THE CROP JOURNAL 8 (2020) 227–242

Available online at www.sciencedirect.com

ScienceDirect

Favorable haplotypes and associated genes for flowering time and photoperiod sensitivity identified by comparative selective signature analysis and GWAS in temperate and tropical maize

Zhiwei Li[a], *Xiaogang Liu*[a], *Xiaojie Xu*[a], *Jiacheng Liu*[a], *Zhiqin Sang*[a,c], *Kanchao Yu*[a,d], *Yuxin Yang*[a], *Wenshuang Dai*[a], *Xin Jin*[a], *Yunbi Xu*[a,b,*]

[a]*Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China*
[b]*International Maize and Wheat Improvement Center (CIMMYT), El Batan 56130, Texcoco, Mexico*
[c]*Xinjiang Academy of Agricultural Reclamation, Shihezi 832000, Xinjiang, China*
[d]*Northeast Agricultural University, Harbin 150030, Heilongjiang, China*

ARTICLE INFO

Article history:
Received 21 March 2019
Received in revised form 21 July 2019
Accepted 10 September 2019
Available online 11 December 2019

ABSTRACT

On the basis of growing environment, maize can largely be classified into temperate and tropical groups, leaving extensive genetic variation and evolutionary signatures in the maize genome. To identify candidate genes governing flowering time and photoperiod sensitivity, selective signature analysis and SNP- and haplotype-based GWAS were performed using 39,350 high-quality SNP markers in temperate and tropical maize groups consisting of 410 inbred lines phenotyped in three representative experiments in different latitudes. Selective signature analysis revealed 106 selective-sweep regions containing 423 candidate genes involved mainly in biological regulation and biosynthesis pathways. Among these genes, 25 overlapped with known genes governing flowering time and photoperiod sensitivity and 37 were also detected by GWAS for days to tassel, anthesis-silk interval, and photoperiod sensitivity measured by days to silking. Only two of the candidate genes governing flowering time overlapped selective signals. Most haplotype alleles within significant haplotype loci showed the same direction of effect on flowering time and photoperiod sensitivity. The inbred lines carrying GATT at HapL499 (haplotype locus 499) on chromosome 1 had relatively short flowering times. Lines carrying CA at HapL4054 on chromosome 10, TA at HapL4055 on chromosome 10, and GTTGT at HapL978 on chromosome 2 were less sensitive to photoperiod than lines carrying other haplotype alleles. Haplotype loci associated with flowering time and photoperiod sensitivity explained respectively 17.5%–18.6% and 11.2%–15.5% of phenotypic variation. Candidate genes and

* *Corresponding author at: Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China and International Maize and Wheat Improvement Center (CIMMYT), El Batan 56130, Texcoco, Mexico.*
E-mail address: y.xu@cgiar.org. (Y. Xu).
Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

https://doi.org/10.1016/j.cj.2019.09.012

favorable haplotypes identified in this study may support the more efficient utilization of maize germplasm groups.

1. Introduction

Archaeological and genetic evidence reveals that maize (*Zea mays* ssp. *mays* L.) was domesticated approximately 10,000 B.P. in the Balsas River basin of southwest Mexico, under high temperature and short day length [1–3]. Following the largely unconscious process of domestication, maize was subjected to targeted improvement, and became adapted to tropical environments mainly through development of landraces, synthetics, and open-pollinated cultivars. More recently, maize has been selected for performance in temperate environments and adapted to modern agro-ecosystems through the development and commercialization of hybrid cultivars. Domestication and post-domestication selection have contributed to large differences between tropical and temperate maize groups, the former showing much richer genetic diversity and more rapid LD (linkage disequilibrium) decay than the latter [4].

Several studies [5–7] have suggested that tropical maize with favorable alleles for resistance to biotic and abiotic stresses should be used to increase genetic diversity for breeding programs. Temperate germplasm with high yield potential can be used as a source of several tolerance-related traits, including high density tolerance and functional stay-green, that can be exploited to improve tropical maize. However, gene introgression between temperate and tropical maize groups has been hindered, mainly by their differences in response to day-length and photoperiod. Flowering time and photoperiod sensitivity are complex, showing a wide range of phenotypic variation influenced by many minor genes, which determine the adaptation of plants to their ecological environments and influence the exchange of germplasm resources across regions with different latitudes [8,9]. Photoperiod sensitivity is usually considered to be a prerequisite for crop survival and reproduction under various environments [10], and is indirectly evaluated by comparing QTL (quantitative trait locus) for flowering time identified under multiple regions in different latitudes [11,12]. Studies of days to heading (DTH) [13] and to flowering (DTF) [14] in response to photoperiod and temperature of rice showed that QTL for DTH and DTF did not coincide with QTL for photo-thermo sensitivity, indicating that photoperiod sensitivity and flowering time were independent. Maize is a wind-pollinated outcrossing species that grows in a range of environments from tropical southwestern Mexico to the Andean highlands and has been widely introduced into tropical and temperate regions [15]. Natural and artificial selection, especially continuous selection over the past century, of key genes affecting flowering time have led to the photoperiod sensitivity of tropical maize and reduced the photoperiod sensitivity of temperate maize [16], such that the time required for maize landraces to mature ranges from 2 to 11 months [17]. Some phenotypic traits of maize have been discovered to be highly associated with photoperiod sensitivity, in particular number of leaves, plant height, and flowering time including days to pollen shed, silking, and anthesis [16,18,19]. These traits can thus be used as efficient indicators of photoperiod sensitivity.

Selective signals and associated genomic regions and candidate genes can be identified and studied using high-density molecular markers such as single nucleotide polymorphisms (SNPs). Although theoretically there are up to four different alleles per SNP locus in a population, SNP arrays can detect only two of them [20,21]. Haplotypes derived from a set of SNPs within a single LD block can be used as markers with more alleles [22]. Haplotype construction compresses multiple SNPs into a haplotype locus and optimizes the design of genomic selection (GS) and genomewide association studies (GWAS) [23]. Haplotype analysis provides two further advantages. The association between a trait and a specific allele depending on *cis* interactions with other loci may not be recognized by SNP-based analysis until the functional haplotypic unit is used in GWAS. Differences in haplotype diversity and frequency across populations may be valuable for identifying variants that are the most likely determinants of phenotypic traits [24,25].

Natural and artificial selection of favorable mutations associated with adaptation leads to reduced polymorphism, increased LD, and increased allele frequency [26–28]. When selection for a beneficial mutation occurs, the genetic variation in the neighboring region will be homogeneous, leaving selective signals in the genome and ultimately shaping the phenotype. With the rapid development of chip technology and high-throughput sequencing technology, selective signals can be identified at the whole-genome level. Selective signatures identification can be divided into three categories according to the source of genomic information and the algorithm employed. (i) Potential selection signatures are identified based on population differentiation, including F_{ST} [29] and *di* [30], by comparison of their allele frequencies among different subgroups. (ii) Potential selection signatures are identified based on LD, including *EHH* (extended haplotype homozygosity) [31], *iHS* (integrated extended haplotype homozygosity score) [32] and *XP-EHH* (cross-population extended haplotype homozygosity) [33], which describes the homozygosity of the extended haplotype shaped by the "hitchhiking effect". (iii) Potential selection signatures are identified based on allele frequency patterns, including Tajima's *D* statistic [34] and the methods related to Tajima's *D*, by distinguishing the difference between θ_π and θ_w that indicates a signature of low-frequency mutation. In addition, emerging approaches commonly employ *XP-CLR* (cross-population composite likelihood ratio), which is used to reveal historical selection by comparing allele frequency spectrum at linked loci [35], and the combination of the θ_π

ratio ($\theta_{\pi,\text{domesticated}}/\theta_{\pi,\text{improved}}$) and F_{ST}, which involves an empirical procedure employing sliding windows with both significantly low or high θ_π ratios and significantly high F_{ST} values to identify potential regions affected by long-term and intensive selection over the whole genome [36–38]. However, small sample sizes weaken the power to detect association of selective signatures with domesticated and improved traits. Recently, large samples have been used to detect selective signals, with respectively 115, 278, and 302 lines included in selective signature analyses of cucumber [39], maize [40], and soybean [41].

Identifying genetic determinants of flowering time and photoperiod sensitivity is usually considered to be a prerequisite for successful exchange of germplasm resources across regions adapted to different latitudes. It is thus desirable to discover selective signatures regulating flowering time and photoperiod sensitivity that distinguish temperate from tropical maizes. In this study, we sought to identify genomewide changes between temperate and tropical maize groups by selective signature analysis and GWAS.

2. Materials and methods

2.1. Plant materials and data generation

The diverse maize germplasm panel used in the study (Table S1), consisted of 410 maize inbred lines. Of these, 238 were temperate inbred lines selected from breeding programs and germplasm collections to represent maize germplasm improved for and adapted to temperate environments and 172 were tropical inbred lines selected from CIMMYT to represent maize germplasm with high genetic diversity but less improved and more adapted to the original environment where maize was domesticated.

The panel had been genotyped with the Maize 55 K SNP Affymetrix Axiom Genotyping Array [42]. From this SNP dataset, 39,350 SNPs were selected as high-quality SNPs with minor-allele frequency (MAF) greater than 0.05 and maximum missing rate less than 20%. They were evenly distributed across the maize genome, with chromosomes carrying 2850 (chromosome 10) to 6050 (chromosome 1) SNPs.

2.2. Field trial and phenotyping

The diverse maize panel was planted in 2014 in three locations, Shunyi, Beijing (40°13′N; summer), Xinxiang, Henan (35°18′N; summer) and Sanya, Hainan (18°09′N; winter), with different environments that affect maize flowering, using a randomized block design with two replications. Each two-row plot was 3 m long with 60 cm between rows and 25 cm between plants. The number of days from planting to the time at which more than 50% of the plants in a plot displayed tasseling, silking, and anthesis was measured as days to tassel (DTT), days to silk (DTS), and days to anthesis (DTA), respectively. The interval between DTA and DTS for the plot was calculated as anthesis-silk interval (ASI). Photoperiod sensitivity was measured as the relative difference index (RD) [43,44], calculated for each flowering trait phenotyped in pairs of environments (Shunyi vs. Sanya; Xinxiang vs. Sanya) as follows:

$$RD_i(\%) = [(L_i - S_i)/S_i] \times 100\%$$

where L_i is the mean performance of a trait under long-day conditions (Shunyi or Xinxiang) and S_i is the mean performance of a trait under short-day conditions (Sanya, Hainan). The RDs for DTT, DTS, and DTA are abbreviated as RD_DTT, RD_DTS, and RD_DTA, respectively.

2.3. Population structure analysis

Using TASSEL software [45], a SNP array of 410 maize accessions was constructed and used to construct a cladogram of SNPs across the whole genome using genetic distances between inbred lines. The online toolkit iTOL (https://itol.embl.de/) [46], was then applied to generate the neighbor-joining tree which displayed the genetic distance gained above. Principal component analysis (PCA) was conducted using TASSEL.

2.4. Linkage disequilibrium analysis

Squared correlation coefficients (r^2) were calculated with PLINK software [47]. The parameters were set to "–ld-window-r2 0 –ld-window 999999 MAF > 0.05". The mean r^2 was computed for pairs of SNPs within 20-kb intervals across the genome.

2.5. Calculation of F_{ST}, θ_π, and Tajima's D statistic

The genetic parameter population-differentiation statistic F_{ST} was used to measure genetic distance or population differentiation using genetic polymorphism data [29]. Nucleotide diversity (θ_π) was employed to measure the degree of diversity within a population [48]. Tajima's D statistic was computed to distinguish random from non-random evolution of DNA sequences [34]. To identify selective signals across the whole genome between temperate and tropical maize groups, F_{ST}, θ_π, and Tajima's D were calculated as indicators of selective signatures by a sliding window method (including 10 SNPs per sliding window in steps of 2 SNPs). F_{ST} was calculated for high quality SNPs as follows [49]:

$$F_{ST} = \frac{\pi_{\text{Between}} - \pi_{\text{Within}}}{\pi_{\text{Between}}}$$

where π_{Between} represents genetic differences across populations and π_{Within} represents genetic differences within populations.

θ_π was calculated according to Nei and Li [48], using

$$\theta_\pi = \sum_{ij} x_i x_j \pi_{ij} = 2 * \sum_{i=2}^{n} \sum_{j=1}^{i-1} x_i x_j \pi_{ij}$$

where x_i and x_j are the frequencies of the ith and jth sequences, respectively, π_{ij} is the number of nucleotide

differences at each site between the *i*th and *j*th sequences, and *n* is the number of sequences over all samples.

Tajima's *D* was calculated as follows [34]:

$$D = \frac{\left(\pi - \frac{S}{a_1}\right)}{\sqrt{(V\left(\pi - \frac{S}{a_1}\right)}}$$

where $a_1 = \sum_{i=1}^{n-1} \frac{1}{i}$, *n* denotes the number of samples, *S* denotes the number of segregating sites, *V* denotes the variance of $\pi - \frac{S}{a_1}$, and π is the sum of the pairwise differences per site between sequences.

The θ_π ratio ($\theta_{\pi,\text{tropical}}/\theta_{\pi,\text{temperate}}$) was calculated and used to evaluate the direction of selection. θ_π, F_{ST}, and Tajima's *D* statistic were calculated using the PopGenome package in R [50].

2.6. Identification of putative selective regions

To identify putative selective signatures associated with the temperate maize group, selective signature analysis using temperate and tropical maize groups was performed to identify selective signatures over the genome. The allele frequencies of SNP loci were used to identify selective sweep regions that were probably shaped by continuous selection. The top 5% values of θ_π ratio were adopted to identify putative selective regions. Likewise, the top 5% values of F_{ST} were adopted to confirm highly differentiated regions. The intersections of windows based on F_{ST} and θ_π ratio were assigned as potential selective regions.

2.7. Functional enrichment analysis of candidate genes

To identify biological functions of candidate genes associated with selection, annotation and enrichment analysis were performed by submitting gene information to Gene Ontology (GO) using the agriGO online platform (http://bioinfo.cau.edu.cn/agriGO/) [51]. A hypergeometric distribution adjusted by false discovery rate (FDR) was employed to check the relationship of different genes in identical GO terms and return a *P*-value for each term. GO terms with *P*-value < 0.05 were taken as those in which candidate genes were significantly enriched.

2.8. SNP-based GWAS

GWAS was performed with TASSEL for seven traits associated with flowering time and photoperiod sensitivity, including DTT, DTS, DTA, ASI, RD_DTT, RD_DTS, and RD_DTA. SNP loci were selected for identification of candidate genes based on a threshold of $P \leq 0.0001$. To correct false positives, a mixed linear model (MLM, PCA + K) was applied in GWAS for both temperate and tropical maize groups together. The population structure was determined by PCA. Thus, kinship matrix (K) was treated as a random effect in MLM.

2.9. Haplotype construction and haplotype-based GWAS

Haplotype loci were identified from the high quality SNP set with 39,350 SNPs using Haploview's interpretation to define haplotype loci as recommended by Gabriel S [52], using PLINK. The parameters were set to "–allow-extra-chr –blocks no-pheno-req –blocks-max-kb 1000 –geno 0.2 –blocks-strong-lowci 0.8 –out –vcf ", which means that two sites are considered to be in a block of LD if the bottom of the confidence interval of r^2 is greater than 0.8. The R package *GHap* was employed to call haplotypes and identify haplotype alleles across all inbred lines. A haplotype genotype matrix, generated by the *ghap.haplotyping* program, was transformed to PLINK format with the *ghap.hap2tped* program, and converted to VCF format for haplotype-based GWAS. Phenotypic variation explained was estimated for both SNPs and haplotypes using marker R^2, which takes genetic background and population structure into account [53,54]. Taking traits related to flowering time and photoperiod as dependent variable, multiple regression analysis was carried out with significantly associated haplotype loci, respectively.

3. Results

3.1. Genetic variation in flowering time and photoperiod sensitivity

Table 1 shows phenotypic means for maize flowering time (DTT, DTS, DTA, and ASI) and photoperiod sensitivity (relative difference index traits, RD_DTT, RD_DTS, and RD_DTA). Tropical inbreds had relatively long flowering times. With decreased latitude, the number of days to flowering was shortened for both temperate and tropical inbred lines. Compared with temperate maize, tropical maize was relatively sensitive to photoperiod as revealed by both comparisons, Shunyi vs. Sanya and Xinxiang vs. Sanya, and temperate maize showed a greater difference in longer days and lower temperature during the crop season (Fig. S1, Fig. S2). The broad-sense heritability estimates of flowering time and photoperiod sensitivity were moderate to high except for ASI (Table 1). Both flowering time and photoperiod sensitivity showed wide genetic variation for GWAS and associated analyses.

3.2. Phylogenetic relationship and LD revealed by SNPs

The inbred lines could be clearly divided into two major groups, temperate and tropical, although a few lines were mismatched in the neighbor-joining tree (Fig. S3). PCA visualization using the first three eigenvectors shows a similar classification result (Fig. S4). The levels of LD were attenuated to half of their maximum values at 40 kb for the tropical and 120 kb for the temperate group (Fig. S5). The mean distance of LD decay in the temperate group was thus three times that in the tropical group, supporting that tropical maize underwent intensive recombination with more genetic diversity.

3.3. Selective signatures and associated candidate genes revealed by GO analysis

F_{ST}, θ_π ratio, and Tajima's *D* across the whole maize genome are shown in a circus plot (Fig. S6). A total of 106 putative selective-sweep regions were identified (Fig. 1), each

Table 1 – Phenotypic means and heritability for flowering time and photoperiod sensitivity.

	Flowering time per se (day)								Photoperiod sensitivity-RD (%)					
	DTT		DTS		DTA		ASI		RD_DTT		RD_DTS		RD_DTA	
	TEM	TRO	TEM	TRO	TEM	TRO	TEM	TRO	TEM	TRO	TEM	TRO	TEM	TRO
Shunyi	65.8^{B}_{a}	79.0^{A}_{a}	70.1^{B}_{a}	84.0^{A}_{a}	68.0^{B}_{a}	82.0^{A}_{a}	2.4^{B}_{a}	4.3^{A}_{a}						
Xinxiang	57.0^{B}_{b}	64.5^{A}_{b}	62.2^{B}_{b}	69.3^{A}_{b}	62.2^{B}_{b}	69.3^{A}_{b}	0.6^{A}_{c}	0.5^{A}_{c}						
Sanya	55.4^{B}_{b}	62.8^{A}_{c}	59.8^{B}_{c}	67.9^{A}_{c}	58.0^{B}_{c}	66.2^{A}_{c}	2.0^{A}_{b}	2.0^{A}_{b}						
Shunyi vs. Sanya									19.3^{B}_{a}	26.3^{A}_{a}	17.7^{B}_{a}	26.5^{A}_{a}	17.7^{B}_{a}	24.2^{A}_{a}
Xinxiang vs. Sanya									3.25^{A}_{b}	3.44^{A}_{b}	2.55^{B}_{b}	4.41^{A}_{b}	5.30^{B}_{b}	7.61^{A}_{b}
Heritability	0.86		0.81		0.85		0.26		0.84		0.75		0.80	

TEM, temperate maize group; TRO, tropical maize group. Uppercase letters represent significance between temperate and tropical maize groups at $P < 0.05$ for a given trait in a given environment and lowercase letters represent significance at $P < 0.05$ for a given trait across environments using the least significant difference (LSD) test.

consisting of four genes in a region of mean size 208.7 kb, covering in total 1.07% of the maize genome. These regions were validated by a significantly lower level of Tajima's *D* (*P*-values 2.2×10^{-16}) in the temperate maize group.

The identified selective-sweep regions contained 423 protein-coding genes, linked with only 53 gene models. Taking a maize genome annotated with GO as background reference, 37 significant GO terms were detected in the candidate genes of selective-sweep regions, of which 30 were categorized as biological process, 6 as molecular function, and 1 as cellular component (Fig. 2). GO enrichment analysis indicated that candidate genes from selective-sweep regions in temperate maize were significantly enriched in biosynthesis and regulation of biological process pathways. Specifically, three genes identified for flowering in previous studies [55,56] were overrepresented in multiple GO terms. Both GRMZM2G124421 and GRMZM2G129034 displayed overrepresentation in GO:0003700 (transcription factor activity), GO:0005634 (nucleus), GO:0006355 (regulation of transcription), while GRMZM2G180555 showed a strong

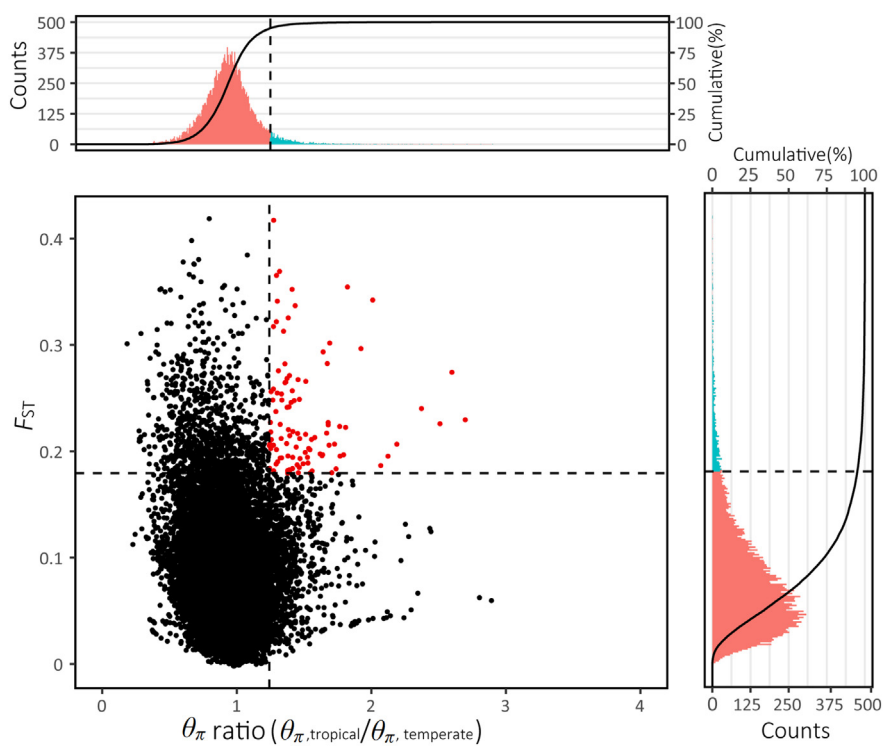

Fig. 1 – Genome-wide selective sweep analysis of the temperate maize group. Distribution of the θ_π ratio ($\theta_{\pi,\text{tropical}}/\theta_{\pi,\text{temperate}}$) and F_{ST} values were calculated in 10-SNP sliding windows with 2-SNP steps. The horizontal and vertical lines represent threshold lines of the top 5% of the F_{ST} and θ_π ratio values, respectively. Points (red) located in the top right sector represent selective signatures of the temperate maize group. Blue and red bins in the histograms of F_{ST} (right) and θ_π ratio (top) represent levels respectively higher and lower than the threshold line.

Fig. 2 – GO enrichment analysis of candidate genes of selected regions in temperate group. Blue, green, and red bars indicate biological process, molecular function, and cellular component, respectively.

overrepresentation in GO:0005634 (nucleus) and GO:0003677 (DNA binding). The identification of candidate genes within selective signatures and the enrichment analysis of the candidate genes show that candidate genes associated with biological regulation and biosynthesis pathways, especially for regulation of flowering time, experienced more selection in the process of differentiation between temperate and tropical maize groups.

3.4. *Selective-sweep regions associated with flowering time and photoperiod sensitivity*

Previous studies have shown that the process of maize improvement and adaptation is not strictly sequential, but overlapped [57]. Several genes associated with flowering time have been reported in recent studies [55,56]. Comparison with the physical loci of previously identified flowering-time genes revealed that 25 known flowering genes overlapped the selective sweep regions identified in this study (Fig. 3, Table S2).

A 160-kb genomic region around each significant SNP from the 63 moderate GWAS signals was analyzed for DTT, DTS, DTA, and ASI (Fig. S7). A total of 16 genes involved in flowering time were found to overlap with previously reported genes identified in GWAS (Table S3). Comparison of each GWAS-identified locus with selective-sweep regions as described for soybean [41], cotton [58], and sesame [59] revealed 35 candidate genes that may regulate flowering time, of which 27 were associated with ASI and eight with DTT, and two (GRMZM2G015384 and GRMZM2G031447) overlapped known flowering-time genes (Fig. 3a, Table 2). The allele frequency distribution and nucleotide diversity at the significant GWAS-identified loci showed similar results that these loci identified by selective signatures and GWAS signals have undergone selection.

A previous study [60] suggested that several traits (DTS, DTT, DTA, ASI, plant height and ear height) reliably reflects the photoperiod sensitive characteristics of different maize groups, so that these traits can be employed as a reliable indicator for photoperiod sensitivity measured by their relative difference under different day lengths. Diverse ecological types and various traits in maize germplasm have different levels of photoperiod sensitivity, indicating its complexity. In the present study, RD_DTT, RD_DTS, and RD_DTA measured for temperate maize under two contrasting day-length conditions, in Shunyi and Sanya, were significantly lower than those for tropical maize (Welch's *t*-test with *P*-values of 7.051×10^{-13}, 2.828×10^{-13}, and 1.296×10^{-11}, respectively). However, no significant difference was found between temperate and tropical maize groups under the other pair of contrasting conditions, Xinxiang, Henan and Sanya. Under the Shunyi vs. Sanya contrast, 6, 5, and 7 candidate loci were identified by GWAS to be associated with RD_DTT, RD_DTS, and RD_DTA, respectively (Fig. S8a). In the Xinxiang vs. Sanya contrast, however, only 1, 2, and 3 SNPs were identified for RD_DTT, RD_DTS, and RD_DTA, respectively (Fig. S8b). Yet no significant SNPs were shared between the two sets of GWAS. When a 160 kb window around each significant SNP was used to identify candidate genes, two candidate genes were identified by both GWAS and selective sweep regions (Table 2). These results can be used to functionally characterize candidate genes in association with flowering time and photoperiod sensitivity for improvement of tropical maize.

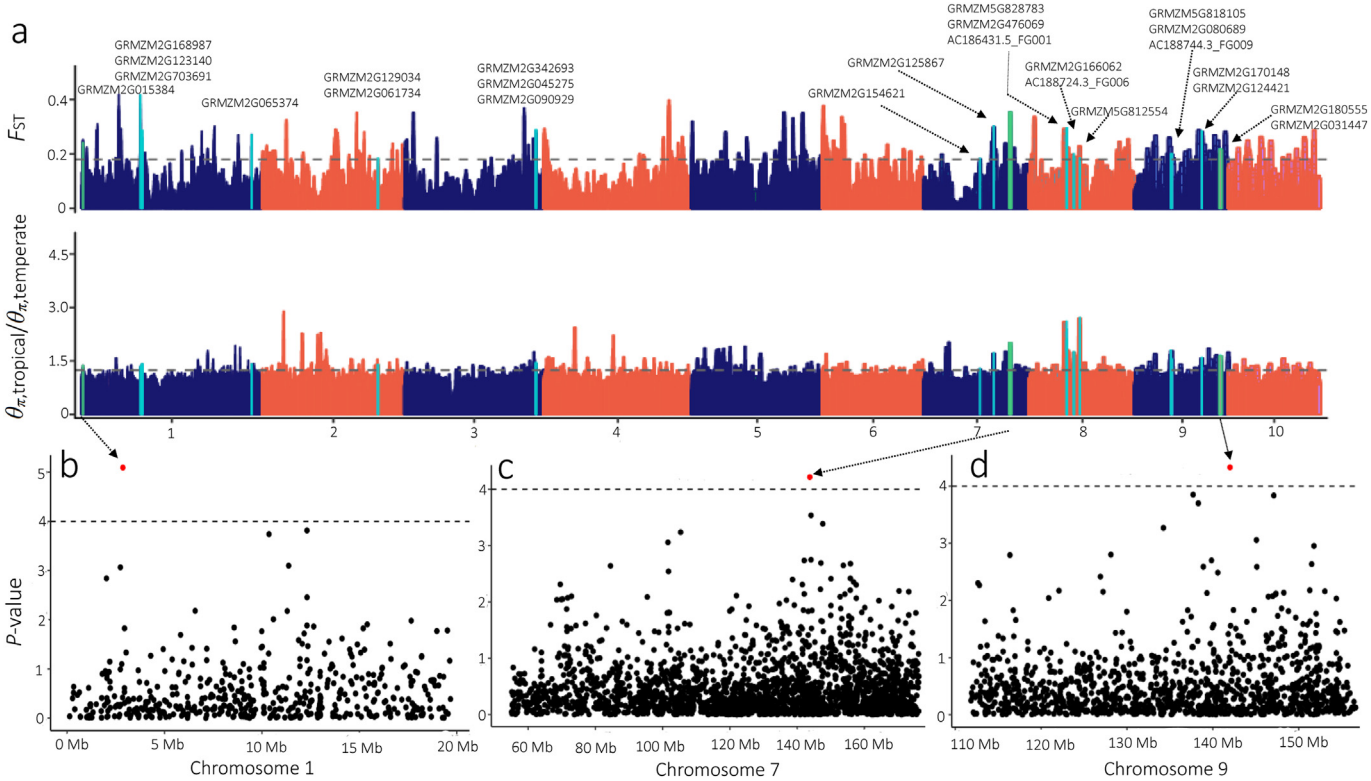

Fig. 3 – Significant GWAS signals that overlapped with strong selective signatures revealed by the distributions of F_{ST} and θ_{π} ratio ($\theta_{\pi,tropical}/\theta_{\pi,temperate}$). The distributions of F_{ST} (top) and θ_{π} ratio ($\theta_{\pi,tropical}/\theta_{\pi,temperate}$) (middle) were used to identify selective signals. The horizontal broken lines in the F_{ST} plot and θ_{π} ratio ($\theta_{\pi,tropical}/\theta_{\pi,temperate}$) plot represent the threshold of selective sweep ($F_{ST} \geq 0.18$; θ_{π} ratio ≥ 1.24). Selective signatures that overlapped maize flowering-time genes identified in previous studies are labeled with turquoise bars and the gene names assigned by MaizeGDB (https://www.maizegdb.org/) are shown above. Selective signatures overlapping significant signals of GWAS are shown in green. Three significant GWAS signals that overlapped strong selective signatures are shown in the bottom plots.

Table 2 – Selected key genes shared by GWAS signals and selective-sweep regions identified in this study.

Gene ID	Annotation	Position	Identified by[b]
GRMZM2G015384[a]	arp4 (actin related protein like4)	Chr1:2874117–2894252	SK and SG_ASI
GRMZM2G031447[a]	Carboxyesterase 17	Chr9:141795429–141797022	SK and SG_ASI
GRMZM2G085691	Transducin/WD40 repeat-like superfamily protein	Chr1:2683790–2690223	SG_ASI
GRMZM2G388684	(XTH32) xyloglucan endotransglucosylase/hydrolase 32	Chr1:2690772–2692141	SG_ASI
GRMZM2G085827	4′-Phosphopantetheinyl transferase superfamily	Chr1:2725603–2734211	SG_ASI
GRMZM2G085872	CemA-like proton extrusion protein-related	Chr1:2734521–2739053	SG_ASI
GRMZM2G085885	Cyclophilin-like peptidyl-prolyl *cis*-trans isomerase family protein	Chr1:2739479–2748501	SG_ASI
GRMZM2G086242	CemA-like proton extrusion protein-related	Chr1:2740244–2744906	SG_ASI
GRMZM2G086269	Plant protein of unknown function (DUF827)	Chr1:2751021–2753499	SG_ASI
GRMZM2G042552	SNF7 family protein	Chr1:2781765–2783974	SG_ASI
GRMZM2G042539	Expressed protein	Chr1:2784642–2785721	SG_ASI
GRMZM2G042250	Homeobox-leucine zipper family protein	Chr1:2791599–2798512	SG_ASI
GRMZM2G041881	btf3 (BTF3 homolog)	Chr1:2798951–2801848	SG_ASI
GRMZM5G802392	Expressed protein	Chr1:2839338–2840257	SG_ASI
GRMZM2G006673	Nuclear matrix protein-related	Chr1:2856655–2869296	SG_ASI
GRMZM2G007025	Peptidoglycan-binding LysM domain-containing protein	Chr1:2871385–2874062	SG_ASI
GRMZM2G436286	Phosphoprotein phosphatase inhibitors	Chr1:2919564–2920987	SG_ASI
GRMZM2G137366	met7 (DNA methyl transferase7)	Chr1:2930055–2935014	SG_ASI
GRMZM2G137352	Remorin family protein	Chr1:2949522–2951161	SG_ASI
GRMZM2G137329	Bifunctional inhibitor/lipid-transfer protein	Chr1:2993011–2993734	SG_ASI
GRMZM2G137236	Adaptor protein complex AP-2 alpha subunit	Chr1:3002518–3020918	SG_ASI
GRMZM2G436226	Nucleic acid binding; zinc ion binding; DNA binding	Chr1:3022700–3026815	SG_ASI
GRMZM2G091586	Cyclin-like family protein	Chr7:143551162–143557193	SG_DTT
GRMZM5G852968	No specific description	Chr7:143560266–143564142	SG_DTT
GRMZM2G390109	Clathrin adaptor complexes medium subunit family protein	Chr7:143565828–143566639	SG_DTT
GRMZM2G135782	PIN domain-like family protein	Chr7:143666299–143668769	SG_DTT
GRMZM2G085948	RING/U-box superfamily protein	Chr7:143729195–143733595	SG_DTT
GRMZM2G449901	No specific description	Chr7:143781954–143783080	SG_DTT
GRMZM2G150674	Expressed protein	Chr7:143783551–143785398	SG_DTT
GRMZM2G449875	(RHA3A) RING-H2 finger A3A	Chr7:143786008–143786918	SG_DTT
GRMZM2G031394	Regulator of Vps4 activity in the MVB pathway protein	Chr9:141804747–141806396	SG_ASI
GRMZM2G068613	Regulator of Vps4 activity in the MVB pathway protein	Chr9:141837577–141839240	SG_ASI
GRMZM2G068647	Regulator of Vps4 activity in the MVB pathway protein	Chr9:141853092–141854752	SG_ASI
GRMZM2G177610	Regulator of Vps4 activity in the MVB pathway protein	Chr9:141900811–141902435	SG_ASI
GRMZM2G177599	LYR family of Fe/S cluster biogenesis protein	Chr9:141902891–141903466	SG_ASI
GRMZM2G130379	Rubredoxin-like superfamily protein	Chr2:224257152–224258761	SG_RD_DTS
GRMZM2G410567	(GH3.17) Auxin-responsive GH3 family protein	Chr2:224301336–224307757	SG_RD_DTS

[a] Li et al. [55].
[b] SK, known genes located in selective-sweep regions; SG denotes selective signature analysis and genome-wide association study and RD denotes relative difference. ASI, anthesis–silk interval; DTT, days to tassel; DTS, days to silk.

3.5. *Haplotype-based GWAS and favorable haplotypes identified*

From 39,350 SNPs, 4166 haplotype loci were constructed, each containing 2 to 75 SNPs with a means of 3.5 SNPs and 2.8 alleles per haplotype. Haplotype-based GWAS was performed for four flowering traits and three photoperiod sensitivity traits using the data collected in three locations. Using a Q+K model, we consistently identified 16 different haplotype loci surpassing the significance threshold (–lg $P \geq 4$). Of 159 candidate genes within the significant haplotype, only eight overlapped known genes associated with flowering time (Table S3).

Multiple regression revealed that significant haplotype loci explained 18.6%, 17.5%, and 17.7% of phenotypic variation at

Table 3 – Multiple regression models for DTT, DTS, and DTA determined by haplotype loci significantly associated with flowering time.

Trait	HapL499	HapL797	HapL1153	HapL3588	HapL3743	HapL3824	Adjusted R^2 (%)	*P*–value
Chr.	1	2	3	9	9	9		
DTT	–0.05	0.05	0.15	–0.14	0.10	0.09	18.6	2.5×10^{-10}
DTS	–0.05	0.05	0.06	–0.12	0.07	0.09	17.5	1.2×10^{-9}
DTA	–0.05	0.05	0.07	–0.13	0.07	0.10	17.7	8.9×10^{-10}

DTT, days to tassel; DTS, days to silk; DTA, days to anthesis; HapL, haplotype locus.

Table 4 – Multiple regression models for RD_DTT, RD_DTS, and RD_DTA determined by haplotype loci associated with photoperiod sensitivity.

Trait	HapL524	HapL978	HapL1994	HapL2939	HapL4054	HapL4055	Adjusted R^2 (%)	*P*–value
Chr.	1	2	5	7	10	10		
RD_DTT	−0.10	0.08	−0.10	−0.09	0.45	−0.29	14.8	1.2×10^{-7}
RD_DTS	−0.09	0.20	−0.13	−0.03	0.38	−0.40	11.2	1.3×10^{-5}
RD_DTA	−0.15	0.14	−0.10	−0.05	0.47	−0.26	15.5	4.5×10^{-8}

RD_DTT, relative difference for days to tassel; RD_DTS, relative difference for days to silk; RD_DTA, relative difference for days to anthesis; HapL, haplotype locus.

the 0.01 significance level for DTT, DTS, and DTA, respectively (Table 3). For photoperiod sensitivity, however, significant haplotype loci explained 14.8%, 11.2%, and 15.5% of the phenotypic variation for RD_DTT, RD_DTS, and RD_DTA, respectively (Table 4).

No significant haplotype loci were shared between flowering time and photoperiod sensitivity, indicating that haplotype loci associated with the two trait types were independent of one another. For seven haplotype loci significantly associated with at least two traits of flowering time and photoperiod sensitivity, the great majority of haplotype alleles showed the same effect direction (Table 5).

For the four haplotype loci that were identified using multi-location data for multiple traits, HapL499 on chromosome 1, significantly associated with flowering time, contained one gene (Fig. 4a) while HapL978 on chromosome 2, significantly associated with photoperiod sensitivity, contained five genes (Fig. 4b), which have not been reported previously except for GRMZM2G166337 annotated as a transcription factor. No candidate genes were found within HapL4054 and HapL4055 on chromosome 1 to be significantly associated with photoperiod sensitivity. The inbred lines carrying GATT at HapL499 had relatively short DTT, DTS, and DTA compared to other three haplotype alleles (Table 5). HapL4054 and HapL4055 were significantly associated with RD_DTT, RD_DTS, and RD_DTA. CA and CG at HapL4054 and TA and TG at HapL4055 showed opposite effects on photoperiod sensitivity. Inbred lines carrying CA and TA were less sensitive to photoperiod than other haploytype alleles in same haplotype locus (Fig. 5). Thus, haplotype alleles GATT, GTTGT, CA, and TA, which had negative effects on flowering time and photoperiod

Table 5 – Significant haplotype loci associated with flowering time and photoperiod sensitivity.

HapL #	Position	Haplotype	Haplotype frequency		Haplotype effect		
			TEM	TRO	DTT/RD_DTT	DTS/RD_DTS	DTA/RD_DTA
499	Chr1:248429068–248500131	GAGC	71.0	54.7	1.1^{a}	1.0^{a}	1.1^{a}
		TGGC	10.5	19.2	0.6^{a}	1.1^{a}	0.9^{a}
		TAGC	7.6	17.4	0.03^{a}	-0.2^{a}	-0.1^{a}
		GATT	10.5	8.7	-3.3^{b}	-3.5^{b}	-3.5^{b}
797	Chr2:78709142_79647071	CTTTAGCGGACC	3.4	13.3	4.0^{a}	1.6^{a}	4.2^{a}
		TTTTAGCGGACC	13.5	1.7	0.2^{b}	1.4^{a}	0.5^{b}
		CTCCAGCGGGAT	23.9	23.3	0.3^{b}	0.5^{a}	0.2^{b}
		CCTCAGGAAGCC	8.4	2.0	-0.1^{b}	-0.7^{a}	-1.7^{b}
		CTCCAACGGGAT	36.1	45.9	-1.6^{b}	-0.9^{a}	-1.6^{b}
3588	Chr9:33138398_33151443	CGGAATGT	28.6	11.1	2.5^{a}	0.7^{a}	2.6^{a}
		CCAAATGT	2.1	11.6	0.1^{ab}	0.4^{a}	-0.2^{ab}
		TGGAGTCT	2.1	20.9	-0.4^{ab}	0.9^{a}	-0.2^{ab}
		CCACACCC	65.5	38.4	-1.6^{b}	-0.7^{a}	-1.8^{b}
3743	Chr9:118450394_118455442	TCA	10.9	8.7	3.4^{a}	1.7^{a}	2.6^{a}
		GAG	71.8	70.3	-0.8^{b}	-0.8^{a}	-0.5^{b}
		GCA	15.1	18.6	-1.5^{b}	-0.1^{a}	-1.2^{b}
978	Chr2:205953303_205953528	GTTGC	**40.0**	**64.5**	**1.6^{a}**	**1.0^{a}**	**1.3^{a}**
		ACCAC	**10.9**	**2.3**	**0.6^{a}**	**-1.2^{a}**	**3.8^{a}**
		GTTGT	**37.8**	**26.1**	**-5.3^{b}**	**-1.5^{a}**	**-5.0^{b}**
4054	Chr10:92518663_92535315	CG	**55.1**	**71.3**	**5.1^{a}**	**4.8^{a}**	**5.3^{a}**
		CA	**31.9**	**16.8**	**-4.4^{b}**	**-4.3^{b}**	**-4.8^{b}**
4055	Chr10:92731794_92736229	TG	**26.5**	**46.5**	**3.4^{a}**	**5.0^{a}**	**3.8^{a}**
		TA	**45.4**	**31.3**	**-4.2^{b}**	**-4.8^{b}**	**-4.3^{b}**

HapL, haplotype locus; DTT, days to tassel; DTS, days to silk; DTA, days to anthesis; Haplotype loci significantly associated with flowering time and photoperiod sensitivity are shown in plain and bold fonts, respectively. Different lowercase letters represent significant differences at $P < 0.05$ by LSD test.

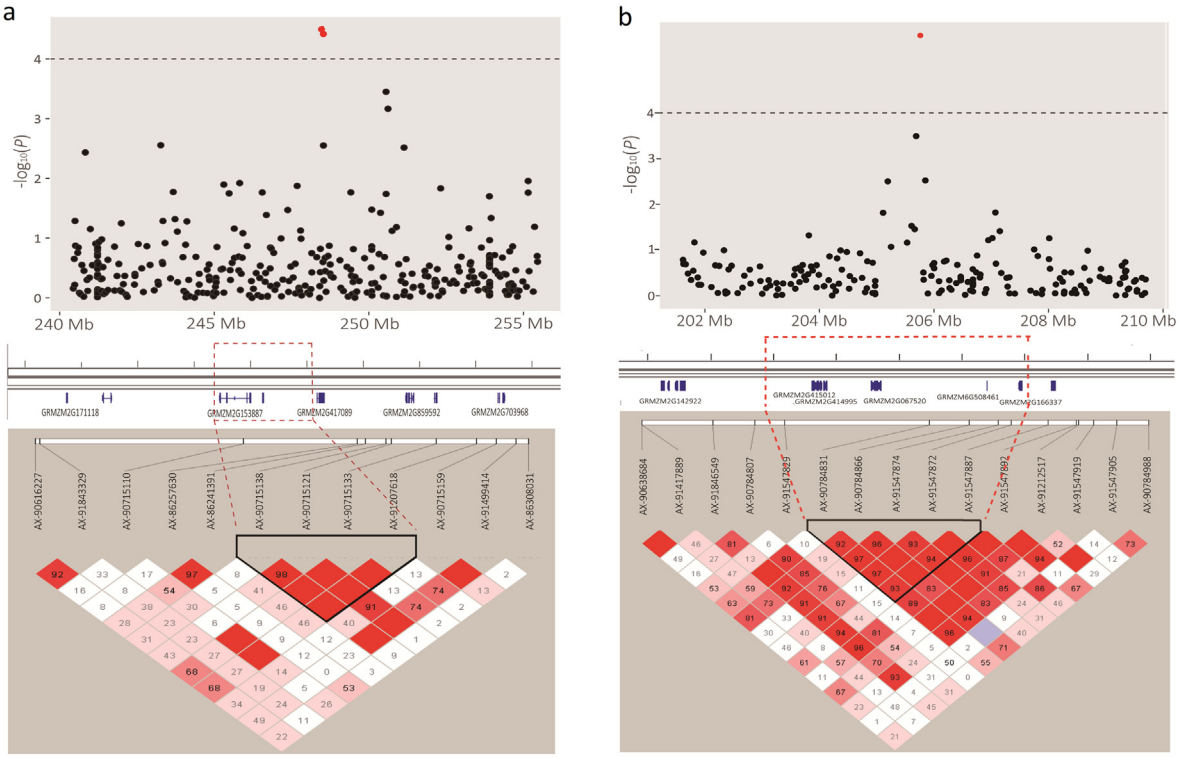

Fig. 4 – Identification of candidate genes as shown by the peaks on chromosomes in GWAS for DTA in HapL499 on chromosome 1 (a) and RD_DTT in HapL978 on chromosome 2 (b). Local Manhattan (top), candidate gene structures and names located in the haplotype locus (middle) and LD heatmap (bottom).

sensitivity, can be considered as favorable haplotypes for shortening flowering time and reducing photoperiod sensitivity. The temperate group showed a higher frequency of the haplotype alleles GATT, GTTGT, CA, and TA than the tropical group.

Also, as described by the result of GWAS for DTT, haplotype-based GWAS increased the proportion of phenotypic variation explained in comparison with SNP-based GWAS (Fig. 6). Similar results were found for the other traits (Fig. S9). Haplotype-based GWAS provides some favorable

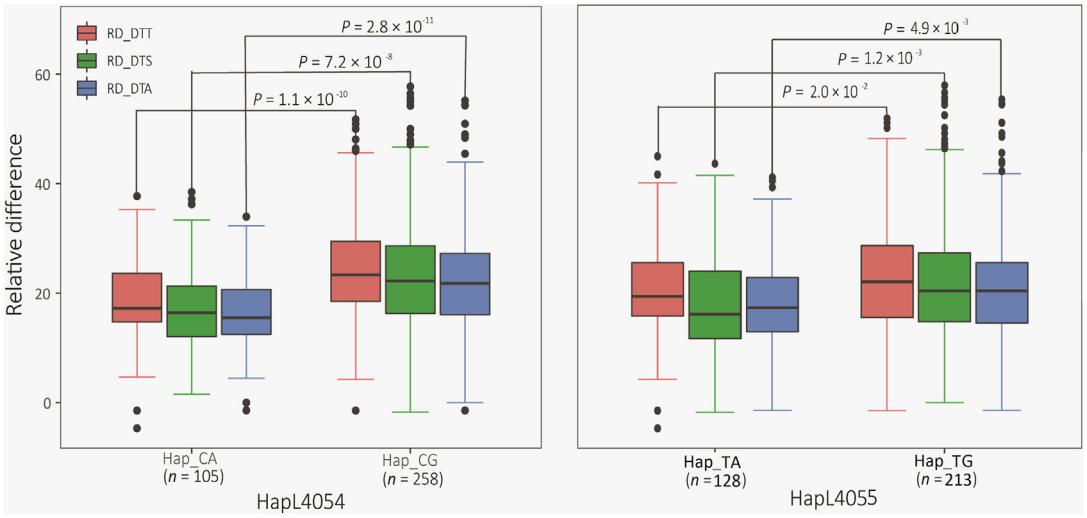

Fig. 5 – Boxplots for RD_DTT, RD_DTS, and RD_DTA based on haplotypes for HapL4054 (left) and HapL4055 (right). Box edges represent 25% quantile (top) and 75% quantile (bottom) while the black bold line represents median value. Statistical significance was determined by Welch's *t*-test.

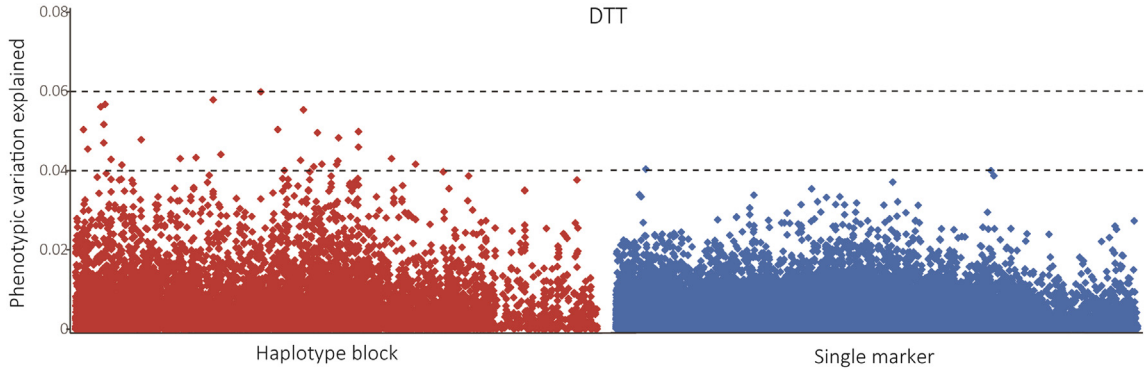

Fig. 6 – Comparison of powers (phenotypic variation explained) in GWAS based on single SNP markers (right) and haplotype loci (left).

haplotype alleles for introducing into tropical maize and might identify specific haplotypes associated with phenotypic traits by detecting associations between traits and alleles that depend on *cis* interactions with other loci.

4. Discussion

4.1. *Importance of tropical germplasm resources*

Maize is one of the most diverse plants at both genetic and morphological levels [61–63]. Previous studies have revealed that the genetic diversity of the breeding pool of elite temperate maize germplasm has declined over the past century [64]. However, the decline of genetic diversity can be mitigated by use of untapped sources, including landraces and wild relatives [65]. Tropical and subtropical maize germplasm resources with a wide range of genetic variation in biotic and abiotic stress tolerance are valuable for maize improvement, particularly for temperate maize. However, the introduction and cultivation of tropical maize has been hindered by its lack of adaptability to temperature and photoperiod [5]. In this study, we identified candidate genes and favorable haplotypes using selective signature analysis and GWAS, which can be applied to molecular breeding for using germplasm resources in different regions.

4.2. *Photoperiod sensitivity as a barrier for the use of tropical germplasm*

Teosinte, a presumed ancestor of maize, probably evolved ecotypes with photoperiod sensitivity to coordinate its reproductive stages to water, short-day environmental conditions [66,67]. Maize reduced its photoperiod sensitivity to adapt to long-day conditions when dispersing from its original area to temperate latitudes [68]. Thus, photoperiod sensitivity is an important factor for widening the planting area of diverse maize varieties. A study [13] of rice response to temperate and photoperiod showed that four genome regions associated significantly with DTH had much lower likelihood of odds (LOD) scores for photo-thermo sensitivity (PTS) than a threshold line, while a PTS QTL region similarly showed a much lower LOD score for DTH than the threshold line. This result indicates that the PTS QTL is independent of DTH QTL. Mining candidate genes associated with photoperiod sensitivity will help to understand the genetic changes during domestication and improvement and contribute to reducing the barriers to use of tropical germplasm [5,69,70]. In the present study, SNP- or haplotype-based GWAS identified 11, 9, and 15 significant loci for RD_DTT, RD_DTS, and RD_DTA, respectively. Candidate genes within significant signal regions could be deployed to overcome the barrier of photoperiod sensitivity and improve temperate maize using tropical germplasm resources.

4.3. *Combined use of F_{ST} and θ_π ratio to detect selective signatures*

XP-EHH, an algorithm based on haplotype, was developed [33] to identify recently fixed or high allele frequency in selective sweep regions via across-population comparison. The *XP-EHH* procedure is insensitive to background selection and thus constitutes a less confusing method for systematic detection of positive selection in a genome [71]. In comparison with F_{ST}, however, the *XP-EHH* algorithm requires a denser set of SNP markers. Both *XP-CLR* and F_{ST} identify signatures of domestication and improvement by comparing allele frequency between two or more populations. However, the *XP-CLR* algorithm was developed [72] to reveal "ancient" selective signatures from several thousand years ago. The combined use of F_{ST} and θ_π ratio has been shown [36,73] to be an effective method for identifying selective signals, especially when functional regions responsive to specific environmental conditions are being sought. In the present study, F_{ST}, and θ_π ratio were employed to investigate genome variation between temperate and tropical maize groups. By comparing the 106 selective sweep regions identified for temperate maize in this study with previously reported features (466 involved in maize domestication and 573 involved in maize improvement) [74], we found that 15 selective sweep features overlapped domestication features and eight overlapped improvement features (Table S4), supporting the reliability of our study.

4.4. Advantages of haplotype-based GWAS

The power of association analysis for detecting positive associations between SNPs and traits can be shown by explained phenotypic variation within a population [16]. Phenotypic variance of a specific trait is determined by phenotypic effect and frequency differences among different allelic variants in the test population. SNPs afford a limited number of alleles per locus, in turn limiting the polymorphism information content at each locus [75–77] whereas haplotype data contains multiple haplotypes (2–9) at each haplotype locus and a higher level of allele diversity. Accordingly, in this study the use of haplotypes increased the amount of phenotypic variation explained, as indicated by GWAS for DTT, supporting the advantages of haplotype-based association analysis. SNP-based GWAS identified 87 significant SNPs above threshold for seven traits, many more than the 16 significant haplotype loci identified by haplotype-based GWAS at the same threshold. This result is due largely to the fact that haplotype data have fewer loci, requiring a lower threshold value. In addition, haplotype length was probably underestimated for the temperate maize group but overestimated for the tropical maize group when haplotype loci were identified using the entire sample. In selecting an improved threshold of linkage disequilibrium ($r^2 \geq 0.8$), we expected that bias, if any, would be greatly reduced. A previous report [4] recommended that the combination of SNP- and haplotype-based GWAS is a better option to identify associated signals than either alone.

4.5. Combined use of selective signature analysis and GWAS in genetic studies

In the process of domestication and improvement, long-term natural and artificial selection in maize has resulted in high genetic variation associated with yield, plant type, flowering time, and photoperiod sensitivity, and also left differential footprints on the genome across germplasm groups. Selective signature analysis focuses on detecting signatures of positive natural and artificial selection and mining candidate genes [78]. Although several studies [37,72,79] have searched for selective signatures using diverse algorithms, few candidate genes associated with specific traits have been identified accurately. Given that the selective signatures identified have spanned several or even dozens of kilobases, it may not be possible to identify all candidate genes within selective signatures. GWAS provides a powerful tool to reconnect phenotypic traits back to their underlying genetic factors [75]. Although GWAS is a sensitive means of identifying candidate genes that have experienced selection, selective signature analysis is preferred and can be used to overcome the limitation. In the present study, both selective signature analysis and GWAS were used, with candidate genes identified, of which 35 genes were associated with flowering time and two with photoperiod sensitivity.

4.6. Rapid LD decay in tropical maize

Generally, the process of domestication and improvement reduces effective population size and levels of genetic diversity [80,81]. However, we found that the genetic diversity of temperate maize group ($\pi = 5.0\times10^{-5}$) was slightly higher than that of tropical maize group ($\pi = 4.7\times10^{-5}$), a finding inconsistent with our PCA result. In the PCA plot analysis, only the top two or three principal components could be selected for visualization, and the cumulative interpretation rate was less than 15%, possibly resulting in biased estimates of genetic diversity. Additionally, the temperate maize group included diverse materials from multiple breeding programs, whereas the tropical maize group was selected largely from CIMMYT. LD decay, usually described by decay distance, which indicates the extent of domestication and the intensity of selection, differed between the two groups. When the threshold of r^2 was set to 0.1, the decay distance in temperate germplasm increased to 380 kb, compared to 80 kb in tropical germplasm, indicating that much faster LD decay was detected in tropical maize. Thus, the extent of selection in temperate maize was much greater than that in tropical maize, a finding consistent with the reality that temperate maize has undergone stronger artificial selection in modern breeding. Also, LD decayed to 0.1 within the physical distance of 160 kb, a greater distance than found in an earlier study [82]. The difference between the two LD decay distances may be ascribed to the difference in genome coverage of markers. LD decay varies from group (population) to group (population), especially under different levels of selection intensity. For example, the extents of LD decay in maize landraces, maize cultivars with extensive source and elite maize inbreds are 1 kb [83], 1.5 kb [84], and 100 kb [82], respectively. Commonly, LD decay is expected to be higher in outcrossing than in self-pollinated species. Accordingly, LD decay distances in two self-pollinated species, *Arabidopsis thaliana* and rice (*Oryza sativa* L.), were 50 kb [85] and 100 kb [86], respectively.

4.7. Prospects for improving adaptability of tropical maize

Concerns about the reduction of genetic diversity in commercial hybrids and the depletion of genetic diversity in gene banks make it imperative to introduce exotic germplasm to broaden germplasm resources [87]. Compared with conventional backcrossing and phenotypic selection, rapid-cycle genome-wide marker-assisted selection strategies, such as GS and marker-assisted recurrent selection (MARS), have been shown to be faster and more effective in incorporating favorable alleles and haplotypes [88]. Furthermore, the development of sequencing technology has greatly reduced the cost and time required for genotyping, making it possible to implement MARS and GS on a large scale. Favorable haplotypes for flowering time and photoperiod sensitivity identified in the present study might provide a basis for introducing tropical maize to temperate regions.

5. Conclusions

An understanding of the genetic determinants of flowering time and photoperiod sensitivity is usually considered to be a prerequisite for successful exchange of germplasm resources across regions adapted to different latitudes. We identified selective signatures, candidate genes, and favorable

haplotypes influencing flowering time and photoperiod sensitivity using selective signature analysis and GWAS in large populations. These genes and haplotypes with different allele distributions across germplasm groups or populations provide a resource for understanding genetic differences in the genome and for improving both temperate and tropical maize. Also, the rapid LD decay in the tropical group suggests that the tropical group might contain more rare alleles and undergo more recombination. GO enrichment analysis showed that these genes were mainly enriched in GO categories associated with biological regulation and biosynthesis pathways, indicating that genetic changes between temperate and tropical maize groups have occurred mainly in genomic regions influencing biological regulation and biosynthesis.

Declaration of competing interest

Authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (2016YFD0101803), the Agricultural Science and Technology Innovation Program (ASTIP) of CAAS, and Fundamental Research Funds for Central Non-Profit of Institute of Crop Science, CAAS (1610092016124). Y. Xu was also supported by the Bill and Melinda Gates Foundation and the CGIAR Research Program MAIZE.

Appendix A. Supplementary data

Supplementary data for this article can be found online at https://doi.org/10.1016/j.cj.2019.09.012.

REFERENCES

[1] D.R. Piperno, A.J. Ranere, I. Holst, J. Iriarte, R. Dickau, Starch grain and phytolith evidence for early ninth millennium BP maize from the Central Balsas River Valley, Mexico, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 5019–5024.

[2] Y. Matsuoka, Y. Vigouroux, M.M. Goodman, J.J. Sanchez Gonzalez, E.S. Buckler, J. Doebley, A single domestication for maize shown by multilocus microsatellite genotyping, Proc. Natl. Acad. Sci. USA 99 (2002) 6080–6084.

[3] J. Van Heerwaarden, J. Doebley, W.H. Briggs, J.C. Glaubitz, M. M. Goodman, J.J. Sanchez Gonzalez, J. Ross-Ibarra, Genetic signals of origin, spread, and intro-gression in a large sample of maize landraces, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 1088–1092.

[4] Y. Lu, T. Shah, Z. Hao, S. Taba, S. Zhang, S. Gao, J. Liu, M. Cao, J. Wang, A.B. Prakash, T. Rong, Y. Xu, Comparative SNP and haplotype analysis reveals a higher genetic diversity and rapider LD decay in tropical than temperate germplasm in maize, PLoS ONE (2011) e24861.

[5] Q.O. Oloyede-Kamiyo, S.O. Ajala, A.O. Job, Potential in a collection of adapted and exotic tropical maize inbred lines as resistance source for stem borers, J. Plant Breed. Crop Sci. 10 (2018) 183–190.

[6] P.F. Ribeiro, B. Badu-Apraku, V.E. Gracen, E.Y. Danquah, A.L. Garcia-Oliveira, M.D. Asante, C. Afriyie-Debrah, M. Gedil, Identification of quantitative trait loci for grain yield and other traits in tropical maize under high and low soil-nitrogen environments, Crop Sci. 58 (2018) 321–331.

[7] M.C. Kuki, C.A. Scapim, E.S. Rossi, C.A. Mangolin, A.T. Amaral, R.J.B. Pinto, Genome wide association study for gray leaf spot resistance in tropical maize core, PLoS ONE 13 (2018), e0199539.

[8] C. Giauffret, J. Lothrop, D. Dorvillez, B. Gouesnard, M. Derieux, Genotype × environment interactions in maize hybrids from temperate or highland tropical origin, Crop Sci. 40 (2000) 1004–1012.

[9] E. Stephenson, S. Estrada, X. Meng, J. Ourada, M.G. Muszynski, J.E. Habben, O.N. Danilevskaya, Over-expression of the photoperiod response regulator *ZmCCT10* modifies plant architecture, flowering time and inflorescence morphology in maize, PLoS ONE 14 (2019), e0203728.

[10] C. Wang, Y. Chen, L. Ku, T. Wang, Z. Sun, F. Cheng, L. Wu, Mapping QTL associated with photoperiod sensitivity and assessing the importance of QTL×Environment interaction for flowering time in maize, PLoS ONE 5 (2010), e14068.

[11] R.P. Koester, P.H. Sisco, C.W. Stuber, Identification of quantitative trait loci controlling days to flowering and plant height in two near-isogenic lines of maize, Crop Sci. 33 (1993) 1209–1216.

[12] R. Moutiq, J.M. Ribaut, G.O. Edmeades, M.D. Krakowsky, M. Lee, Elements of genotype–environment interaction: genetic components of the photoperiod response in maize, in: M.S. Kang (Ed.), Quantitative Genetics, Genomics, and Plant Breeding, CABI Publishing, Wallingford, Oxfordshire, the United Kingdom 2002, pp. 257–267.

[13] Y. Xu, Global view of QTL: rice as a model, in: M.S. Kang (Ed.), Quantitative Genetics, Genomics, and Plant Breeding, CABI Publishing, Wallingford, Oxfordshire, the United Kingdom 2002, pp. 109–134.

[14] M. Maheswaran, N. Huang, S.R. Sreerangasamy, S.R. McCouch, Mapping quantitative trait loci associated with days to flowering and photoperiod sensitivity in rice (*Oryza sativa* L.), Mol. Breed. 6 (2000) 145–155.

[15] E.S. Buckler, 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. Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kreso-vich, M.D. McMullen, The genetic architecture of maize flowering time, Science 325 (2009) 714–718.

[16] J. Xu, Y. Liu, J. Liu, M. Cao, J. Wang, H. Lan, Y. Xu, Y. Lu, G. Pan, T. Rong, The genetic architecture of flowering time and photoperiod sensitivity in maize as revealed by QTL review and meta analysis, J. Integr. Plant Biol. 54 (2012) 358–373.

[17] N.N. Kuleshov, World's diversity of phenotypes of maize, J. Am. Soc. Agron. (1933) 688–700.

[18] S. Salvi, S. Castelletti, R. Tuberosa, An updated consensus map for flowering time QTLs in maize, Maydica 54 (2009) 501–512.

[19] X. Zhang, B. Tang, W. Liang, Y. Zheng, F. Qiu, Quantitative genetic analysis of flowering time, leaf number and photoperiod sensitivity in maize (*Zea mays* L.), J. Plant Breed. Crop Sci. 3 (2011) 168–184.

[20] D. Van Inghelandt, A.E. Melchinger, C. Lebreton, B. Stich, Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers, Theor. Appl. Genet. 120 (2010) 1289–1299.

[21] J. Yan, T. Shah, M.L. Warburton, E.S. Buckler, M.D. McMullen, J. Crouch, Genetic characterization and linkage

disequilibrium estimation for a global collection using SNP markers, PLoS ONE 4 (2009), e8451.
[22] Y. Lu, S. Zhang, T. Shah, C. Xie, Z. Hao, X. Li, M. Farkhari, J.M. Ribaut, M. Cao, T. Rong, Y. Xu, Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 19585–19590.
[23] F.B. Mokry, M.E. Buzanskas, M.A. Mudadu, D.A. Grossi, R.H. Higa, R.V. Ventura, A.O. Lima, M. Sargolzaei, S.L.C. Meirelles, F.S. Schenkel, M.V.G.B. Silva, S.C.M. Niciura, M.M. Alencar, D. P. Munari, L.C.A. Regitano, Linkage disequilibrium and haplotype block structure in a composite beef cattle breed, BMC Genomics 15 (2014) S6.
[24] G.C.L. Johnson, L. Esposito, B.J. Barratt, A.N. Smith, J. Heward, G.D. Genova, H. Ueda, H.J. Cordell, I.A. Eaves, F. Dudbridge, R. C.J. Twells, F. Payne, W. Hughes, S. Nutland, H. Stevens, P. Carr, E. Tuomilehto-Wolf, J. Tuomilehto, S.C.L. Gough, D.G. Clayton, J.A. Todd, Haplotype tagging for the identification of common disease genes, Nat. Genet. 29 (2001) 233–237.
[25] D.E. Reich, M. Cargill, S. Bolk, J. Ireland, P.C. Sabeti, D.J. Richter, T. Lavery, R. Kouyoumjian, S.F. Farhadian, R. Ward, E. S. Lander, Linkage disequilibrium in the human genome, Nature 411 (2001) 199–204.
[26] J.M. Smith, J. Haigh, The hitch-hiking effect of a favourable gene, Genet. Res. 23 (1974) 23–35.
[27] N.L. Kaplan, R.R. Hudson, C.H. Langley, The "hitchhiking effect" revisited, Genetics 123 (1989) 887–899.
[28] T. Paape, R.V. Briskine, H.E.L. Lischer, G. Halstead-Nussloch, R.S. Inatsugi, M. Hatekayama, K. Tanaka, T. Nishiyama, R. Sabirov, J. Sese, K.K. Shimizu, Patterns of polymorphism and selection in the subgenomes of the allopolyploid *Arabidopsis kamchatica*, Nat. Commun. 9 (2018) 3909.
[29] S. Wright, The genetical structure of populations, Ann. Eugenics 15 (1949) 323–354.
[30] J.M. Akeya, A.L. Ruheb, D.T. Akeya, A.K. Wongb, C.F. Connellya, J. Madeoya, T.J. Nicholasa, M.W. Neff, Tracking footprints of artificial selection in the dog genome, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 1160–1165.
[31] P.C. Sabeti, D.E. Reich, J.M. Higgins, H.Z. Levine, D.J. Richter, S. F. Schaffner, S.B. Gabriel, J.V. Platko, N.J. Patterson, G.J. McDonald, H.C. Ackerman, S.J. Campbell, D. Altshuler, R. Cooperk, D. Kwiatkowski, R. Ward, E.S. Lander, Detecting recent positive selection in the human genome from haplotype structure, Nature 419 (2002) 832–837.
[32] B.F. Voight, S. Kudaravalli, X. Wen, J.K. Pritchard, A map of recent positive selection in the human genome, PLoS Biol. 4 (2006) e72.
[33] P.C. Sabeti, P. Varilly, B. Fry, J. Lohmueller, E. Hostetter, C. Cotsapas, X. Xie, E.H. Byrne, S.A. McCarroll, R. Gaudet, S.F. Schaffner, E.S. Lander, The International HapMap Consortium. Genome-wide detection and characterization of positive selection in human populations, Nature 449 (2007) 913–919.
[34] F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, Genetics 123 (1989) 585–595.
[35] H. Chen, N. Patterson, D. Reich, Population differentiation as a test for selective sweeps, Genome Res. 20 (2010) 393–402.
[36] M. Li, S. Tian, L. Jin, G. Zhou, Y. Li, Y. Zhang, T. Wang, C.K.L. Yeung, L. Chen, J. Ma, J. Zhang, A. Jiang, J. Li, C. Zhou, J. Zhang, Y. Liu, X. Sun, H. Zhao, Z. Niu, P. Lou, L. Xian, X. Shen, S. Liu, S. Zhang, M. Zhang, L. Zhu, S. Shuai, L. Bai, G. Tang, H. Liu, Y. Jiang, M. Mai, J. Xiao, X. Wang, Q. Zhou, Z. Wang, P. Stothard, M. Xue, X. Gao, Z. Luo, Y. Gu, H. Zhu, X. Hu, Y. Zhao, G. Plastow, J. Wang, Z. Jiang, K. Li, N. Li, X. Li, R. Li, Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars, Nat. Genet. 45 (2013) 1431–1439.
[37] Q. Qiu, L. Wang, K. Wang, Y. Yang, T. Ma, Z. Wang, X. zhang, Z. Ni, F. Hou, R. Long, R. Abbott, J. Lenstra, J. Liu, Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions, Nat. Commun. 6 (2015) 10283.
[38] D. Lan, X. Xiong, T.D. Mipam, C. Fu, Q. Li, Y. Ai, D. Hou, Z. Chai, J. Zhong, J. Li, Genetic diversity, molecular phylogeny and selection evidence of jinchuan yak revealed by whole-genome resequencing, *G3-Genes*, Genom Genet. 8 (2018) 300572.
[39] J. Qi, X. Liu, D. Shen, H. Miao, B. Xie, X. Li, P. Zeng, S. Wang, Y. Shang, X. Gu, Y. Du, Y. Li, T. Lin, J. Yuan, X. Yang, J. Chen, H. Chen, X. Xiong, K. Huang, Z. Fei, L. Mao, L. Tian, T. Städler, S.S. Renner, S. Kamoun, W.J. Lucas, Z. Zhang, S. Huang, A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity, Nat. Genet. 45 (2013) 1510–1518.
[40] Y. Jiao, H. Zhao, L. Ren, W. Song, B. Zeng, J. Guo, B. Wang, Z. Liu, J. Chen, W. Li, M. Zhang, S. Xie, J. Lai, Genome-wide genetic changes during modern breeding of maize, Nat. Genet. 44 (2012) 812–817.
[41] Z. Zhou, Y. Jiang, Z. Wang, Z. Gou, J. Lyu, W. Li, Y. Yu, L. Shu, Y. Zhao, Y. Ma, C. Fang, Y. Shen, T. Liu, C. Li, Q. Li, M. Wu, M. Wang, Y. Wu, Y. Dong, W. Wan, X. Wang, Z. Ding, Y. Gao, H. Xiang, B. Zhu, S. Lee, W. Wang, Z. Tian, Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean, Nat. Biotechnol. 33 (2015) 408.
[42] C. Xu, Y. Ren, Y. Jian, Z. Guo, Y. Zhang, C. Xie, J. Fu, H. Wang, G. Wang, Y. Xu, P. Li, C. Zou, Development of a maize 55 K SNP array with improved genome coverage for molecular breeding, Mol. Breed. 37 (2017) 20.
[43] Y. Xu, L. Luo, Biotechnology and germplasm resource management in rice, in: Shengxiang Tang (Ed.), Rice Germplasm Resources, Hubei Science and Technology Publisher, Wuhan, China 2002, pp. 229–250 , (in Chinese).
[44] M. Maheswaran, N. Huang, S.R. Sreerangasamy, S.R. McCouch, Mapping quantitative trait loci associated with days to flowering and photoperiod sensitivity in rice (*Oryza sativa* L.), Mol. Breed. 6 (2000) 145–155.
[45] P.J. Bradbury, Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S. Buckler, TASSEL: software for association mapping of complex traits in diverse samples, Bioinformatics 23 (2007) 2633–2635.
[46] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation, Bioinformatics 23 (2006) 127–128.
[47] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira, D. Bender, J. Maller, P. Sklar, P.I.W. Bakker, Mark J. Daly, P.C. Sham, PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (2007) 559–575.
[48] M. Nei, W. Li, Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 5269–5273.
[49] R.R. Hudson, M. Slatkin, W.P. Maddison, Estimation of levels of gene flow from DNA sequence data, Genetics 132 (1992) 583–589.
[50] B. Pfeifer, U. Wittelsbu, S.E. Ramos-Onsins, M.J. Lercher, PopGenome: an efficient Swiss army knife for population genomic analyses in R, Mol. Biol. Evol. 31 (2014) 1929–1936.
[51] Z. Du, X. Zhou, Y. Ling, Z. Zhang, Z. Su, agriGO: a GO analysis toolkit for the agricultural community, Nucleic Acids Res. 38 (2010) W64–W70.
[52] S.B. Gabriel, S.F. Schaffner, H. Nguyen, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S. Lander, M.J. Daly, D. Altshuler, The structure of haplotype blocks in the human genome, Science 296 (2002) 2225–2229.
[53] Z. Zhang, E. Ersoz, C. Lai, R.J. Todhunter, H.K. Tiwari, M.A. Gore, P.J. Bradbury, J. Yu, D.K. Arnett, J.M. Ordovas, E.S.

Buckler, Mixed linear model approach adapted for genome-wide association studies, Nat. Genet. 42 (2010) 355–360.
[54] H.A. Agrama, G.C. Eizenga, W. Yan, Association mapping of yield and its components in rice cultivars, Mol. Breed. 19 (2007) 341–356.
[55] Y. Li, C. Li, 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, T. Wang, Identification of genetic variants associated with maize flowering time using an extremely large multi-enetic background population, Plant J. 86 (2016) 391–402.
[56] J.A.R. Navarro, M. Wilcox, J. Burgueño, C. Romay, K. Swarts, S. Trachsel, E. Preciado, A. Terron, H.V. Delgado, V. Vidal, A. Ortega, A.E. Banda, N.O.G. Montiel, I. Ortiz-Monasterio, F.S. Vicente, A.G. Espinoza, G. Atlin, P. Wenzl, S. Hearne, E.S. Buckler, A study of allelic diversity underlying flowering-time adaptation in maize landraces, Nat. Genet. 49 (2017) 476–483.
[57] H. Liu, X. Wang, M.L. Warburton, W. Wen, M. Jin, M. Deng, J. Liu, H. Tong, Q. Pan, X. Yang, J. Yan, Genomic, transcriptomic, and phenomic variation reveals the complex adaptation of modern maize breeding, Mol. Plant 8 (2015) 871–884.
[58] L. Fang, Q. Wang, Y. Hu, Y. Jia, J. Chen, B. Liu, Z. Zhang, X. Guan, S. Chen, B. Zhou, G. Mei, J. Sun, Z. Pan, S. He, S. Xiao, W. Shi, W. Gong, J. Liu, J. Ma, C. Cai, X. Zhu, W. Guo, X. Du, T. Zhang, Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits, Nat. Genet. 49 (2017) 1089–1097.
[59] X. Wei, K. Liu, Y. Zhang, Q. Feng, L. Wang, Y. Zhao, D. Li, Q. Zhao, X. Zhu, X. Zhu, W. Li, D. Fan, Y. Gao, Y. Lu, X. Zhang, X. Tang, C. Zhou, C. Zhu, L. Liu, R. Zhong, Q. Tian, Z. Wen, Q. Weng, B. Han, X. Huan, X. Zhang, Genetic discovery for oil production and quality in sesame, Nat. Commun. 6 (2015) 8609.
[60] Y. Chen, S. Chang, L. Wu, A preliminary genetic study on the basic vegetative growth stage of the hybrids between subtropical and temperate lines in maize, Acta Agric. Boreali-Sin. 15 (2000) 15–20(in Chinese with English abstract).
[61] E.S. Buckler, B.S. Gaut, M.D. McMullen, Molecular and functional diversity of maize, Curr. Opin. Plant. Biol. 9 (2006) 172–176.
[62] Y. Vigouroux, S. Mitchell, Y. Matsuoka, M. Hamblin, S. Kresovich, J.S.C. Smith, J. Jaqueth, O.S. Smith, J. Doebley, An analysis of genetic diversity across the maize genome using microsatellites, Genetics 169 (2005) 1617–1630.
[63] X. Xia, J.C. Reif, A.E. Melchinger, M. Frisch, D.A. Hoisington, D. Beck, K. Pixley, M.L. Warburton, Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: I. Subtropical, tropical midlatitude, and highland maize inbred lines and their relationships with elite US and European maize, Crop Sci. 45 (2005) 2573–2582.
[64] D.N. Duvick, J.S.C. Smith, M. Cooper, Changes in performance, parentage, and genetic diversity of successful corn hybrids, 1930–2000, in: C.W. Smith, J. Betran, E.C.A. Runge (Eds.),Corn: origin, history, Technology, and Production, Wiley, New York, USA 2004, pp. 65–97.
[65] D. Hoisington, M. Khairallah, T. Reeves, J.M. Ribaut, B. Skovmand, S. Taba, M. Warburton, Plant genetic resources: what can they contribute toward increased crop productivity? Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 5937–5943.
[66] J.M. Ribaut, C. Jiang, D. Gonzalez-de-Leon, G.O. Edmeades, D. A. Hoisington, Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies, Theor. Appl. Genet. 94 (1997) 887–896.
[67] H. Campos, M. Cooper, G.O. Edmeades, C. Löffler, J.R. Schussler, M. Ibañez, Changes in drought tolerance in maize associated with fifty years of breeding for yield in the US corn belt, Maydica 51 (2006) 369.
[68] B. Gouesnard, C. Rebourg, C. Welcker, A. Charcosset, Analysis of photoperiod sensitivity within a collection of tropical maize populations, Genet. Resour. Crop. Evol. 49 (2002) 471–481.
[69] K. Liu, M. Goodman, S. Muse, J.S. Smith, E. Buckler, J. Doebley, Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites, Genetics 165 (2003) 2117–2128.
[70] S. Ducrocq, C. Giauffret, D. Madur, V. Combes, F. Dumas, S. Jouanne, D. Coubriche, P. Jamin, L. Moreau, A. Charcosset, Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10, Genetics 183 (2009) 1555–1563.
[71] D. Enard, P.W. Messer, D.A. Petrov, Genome-wide signals of positive selection in human evolution, Genome Res. 24 (2013) 885.
[72] C. He, J. Fu, J. Zhang, Y. Li, J. Zheng, H. Zhang, X. Yang, J. Wang, G. Wang, A gene-oriented haplotype comparison reveals recently selected genomic regions in temperate and tropical maize germplasm, PLoS ONE 12 (2017), e0169806.
[73] M. Wang, L. Tu, M. Lin, Z. Lin, P. Wang, Q. Yang, L. Zhang, Z. Ye, C. Shen, J. Li, K. Guo, X. Zhou, X. Nie, Z. Li, Y. Ma, C. Huang, S. Jin, L. Zhu, X. Yang, L. Min, D. Yuan, Q. Zhang, K. Lindsey, X. Zhang, Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication, Nat. Genet. 49 (2017) 579–587.
[74] M.B. Hufford, X. Xu, J. van Heerwaarden, T. Pyhäjärvi, J.M. Chia, R.A. Cartwright, R.J. Elshire, J.C. Glaubitz, K.E. Guill, S.M. Kaeppler, J. Lai, P.L. Morrell, L.M. Shannon, C. Song, N.M. Springer, R.A. Swanson-Wagner, P. Tiffin, J. Wang, G. Zhang, J. Doebley, M.D. McMullen, D. Ware, E.S. Buckler, S. Yang, J. Ross-Ibarra, Comparative population genomics of maize domestication and improvement, Nat. Genet. 44 (2012) 808–811.
[75] A. Korte, A. Farlow, The advantages and limitations of trait analysis with GWAS: a review, Plant Methods (2013) 9–29.
[76] M.T. Hamblin, M.L. Warburton, E.S. Buckler, Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness, PLoS ONE 12 (2007), e1367.
[77] Y. Lu, J. Yan, C.T. Guimara˜es, S. Taba, Z. Hao, S. Gao, S. Chen, J. Li, S. Zhang, B.S. Vivek, C. Magorokosho, S. Mugo, D. Makumbi, S.N. Parentoni, T. Shah, T. Rong, J.H. Crouch, Y. Xu, Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms, Theor. Appl. Genet. 120 (2009) 93–115.
[78] R. Nielsen, Genomic scans for selective sweeps using SNP data, Genome Res. 15 (2005) 1566–1575.
[79] D. Lan, X. Xiong, T.D. Mipam, C. Fu, Q. Li, Y. Ai, D. Hou, Z. Chai, J. Zhong, J. Li, Genetic diversity, molecular phylogeny, and selection evidence of Jinchuan yak revealed by whole-genome resequencing, G3-Genes Genomes Genet. 8 (2018) 945–952.
[80] The International HapMap Consortium, Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds, Science 324 (2009) 528–532.
[81] J. Ross-Ibarra, P.L. Morrell, B.S. Gaut, Plant domestication, a unique opportunity to identify the genetic basis of adaptation, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 8641–8648.
[82] M. Jung, A. Ching, D. Bhattramakki, M. Dolan, S. Tingey, M. Morgante, A. Rafalski, Linkage disequilibrium and sequence diversity in a 500-kbp region around the *adh1* locus in elite maize germplasm, Theor. Appl. Genet. 109 (2004) 681–689.
[83] M.I. Tenaillon, M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley, B.S. Gaut, Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.), Proc. Natl. Acad. Sci, U.S.A. 98 (2001) 9161–9166.
[84] D.L. Remington, J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S.R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman, E.S. Buckler, Structure of linkage disequilibrium and phenotypic associations in the maize genome, Proc. Natl. Acad. Sci. U.S. A. 98 (2001) 11479–11484.

[85] M. Nordborg, T. Hu, Y. Ishino, J. Jhaveri, C. Toomajian, H. Zheng, E. Bakker, P. Calabrese, J. Gladstone, R. Goyal, M. Jakobsson, S. Kim, Y. Morozov, B. Padhukasahasram, V. Plagnol, N.A. Rosenberg, C. Shah, J.D. Wall, J. Wang, K. Zhao, T. Kalbfleisch, V. Schulz, M. Kreitman, J. Bergelson, The pattern of polymorphism in *Arabidopsis thaliana*, PLoS Biol. 3 (2005) e196.
[86] A.J. Garris, S.R. McCouch, S. Kresovich, Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa*5 locus of rice, *Oryza sativa* L, Genetics 165 (2003) 759–769.
[87] Y. Xu, Molecular Plant Breeding, CAB International, Wallingford, Oxfordshire, the United Kingdom, 2010.
[88] R. Bernardo, J.M. Yu, Prospects for genomewide selection for quantitative traits in maize, Crop Sci. 47 (2007) 1082–1090.