrium between significant SNPs and previously identified QTLs from Li et al. (2016b) for maize flowering traits. Only significant SNPs within 5 Mb of QTL were considered.

			Previously identified region					
Trait	Chr	QTL	Start (Mb)†	End (Mb)†	SNP	Distance (Mb)	N‡	LD (r^2)
GDUSLK	1	CN_QTL_01	2.73	4.98	S1_4749033	–	–	–
	1	USA_QTL_01	0	4.75	S1_4749033	0.03	257	0.10
	2	USA_QTL_13	0	4.69	S2_6177366	1.49	505	0.22
	2	USA_QTL_14	10.53	15.68	S2_6177366	4.35	469	0.15
	4	USA_QTL_34	2.70	6.11	S4_5707348	–	–	–
	7	USA_QTL_59	140.41	145.74	S7_149363044	3.62	230	0.30
	7	USA_QTL_60	150.03	156.21	S7_149363044	0.67	307	0.30
	7	USA_QTL_60	150.03	156.21	S7_158112134	1.90	307	0.29
	7	CN_QTL_40	156.74	160.03	S7_158112134	–	–	–
	10	USA_QTL_75	142.61	145.09	S10_146119225	1.02	262	0.14
ASI	1	CN_QTL_02	10.84	15.16	S1_15995431	0.83	259	0.19
	1	USA_QTL_02	8.62	14.85	S1_15995431	1.14	372	0.17
	1	CN_QTL_04	48.55	64.89	S1_53996436	–	–	–
	1	USA_QTL_04	52.25	68.93	S1_53996436	–	–	–
	2	CN_QTL_13	51.44	63.94	S2_60273238	–	–	–
	2	USA_QTL_17	60.66	86.00	S2_60273238	0.39	294	0.20
RSHDSLK	1	CN_QTL_02	10.84	15.16	S1_15995431	0.83	259	0.19
	1	USA_QTL_02	8.62	14.85	S1_15995431	1.14	372	0.17
	1	USA_QTL_05	8.22	103.39	S1_107308297	3.92	308	0.26
	2	CN_QTL_13	51.44	63.94	S2_60273238	–	–	–
	2	USA_QTL_17	60.66	86.00	S2_60273238	0.39	294	0.20
	3	CN_QTL_20	111.57	128.65	S3_118163262	–	–	–
	7	CN_QTL_37	4.84	8.70	S7_10572080	1.87	216	0.25
	7	CN_QTL_37	4.84	8.70	S7_11335522	2.63	216	0.05
	7	CN_QTL_40	156.74	160.03	S7_163923782	3.89	239	0.25
	9	USA_QTL_69	122.95	132.03	S7_125795791	–	–	–

† Based on B73 RefGen_v2.

‡ Number of SNP markers that fell within QTL region used to estimate LD. When significant SNP was located within QTL region, LD was not calculated.

outside a QTL region and was not in LD with the QTL. When a GWAS identified SNP was located within a QTL region, SNP and QTL were co-localized and thus, the SNP region was deemed overlapping with the previously identified QTL region. The SNPs that fell outside of a QTL region but were in LD with SNPs located within the QTL region were concluded to be overlapping regions. However, novel regions could be declared when a SNP was located outside of the QTL region and was not in LD with the QTL. Nine significant SNPs were found within previously defined QTL regions for both flowering and height related traits. In addition, nine and five SNPs were found to be in LD with defined QTL regions for flowering and height traits, respectively. Finally, eight flowering and two height SNPs were outside defined QTL regions and lacked LD to the QTL region. Since linkage and LD mapping studies rely heavily on the choice of germplasm included, these final significant SNPs could be novel regions effecting flowering and height related traits.

Flowering Trait Associations

Across the three GWAS methods used, 18 SNPs were identified for flowering traits (Supplemental Table S3). Seventeen of these 18 SNPs were found to be located within regions of candidate genes, shown in Table 5. However, due to the abundance of flowering candidate genes, it is likely that an identified SNP is simply in LD with a nearby known flowering gene. For this reason, we viewed 1 Mb around identified SNPs (Li et al., 2016b) and compared to the region previously identified candidate genes listed in Li et al. (2016b) and identified by Chen et al. (2012), Danilevskaya et al. (2008), and Dong et al. (2012). This collection of candidate genes compiled by Li et al. (2016b) was comprised of 919 flowering time related candidate genes or homologs. From flowering trait related SNPs listed in Supplemental Table S3, 17 candidate genes fell within the 1 Mb window threshold (Table 7). A nuclear transcription factor y subunit a1 known as *nfya1* in maize, *GRMZM2G000686*, was identified by ASI and RSHDSLK SNP *S1_15995431*. The ASI and RSHDSLK are both measures of anthesis to silking interval and are highly correlated with $r = 0.994$. *GRMZM2G021416* was identified by *S7_10572080* and *S7_11335522* for trait RSHDSLK. A homolog of *Arabidopsis thaliana WNK5*, *GRMZM2G021416*, which is a serine/threonine-protein kinase, has unknown function in maize but was predicted as a flowering time related gene by Chen

Table 4. Linkage disequilibrium between significant SNPs and previously identified QTLs from Peiffer et al. (2014) for maize height traits. Only significant SNPs within 5 Mb of QTL were considered.

Trait	Chr	Previously identified region†		SNP	Distance (Mb)	N§	LD (r^2)
		Start (Mb)‡	End (Mb)‡				
PH	2	0	5	S2_7657546	2.66	531	0.19
	2	0	7	S2_7657546	0.66	777	0.20
	3	156	164	S3_160516096	–	–	–
	4	231	239	S4_235711113	–	–	–
	7	148	156	S7_155356223	–	–	–
	7	151	159	S7_155356223	–	–	–
EH	1	245	253	S1_244512817	0.49	318	0.28
	2	0	7	S2_7657546	0.66	777	0.20
	2	0	7	S2_7958989	0.96	531	0.19
	3	156	164	S3_160516096	–	–	–
	4	231	239	S4_235711113	–	–	–
	7	139	147	S7_144205699	–	–	–
	7	148	156	S7_144205699	3.79	348	0.33
	7	148	156	S7_154433671	–	–	–
	7	151	159	S7_154433671	–	–	–
REHPH	1	79	87	S1_88000915	1.00	172	0.20

† QTL regions were not defined by Peiffer et al. (2014). To define regions, 4 Mb up- and downstream of the QTL position was used.

‡ Based on B73 RefGen_v2.

§ Number of SNP markers that fell within QTL boundaries used to estimate LD. When significant SNP was located within QTL region, LD was not calculated.

et al. (2012). Ten of the candidate genes, *GRMZM2G082227*, *GRMZM2G327059*, *GRMZM2G000686*, *GRMZM2G174784*, *GRMZM2G106548*, *GRMZM2G147716*, *GRMZM2G384528*, *GRMZM2G142999*, *GRMZM2G076602*, and *GRMZM2G058588*, are transcription factors that are likely to be controlling the expression of other flowering genes. *GRMZM2G033885*, which has a role in photosystem II, was identified on chromosome 7. Khan et al. (2010) identified *GRMZM2G033885* as playing a role in the circadian clock of maize, most likely regulating the start and stop of photosystem II. A candidate gene identified for EH, *GRMZM2G012546*, is also known to have effects on flowering time. Ear height was significantly correlated ($P < 0.001$) with GDUSHD and GDUSLK, supplying a reason for the overlap in identified candidate genes. *GRMZM2G012546* is a gibberellin receptor-like candidate that can affect *FT*-like gene expressions in maize and other crops (Osnato et al., 2012; Song et al., 2012; Li et al., 2016a).

In summary, many flowering time candidate genes were discovered for GDUSLK, ASI, and RSHDSLK traits. Some of these candidate genes are known to have an effect on flowering and height related traits. One candidate gene was identified for ear height that has known effects on *FT*-like flowering gene expressions.

Height Trait Associations

Supplemental Table S3 shows the 12 SNPs that were identified though association analyses for PH. Four of these SNPs were located within candidate genes (Table 6) *GRMZM2G427635*, *GRMZM2G121074*, *GRMZM2G101221*, and *GRMZM2G044460*. Most interestingly, *GRMZM2G121074* is the closest maize homolog to severe depolymerization of actin (*sad1*), and has been shown to affect cell number in the shoot apical meristem (SAM) in maize (Leiboff et al., 2015). The increased number of cells in the SAM could, therefore, affect overall PH. Five additional candidate genes (Table 8) were identified within 1 Mb of identified SNPS. Maize gene *GRMZM2G025742* (*pin5*) is a putative auxin efflux carrier that has been shown to be localized in the endoplasmic reticulum, suggesting that it plays a key role in intercellular auxin homeostasis (Mravec et al., 2009; Wabnik et al., 2011; Forestan et al.,

Table 5. Candidate genes for significant SNPs found in GWAS of flowering traits. Candidate genes were only considered when an identified SNP fell within gene regions.

Trait	Chr	Start†	B73 Gene ID	*Zm* Gene ID	Function‡
GDUSLK	2	6.17	GRMZM2G072850		ABC transporter B family member 19-like
	4	5.70	GRMZM2G011364		sec12-like protein 2
	7	149.36	GRMZM2G369340		
	7	158.11	GRMZM2G103276		
	10	146.12	GRMZM2G104260	ereb190	ethylene-responsive transcription factor ERF003-like
ASI	1	15.99	GRMZM2G156486		
	1	53.99	GRMZM2G124209		probable apyrase 1
	2	60.27	GRMZM2G157269		acetate—CoA ligase ACS, chloroplastic/glyoxysomal-like
RSHDSLK	1	15.99	GRMZM2G156486		
	1	107.30	GRMZM2G469795		
	2	60.27	GRMZM2G157269		acetate—CoA ligase ACS, chloroplastic/glyoxysomal-like
	3	118.16	GRMZM2G075900		H+-translocating pyrophosphatase
	7	10.57	GRMZM2G030138		
	7	11.34	GRMZM2G102075		40S ribosomal protein S15
	7	163.92	GRMZM2G370081		protein LONGIFOLIA 1-like

† Based on B73 RefGen_v2 in Mb.

‡ Obtained from MaizeGDB (www.maizegdb.org), NCBI (www.ncbi.nlm.nih.gov), and Gramene (www.gramene.org).

Table 6. Candidate genes for significant SNPs found in GWAS of height traits. Candidate genes were only considered when an identified SNP fell within gene regions.

Trait	Chr	Start†	B73 Gene ID	*Zm* Gene ID	Function‡
PH	1	163.27	GRMZM2G427635		probable L-gulonolactone oxidase 6
	2	7.66	GRMZM2G121074	sda1	protein SDA1 homolog
	3	138.77	GRMZM2G101221		peroxidase superfamily protein
	3	160.52	GRMZM2G044460		
EH	1	244.51	GRMZM2G150169		
	2	7.66	GRMZM2G121074	sda1	protein SDA1 homolog
	2	7.96	GRMZM2G155217		putative HLN DNA-binding domain superfamily protein
	2	223.89	GRMZM2G151227	whp1	chalcone synthase/white pollen 1
	2	223.89	GRMZM5G818346		
	2	234.03	GRMZM5G871297		
	2	234.80	GRMZM2G020805	tcptf43	TCP-transcription factor 43
	2	235.06	GRMZM2G409771		
	2	235.22	GRMZM2G178136		
	2	235.85	GRMZM2G148693	zap1	zea apetala homolog 1
	2	235.89	GRMZM2G143602	cka1	CK2 protein kinase alpha 1
	2	235.92	GRMZM2G048775		peroxidase 35-like
	2	236.29	GRMZM2G104032		probable fucosyltransferase 8
	2	236.48	GRMZM2G452121		L-type lectin-domain containing receptor kinase IV.1-like
	3	138.63	GRMZM2G087513		O-fucosyltransferase family protein
	3	138.77	GRMZM2G101221		peroxidase superfamily protein
	3	160.52	GRMZM2G044460		
	5	51.43	GRMZM2G177549		
	7	144.20	GRMZM2G098764		DNA-directed RNA polymerase III subunit 22.9 kDa polypeptide
	10	77.11	GRMZM2G171236		mitochondrial NADH ubiquinone oxidoreductase 13kD-like subunit
REHPH	1	88.00	GRMZM2G071172		
	2	236.29	GRMZM2G104032		probable fucosyltransferase 8
	2	236.48	GRMZM2G452121		L-type lectin-domain containing receptor kinase IV.1-like
AEPH	5	51.43	GRMZM2G177570		tyrosine-sulfated glycopeptide receptor 1-like
	5	112.63	GRMZM2G137046	bzip61	bZIP transcription factor 61

† Based on B73 RefGen_v2 in Mb.

‡ Obtained from MaizeGDB (www.maizegdb.org), NCBI (www.ncbi.nlm.nih.gov), and Gramene (www.gramene.org).

2012). Auxins play a pivotal role in the growth of maize by influencing many plant processes like SAM development, vascular elongation, lateral root initiation, embryogenesis, and flower and fruit development (Kriechbaumer et al., 2006; De Smet and Jürgens, 2007), with stems, branches, and lateral organs being the final products of the SAMs activity (Gallavotti, 2013). Additionally, two auxin response factor (ARF) candidate genes, *GRMZM2G056120* (*arftf11*) and *GRMZM2G028980* (*arftf16*), were found near associated PH SNPs. The ARFs specifically bind to auxin response elements (Liu et al., 2011; Wang et al., 2012), thus regulating auxin levels. A small auxin upregulated RNA (SAUR) family member, *GRMZM2G471304* (*saur45*), was also identified. The SAURs are the largest family of early auxin response genes, with 79 in maize (Chen et al., 2014). They regulate plant growth and development by mediating auxin's genomic response in the plant, through inhibiting phosphatases to activate plasma membrane H^+-ATPases and promote cell elongation (Ren and Gray, 2015). A gibberellin oxidase, *GRMZM2G022679* (*ga2ox3*) was found as a candidate gene effecting PH in the BGEM lines. Gibberellin oxidases regulate plant growth by inactivating endogenous bioactive gibberellins, thus altering plant height (Wang et al., 2013).

EH associated SNPs fell within the boundaries of 20 candidate genes (Table 6). Identified as a PH candidate gene, *GRMZM2G121074*, was also identified for EH. However, no other candidate genes identified by this method were known to have an effect on height related traits. An additional nine candidate genes were identified within 1 Mb of the EH associated SNPs (Table 8). Maize genes *GRMZM2G025742* (*pin5*), *GRMZM2G471304* (*saur45*), and *GRMZM2G028980* (*arftf16*) were in common with candidate genes identified for PH. Additionally, three auxin related candidate genes were found and included: *GRMZM2G077401*, an ARF, *GRMZM2G009103*, an auxin efflux carrier homologous to *Arabidopsis thaliana WAT1* (see Ranocha et al. (2013) for *WAT1* details), and *GRMZM2G095839* (*saur64*), a SAUR family member. A gibberellin receptor-like candidate gene has known effects for *FT*-like flowering gene expressions, *GRMZM2G012546*, was also identified. A gibberellin receptor GID1L2 *GRMZM5G817777*, (gibberellin-insensitive dwarf protein1 L2 family [Jiang et al., 2014; Li et al.,

Table 7. Candidate genes for significant SNPs found in GWAS of flowering traits. Candidate genes were considered when an identified SNPs were within 1 Mb of previously published candidate gene.§

Trait	Chr	Start†	B73 Gene ID	*Zm* Gene ID	Function‡
GDUSLK	1	4.55	GRMZM2G082227	abi8	ABI3-VP1-transcription factor 8
GDUSLK	1	5.09	GRMZM2G327059	hb30	homeobox-transcription factor 30
ASI	1	15.81	GRMZM2G000686	nfya1	nuclear transcription factor y subunit a1
RSHDSLK	1	15.81	GRMZM2G000686	nfya1	nuclear transcription factor y subunit a1
GDUSLK	2	5.51	GRMZM2G174784	ereb197	AP2/EREBP transcription factor 197
RSHDSLK	3	118.96	GRMZM2G444621		calmodulin-like protein 8
GDUSLK	4	5.96	GRMZM2G145213		14–3-3-like protein
RSHDSLK	7	11.57	GRMZM2G021416		probable serine/threonine-protein kinase/WNK5-like
RSHDSLK	7	11.57	GRMZM2G021416		probable serine/threonine-protein kinase/WNK5-like
GDUSLK	7	149.92	GRMZM2G134797	ndk1	nucleotide diphosphate kinase 1
GDUSLK	7	157.28	GRMZM2G033885	psb29	photosystem II subunit29
GDUSLK	7	157.57	GRMZM2G057281		chlorophyll a-b binding protein-LHCII type III
GDUSLK	7	158.37	GRMZM2G106548	gras54	GRAS-transcription factor 54
RSHDSLK	7	164.41	GRMZM2G147716	mads67	MADS-transcription factor 67
RSHDSLK	7	164.69	GRMZM2G384528	ca3p4	CCAAT-HAP3-transcription factor 34
RSHDSLK	7	164.77	GRMZM2G351482		probable carboxylesterase 18
RSHDSLK	9	125.56	GRMZM2G142999	abi38	ABI3-VP1-transcription factor 38
GDUSLK	10	145.35	GRMZM2G076602	ereb212	AP2-EREBP transcription factor 212
GDUSLK	10	146.29	GRMZM2G058588	sbp28	SBP-transcription factor 28

† Based on B73 RefGen_v2 in Mb.

‡ Obtained from MaizeGDB (www.maizegdb.org), NCBI (www.ncbi.nlm.nih.gov), Gramene (www.gramene.org).

§ Candidate genes were compared to the findings of Chen et al. (2012), Danilevskaya et al. (2008), Dong et al. (2012), Hung et al. (2012), and Li et al. (2016b).

Table 8. Candidate genes for significant SNPs found in GWAS of height traits. Candidate genes were considered when an identified SNPs were within 1 Mb of previously published candidate gene.§

Trait	Chr	Start†	B73 Gene ID	*Zm* Gene ID	Function‡
EH	2	234.10	GRMZM2G012546		probable carboxylesterase 15
EH	2	234.11	GRMZM5G817777		gibberellin receptor GID1L2
EH	2	234.43	GRMZM2G021051	ga20ox2	gibberellin 20-oxidase 2
AEPH	2	277.64	GRMZM2G006964	ga2ox2	gibberellin 2-oxidase 2
AEPH	2	279.80	GRMZM2G122614	arftf6	auxin response factor-transcription factor 6
AEPH	2	281.52	GRMZM2G089806		auxin-induced protein 15A-like
PH	3	160.80	GRMZM2G025742	pin5	auxin efflux carrier family protein
EH	3	160.80	GRMZM2G025742	pin5	auxin efflux carrier family protein
PH	3	196.06	GRMZM2G022679	ga2ox3	gibberellin 2-oxidase 3
PH	3	196.57	GRMZM2G056120	arftf11	auxin response factor-transcription factor 11
EH	3	208.54	GRMZM2G077401		putative auxin-response protein
PH	4	236.38	GRMZM2G471304	saur45	SAUR45-auxin responsive SAUR family member
EH	4	236.38	GRMZM2G471304	saur45	SAUR45-auxin responsive SAUR family member
PH	4	236.47	GRMZM2G028980	arftf16	auxin response factor-transcription factor 16
EH	4	236.47	GRMZM2G028980	arftf16	auxin response factor-transcription factor 16
AEPH	5	101.59	GRMZM2G379490		auxin induced protein 15A-like
EH	7	144.30	GRMZM2G095839	saur64	SAUR64-auxin responsive SAUR family member
EH	7	153.85	GRMZM2G009103		WAT1-related protein/auxin efflux

† Based on B73 RefGen_v2 in Mb.

‡ Obtained from MaizeGDB (www.maizegdb.org), NCBI (www.ncbi.nlm.nih.gov) and Gramene (www.gramene.org).

§ Candidate genes were compared to the findings of Song et al. (2012), Teng et al. (2013), and Wallace et al. (2016).

2016a]) was found on chromosome 2. A gibberellin oxidase, *GRMZM2G021051* (*ga20ox2*), was also found.

The SNPs associated with AEPH and REHPH were found to be in the boundaries of two and three candidate genes, respectively (Table 6); however, these genes have not been linked to height related traits. Four additional height related candidate genes were found within 1 Mb associated SNPs for trait AEPH (Table 8) but no additional

candidate genes were found for REHPH. Two auxin induced protein candidate genes, *GRMZM2G089806* and *GRMZM2G379490*, were identified. Auxin enhances the binding of auxin induced proteins to receptor proteins (Maraschin et al., 2009), thus regulating receptor protein expressions and altering plant height. ARF, *GRMZM2G122614* (*arftf6*), was also identified as an auxin candidate gene affecting AEPH. Gibberellin oxidase, *GRMZM2G122614* (*ga2ox2*), was the only gibberellin related candidate gene identified for AEPH.

In summary, candidate genes for four height related traits were discovered on chromosomes 1, 2, 3, 4, 5, 7, and 10. Many of these candidate genes were related to plant hormones auxin and gibberellin.

CONCLUSIONS

Flowering and height traits are two of the most studied traits in maize. In this study, we used a diverse panel of exotic derived DH lines, representing 52 exotic maize races, to investigate flowering and height traits to aid in the understanding of adapting exotic germplasm to the U.S. Corn Belt. Nineteen and eighteen flowering and height related candidate genes, respectively, were found, with an additional 46 candidate genes identified from associated SNPs falling within candidate gene boundaries. Future work is required to validate or disprove these 46 candidate genes as potential flowering and height related genes. Future work should also include the investigation of candidate genes similar to *GRMZM2G012546*, as the effects of similar genes could alter flowering time and alter height related traits. Additionally, validation is needed for SNP markers found in the association analyses. After validation, breeding programs, such as the USDA GEM project, could use these markers as a selection tool to speed the adaptation of exotic germplasm to the central U.S.

Supplemental information

Supplemental Table S1. BGEM code, exotic donor race and accession number, country of origin, and elevation for 252 BGEM lines.

Supplemental Table S2. Variance estimates of G, GxE, and error of flowering and plant architecture traits for 252 BGEM lines.

Supplemental Table S3. List of all significant markers found using three different association mapping analyses for flowering and plant architecture traits.

Supplemental Table S4. Trait BLUPs of flowering and height traits for 232 BGEM lines and recurrent parents.

Supplemental Figure S1. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUSLK.

Supplemental Figure S2. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUASI.

Supplemental Figure S3. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for RSHDSLK.

Supplemental Figure S4. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for PH.

Supplemental Figure S5. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for EH.

Supplemental Figure S6. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for AEPH.

Supplemental Figure S7. Comparison of QQ plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for REHPH.

Supplemental Figure S8. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUSHD.

Supplemental Figure S9. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUSLK.

Supplemental Figure S10. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for GDUASI.

Supplemental Figure S11. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for RSHDSLK.

Supplemental Figure S12. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for PH.

Supplemental Figure S13. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for EH.

Supplemental Figure S14. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for AEPH.

Supplemental Figure S15. Manhattan plots of (A) GLM+PCA, (B) MLM, and (C) FarmCPU for REHPH.

Conflict of interest disclosure

The authors declare that there is no conflict of interest.

Acknowledgments

The authors would like to thank the staff of the Iowa State Doubled Haploid Facility and the USDA-ARS Germplasm Enhancement of Maize project. Without the aid of these two groups, completing this work would not have been possible. This work was supported by the Raymond F. Baker Center for Plant Breeding, Iowa State University, Iowa State University Doubled Haploid Facility, the USDA-ARS Germplasm Enhancement of Maize Project, and the Hatch Multi-State Project NC007.

References

Andorf, C.M., E.K. Cannon, J.L. Portwood, J.M. Gardiner, L.C. Harper, M.L. Schaeffer, B.L. Braun, D.A. Campbell, V.A. Vinnakota, V.V. Sribalusu, M. Huerta, K. Tak Cho, K. Wimalanathan, J.D. Richter, E.D. Mauch, B.S. Rao, S.M. Birkett, T.Z. Sen, and C.J. Lawrence-Dill. 2015. MaizeGDB update:new tools, data and interface for the maize model organism database. Nucleic Acids Res. 44:195–201.

Blackman, B.K., D.A. Rasmussen, J.L. Strasburg, A.R. Raduski, J.M. Burke, S.J. Knapp, S.D. Michaels, and L.H. Rieseberg. 2011. Contributions of flowering time genes to sunflower domestication and improvement. Genetics 187:271–287. doi:10.1534/genetics.110.121327

Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez, M. Nordborg, J. Bergelson, J. Cuguen, and F. Roux. 2010. Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS Genet. 6:e1000940. doi:10.1371/journal.pgen.1000940

Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, and E.S. Buckler. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. doi:10.1093/bioinformatics/btm308

Brenner, E.A., M. Blanco, C. Gardner, and T. Lübberstedt. 2012. Genotypic and phenotypic characterization of isogenic doubled haploid exotic introgression lines in maize. Mol. Breed. 30:1001–1016. doi:10.1007/s11032-011-9684-5

Buckler, E.S., J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J.C. Glaubitz, M.M. Goodman, C. Harjes, K. Guill, D.E. Kroon, S. Larsson, N.K. Lepak, H. Li, S.E. Mitchell, G. Pressoir, J.A. Peiffer, M.O. Rosas, T.R. Rocheford, M.C. Romay, S. Romero, S. Salvo, H.S. Villeda, H.S. de Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich, and

M.D. McMullen. 2009. The genetic architecture of maize flowering time. Science 325:714–718. doi:10.1126/science.1174276
Chen, C., G. DeClerck, F. Tian, W. Spooner, S. McCouch, and E.S. Buckler. 2012. PICARA, an analytical pipeline providing probabilistic inference about a prior candidates genes underlying genome-wide association QTL in plants. PLoS One 7:e46596. doi:10.1371/journal.pone.0046596
Chen, Y., X. Hao, and J. Cao. 2014. Small auxin upregulated RNA (SAUR) gene family in maize: Identification, evolution, and its phylogenetic comparison with Arabidopsis, rice, and sorghum. J. Integr. Plant Biol. 56:133–150. doi:10.1111/jipb.12127
Cheverud, J.M. 2001. A simple correction for multiple comparisons in interval mapping genome scans. Heredity 87:52–58. doi:10.1046/j.1365-2540.2001.00901.x
Cohen, J. 2003. Human population: The next half century. Science 302:1172–1175. doi:10.1126/science.1088665
Danilevskaya, O.N., X. Meng, D.A. Selinger, S. Deschamps, P. Hermon, G. Vansant, R. Gupta, E.V. Ananiev, and M.G. Muszynski. 2008. Involvement of the MADS-box gene ZMM4 in floral induction and inflorescence development in maize. Plant Physiol. 147:2054–2069. doi:10.1104/pp.107.115261
De Smet, I., and G. Jürgens. 2007. Patterning the axis in plants-auxin in control. Curr. Opin. Genet. Dev. 17:337–343. doi:10.1016/j.gde.2007.04.012
Dickerson, G.E. 1969. Techniques for research in quantitative animal genetics In: Techniques and Procedures in Animal Science Research. Am. Soc. Anim. Sci, New York, NY.
Dong, Z., O. Danilevskaya, T. Abadie, C. Messina, N. Coles, and M. Cooper. 2012. A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. PLoS One 7:e43450. doi:10.1371/journal.pone.0043450
Durand, E., S. Bouchet, P. Bertin, A. Ressayre, P. Jamin, A. Charcosset, C. Dillman, and M.I. Tenaillon. 2012. Flowering time in maize: Linkage and epistasis at a major effect locus. Genetics 190:1547–1562. doi:10.1534/genetics.111.136903
Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6(5):e19379. doi:10.1371/journal.pone.0019379
Forestan, C., S. Farinati, and S. Varotto. 2012. The maize PIN gene family of auxin transporters. Front. Plant Sci. 3:16. doi:10.3389/fpls.2012.00016
Gallavotti, A. 2013. The role of auxin in shaping shoot architecture. J. Exp. Bot. 64:2593–2608. doi:10.1093/jxb/ert141
Gao, X., L.C. Becker, D.M. Becker, J.D. Starmer, and M.A. Province. 2010. Avoiding the high bonferroni penalty in genome-wide association studies. Genet. Epidemiol. 34:100–105.
Gao, X., J. Starmer, and E.R. Martin. 2008. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet. Epidemiol. 32:361–369. doi:10.1002/gepi.20310
Gazzani, S., A.R. Gendall, C. Lister, and C. Dean. 2003. Analysis of the molecular basis of flowering time in Arabidopsis accessions. Plant Physiol. 132:1107–1114. doi:10.1104/pp.103.021212
Halford, N.G., and C.H. Foyer. 2015. Producing a road map that enables plants to cope with future climate change. J. Exp. Bot. 66:3433–3434. doi:10.1093/jxb/erv277
Hallauer, A.R., and M.J. Carena. 2014. Adaptation of tropical maize germplasm to temperate environments. Euphytica 196:1–11. doi:10.1007/s10681-013-1017-9
Hallauer, A.R., M.J. Carena, and J.B. Miranda Filho. 1988. Quantitative Genetics in Maize Breeding, Handbook of Plant Breeding 6, Vol. 2. Springer Science + Business Media, LLC, New York. p. 145.
Harrell Jr, F.E., C. Dupont, and many others. 2016. Hmisc: Harrell Miscellaneous. R package version 3.17-4. http://CRAN.R-project.org/package=Hmisc
Hu, S. 2016. Brassinosteroid and gibberellin control of plant height in maize (*Zea mays. L*) (Doctoral Thesis).
Hung, H.Y., L.M. Shannon, F. Tian, P.J. Bradbury, C. Chen, S.A. Flint-Garcia, M.D. McMullen, D. Ware, E.S. Buckler, J.F. Doebley, and J.B. Holland. 2012. ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc. Natl. Acad. Sci. USA 109:E1913–E1921. doi:10.1073/pnas.1203189109
Izawa, T., Y. Takahashi, and M. Yano. 2003. Comparative biology comes into bloom: Genomic and genetic comparison of flowering pathways in rice and Arabidopsis. Curr. Opin. Plant Biol. 6:113–120. doi:10.1016/S1369-5266(03)00014-1
Jiang, G., Y. Xiang, J. Zhao, D. Yin, X. Zhao, L. Zhu, and W. Zhai. 2014. Regulation of inflorescence branch development in rice through a novel pathway involving pentatricopeptide repeat protein sped1-D. Genetics 197:1395–1407. doi:10.1534/genetics.114.163931
Jung, C., and A.E. Müller. 2009. Flowering time control and applications in plant breeding. Trends Plant Sci. 14:563–573. doi:10.1016/j.tplants.2009.07.005
Khan, S., S.C. Rowe, and F.G. Harmon. 2010. Coordination of the maize transcriptome by a conserved circadian clock. BMC Plant Biol. 10:126. doi:10.1186/1471-2229-10-126
Kong, F., B. Liu, Z. Xia, S. Sato, B.M. Kim, S. Watanabee, T. Yamada, S. Tabata, A. Kanazawa, K. Harada, and J. Abe. 2010. Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. Plant Physiol. 154:1220–1231. doi:10.1104/pp.110.160796
Kriechbaumer, V., W.J. Park, A. Gierl, and E. Glawischnig. 2006. Auxin biosynthesis in maize. Plant Physiol. 8:334–339.
Lee, M., N. Sharopova, W.D. Beavis, D. Grant, M. Katt, D. Blair, and A. Hallauer. 2002. Expanding the genetic map of maize with the B73 x Mo17 (IBM) population. Plant Mol. Biol. 48:453–461. doi:10.1023/A:1014893521186
Leiboff, S., X. Li, H.C. Hu, N. Todt, J. Yang, X. Li, X. Yu, G.J. Muehlbauer, M.C.P. Timmermans, J. Yu, P.S. Schnable, and M.J. Scanlon. 2015. Genetic control of morphometric diversity in the maize shoot apical meristem. Nat. Commun. 6:8974. doi:10.1038/ncomms9974
Li, J., and L. Ji. 2005. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity 95:221–227. doi:10.1038/sj.hdy.6800717
Li, X., D.L. Yang, L. Sun, Q. Li, B. Mao, and Z. He. 2016a. The systemic acquired resistance regulator OsNPR1 attenuates growth by repressing auxin signaling through promoting IAA-amido synthase expression. Plant Physiol. 172:546–558. doi:10.1104/pp.16.00129
Li, Y.X., P.J. Bradbury, X. Liu, F. Lu, C.M. Romay, J.C. Glaubitz, X. Wu, B. Peng, Y. Shi, Y. Song, D. Zhang, E.S. Buckler, Z. Zhang, Y. Li, and T. Wang. 2016b. Identification of genetic variants associated with flowering time using an extremely large multi-genetic background population. Plant J. 86:391–402. doi:10.1111/tpj.13174
Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler, and Z. Zhang. 2012. GAPIT: Genome association and prediction integrated tool. Bioinformatics 28:2397–2399. doi:10.1093/bioinformatics/bts444
Liu, X., M. Haung, B. Fan, E.S. Buckler, and Z. Zhang. 2016. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLoS Genet. 12(2):e1005767. doi:10.1371/journal.pgen.1005767
Liu, Y., H.Y. Jiang, W. Chen, Y. Qian, Q. Ma, B. Cheng, and S. Zhu. 2011. Genome-wide analysis of the auxin response factor (ARF) gene family in maize (Zea mays). Plant Growth Regul. 63:225–234. doi:10.1007/s10725-010-9519-0
Lopes, M.S., I. El-Basyoni, P.S. Baenzinger, S. Singh, C. Royo, K. Ozbek, H. Aktas, E. Ozer, F. Ozdemit, A. Manickavelu, T. Ban, and P. Vikram. 2015. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J. Exp. Bot. 66:3477–3486. doi:10.1093/jxb/erv122
Maraschin, F.S., J. Memelink, and R. Offringa. 2009. Auxin-induced, SCF(TIR1)-mediated poly-ubiquitination marks AUX/IAA proteins for degradation. Plant J. 59:100–109. doi:10.1111/j.1365-313X.2009.03854.x
Méndez-Vigo, B., J.M. Martínez-Zapater, and C. Alonso-Blanco. 2013. The flowering repressor SVP underlies a novel Arabidopsis thaliana QTL interacting with the genetic background. PLoS Genet. 9:e1003289.
de Mendiburu, F. 2015. agricolae: Statistical Procedures for Agricultural Research. R package version 1.2-3. http://CRAN.R-project.org/package=agricolae
Menkir, A., T. Rocheford, B. Maziya-Dixon, and S. Tanumihardjo. 2015. Exploiting natural variation in exotic germplasm for increasing provitamin-A carotenoids in tropical maize. Euphytica 205:203–217. doi:10.1007/s10681-015-1426-z

Moskvina, V., and K.M. Schmidt. 2008. On multiple-testing correction in genome-wide association studies. Genet. Epidemiol. 32:567–573. doi:10.1002/gepi.20331

Mravec, J., P. Skupa, A. Bailly, K. Hoyerova, P. Krecek, A. Bielach, J. Petrasek, J. Zhang, V. Gaykova, Y.D. Stierhof, P.I. Dobrev, K. Schwarzerova, J. Rolcik, D. Seifertova, C. Luschnig, E. Benkova, E. Zazimalova, Geisler, M., and J. Friml. 2009. Subcellular homeostasis of phytohormone auxin is mediated by the ER- localized PIN5 transporter. Nature 459:1136–1140. doi:10.1038/nature08066

Murphy, R.L., R.R. Klein, D.T. Morishige, J.A. Brady, W.L. Rooney, F.R. Miller, D.V. Dugas, P.E. Klein, and J.E. Mullet. 2011. Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiod flowering in sorghum. Proc. Natl. Acad. Sci. USA 108:16469–16474. doi:10.1073/pnas.1106212108

Muszynski, M.G., T. Dam, B. Li, D.M. Shirbroun, Z. Hou, E. Bruggemann, R. Archibald, E.V. Ananiev, and O.N. Danilevskaya. 2006. delayed flowering1 encodes a basic leucine zipper protein that mediates floral inductive signals at the shoot apex in maize. Plant Physiol. 142:1523–1536. doi:10.1104/pp.106.088815

Nemoto, Y., M. Kisaka, T. Fuse, M. Yano, and Y. Ogihara. 2003. Characterization and functional analysis of three wheat genes with homology to the CONSTANS flowering time gene in transgenic rice. Plant J. 36:82-93.

Nyholt, D.R. 2004. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium and each other. Am. J. Hum. Genet. 74:765–769. doi:10.1086/383251

Osnato, M., C. Castillejo, L. Matías-Hernández, and S. Pelaz. 2012. TEMPRANILLO genes link photoperiod and gibberellin pathways to control flowering in Arabidopsis. Nat. Commun. 3:808. doi:10.1038/ncomms1810

Peiffer, J.A., M.C. Romay, M.A. Gore, S.A. Flint-Garcia, Z. Shang, M.J. Millard, C.A.C. Gardner, M.D. McMullen, J.B. Holland, P.J. Bradbury, and E.S. Buckler. 2014. The genetic architecture of maize height. Genetics 196:1337–1356. doi:10.1534/genetics.113.159152

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ranocha, P., O. Dima, R. Nagy, J. Felten, C. Corratgé-Faillie, O. Novák, K. Morreel, B. Lacombe, Y. Martinez, S. Pfrunder, X. Jin, J.P. Renou, J.B. Thibaud, K. Ljung, U. Fischer, E. Martinoia, W. Boerjan, and D. Goffner. 2013. Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homeostasis. Nat. Commun. 4:2625. doi:10.1038/ncomms3625

Ranum, P., J.P. Peña-Rosas, and M.N. Garcis-Casal. 2014. Global maize production, utilization, and consumption. Ann. N. Y. Acad. Sci. 1321:105–112.

Ren, H., and W.M. Gray. 2015. SAUR proteins as effectors of hormonal and environmental signals in plant growth. Mol. Plant 8:1153–1164. doi:10.1016/j.molp.2015.05.003

Röber, F.K., G.A. Gordillo, and H.H. Geiger. 2005. In: vivo haploid induction in maize-Performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50:275–283.

Romay, C.M., M.J. Millard, J.C. Glaubitz, J.A. Peiffer, K.L. Swarts, T.M. Casstevens, R.J. Elshire, C.B. Acharya, S.E. Mitchell, S.A. Flint-Garcia, M.D. McMullen, J.B. Holland, E.S. Buckler, and C.A. Gardner. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. 14:R55. doi:10.1186/gb-2013-14-6-r55

Russell, W.A. 1972. Registration of B70 and B73 parental lines of maize. Crop Sci. 12:721. doi:10.2135/cropsci1972.0011183X001200050085x

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, K.A. Fengler, R. Meeley, E.V. Ananiev, S. Svitashev, E. Bruggemann, B. Li, C.F. Hainey, S. Radovic, G. Zaina, J.-Antoni Rafalski, S.V. Tingey, G.H. Miao, R.L. Phillips, and R. Tuberosa. 2007. Conserved noncoding genomic sequences associated with a flowering- time quantitative trait locus in maize. Proc. Natl. Acad. Sci. USA 104:11376–11381. doi:10.1073/pnas.0704145104

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T.A. Graves, et al. 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112–1115.

Song, Y.H., I. Lee, S.Y. Lee, T. Imaizumi, and J.C. Hong. 2012. CONSTANS and ASYMMETRIC LEAVES 1 complex is involved in the induction of FLOWERING LOCUS T in photoperiodic flowering in Arabidopsis. Plant J. 69:332–342. doi:10.1111/j.1365-313X.2011.04793.x

Teng, F., L. Zhai, R. Liu, W. Bai, L. Wang, D. Huo, Y. Tao, Y. Zheng, and Z. Zhang. 2013. ZmGA3ox2, a candidate gene for a major QTL, qPH3.1, for plant height in maize. Plant J. 73:405–416. doi:10.1111/tpj.12038

Turner, S. 2014. qqman: Q-Q and manhattan plots for GWAS data. R package version 0.1.2. http://CRAN.R-project.org/package=qqman

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

Voorend, W., H. Nelissen, R. Vanholme, A. De Vliegher, F. Van Breusegem, W. Boerjan, I. Roldán-Ruiz, H. Muylle, and D. Inzé. 2016. Overexpression of GA20-OXIDASE1 impacts plant height, biomass allocation and saccharification efficiency in maize. Plant Biotechnol. J. 14:997–1007. doi:10.1111/pbi.12458

Wabnik, W., J. Kleine-Vehn, W. Govaerts, and J. Friml. 2011. Prototype cell-to-cell auxin transport mechanism by intercellular auxin compartmentalization. Trends Plant Sci. 16:468–475. doi:10.1016/j.tplants.2011.05.002

Wallace, J., X. Zhang, Y. Beyene, K. Semagn, M. Olsen, B. Prasanna, and E.S. Buckler. 2016. Genome-wide association for plant height and flowering time across 15 tropical maize populations under managed drought stress and well-watered conditions in sub-Saharan Africa. Crop Sci. 56:2365–2378. doi:10.2135/cropsci2015.10.0632

Wang, Y., D. Deng, H. Ding, X. Xu, R. Zhang, S. Wang, Y. Bian, Z. Yin, and Y. Chen. 2013. Gibberellin biosynthetic deficiency is responsible for maize dominant dwarf11 (D11) mutant phenotype: Physiological and transcriptomic evidence. PLoS One 8:e66466. doi:10.1371/journal.pone.0066466

Wang, Y., D. Deng, Y. Shi, N. Miao, Y. Bian, and Z. Yin. 2012. Diversity, phenology, and evolution of auxin response factor (ARF) family: Insights gain from analyzing maize ARF genes. Mol. Biol. Rep. 39:2401–2415. doi:10.1007/s11033-011-0991-z

Wei, F., E. Coe, W. Nelson, A.K. Bharti, F. Engler, E. Butler, H.R. Kim, J.L. Goicoechea, M. Chen, S. Lee, G. Fuks, H. Sanchez-Villeda, S. Schroeder, Z. Fang, M. McMullen, G. Davis, J.E. Bowers, A.H. Paterson, M. Schaeffer, J. Gardiner, K. Cone, J. Messing, C. Soderlund, and R.A. Wing. 2007. Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genet. 3:e123. doi:10.1371/journal.pgen.0030123

Wolfinger, R.D., W.T. Federer, and O. Codero-Brana. 1997. Recovering Information in Augmented Designs, Using SAS PROC GLM and PROC MIXED. Agron. J. 89:856–859. doi:10.2134/agronj1997.00021962008900060002x

Wright, S.I., I.V. Bi, S.G. Schroeder, M. Yamasaki, J.F. Doebley, M.D. McMullen, and B.S. Gaut. 2005. The effects of artificial selection on the maize genome. Science 308:1310–1314. doi:10.1126/science.1107891

Yang, Q., Z. Li, W. Li, L. Ku, C. Wang, J. Ye, K. Li, N. Yang, Y. Li, T. Zhong, J. Li, Y. Chen, J. Yan, Z. Yang, and M. Xu. 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. doi:10.1073/pnas.1310949110

Yu, J., G. Pressoir, W.H. Briggs, I.V. Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203–208. doi:10.1038/ng1702

Zuber, M.S. 1973. Registration of 20 maize parental lines. Crop Sci. 13:779–780. doi:10.2135/cropsci1973.0011183X001300060085x