ZmCCT9 enhances maize adaptation to higher latitudes

Cheng Huang[a], Huayue Sun[a], Dingyi Xu[a], Qiuyue Chen[a], Yameng Liang[a], Xufeng Wang[a], Guanghui Xu[a], Jinge Tian[a], Chenglong Wang[a], Dan Li[a], Lishuan Wu[a], Xiaohong Yang[a], Weiwei Jin[a], John F. Doebley[b,1], and Feng Tian[a,1]

[a]National Maize Improvement Center of China, Beijing Key Laboratory of Crop Genetic Improvement, Laboratory of Crop Heterosis and Utilization, Joint International Research Laboratory of Crop Molecular Breeding, China Agricultural University, Beijing 100193, China; and [b]Department of Genetics, University of Wisconsin, Madison, WI 53706

Contributed by John F. Doebley, December 3, 2017 (sent for review October 16, 2017; reviewed by Nathan M. Springer and Jonathan F. Wendel)

From its tropical origin in southwestern Mexico, maize spread over a wide latitudinal cline in the Americas. This feat defies the rule that crops are inhibited from spreading easily across latitudes. How the widespread latitudinal adaptation of maize was accomplished is largely unknown. Through positional cloning and association mapping, we resolved a flowering-time quantitative trait locus to a Harbinger-like transposable element positioned 57 kb upstream of a CCT transcription factor (*ZmCCT9*). The Harbinger-like element acts in *cis* to repress *ZmCCT9* expression to promote flowering under long days. Knockout of *ZmCCT9* by CRISPR/Cas9 causes early flowering under long days. *ZmCCT9* is diurnally regulated and negatively regulates the expression of the florigen *ZCN8*, thereby resulting in late flowering under long days. Population genetics analyses revealed that the Harbinger-like transposon insertion at *ZmCCT9* and the CACTA-like transposon insertion at another CCT paralog, *ZmCCT10*, arose sequentially following domestication and were targeted by selection for maize adaptation to higher latitudes. Our findings help explain how the dynamic maize genome with abundant transposon activity enabled maize to adapt over 90° of latitude during the pre-Columbian era.

maize adaptation | flowering time | transposable element | domestication | *ZmCCT9*

Flowering time is a major determinant of the local adaptation of crops (1–10). Maize (*Zea mays* ssp. *mays*) was domesticated in southwestern Mexico ∼9,000 y ago from its wild progenitor, teosinte (*Zea mays* ssp. *parviglumis*), a tropical species that exhibits substantial photoperiod sensitivity and requires short-day (SD) conditions to flower (6, 11). As a result of natural and artificial selection, maize has evolved a reduced photoperiod sensitivity to adapt to long-day (LD) environments from latitude 58° north in Canada to 40° south in Chile (12–14). This widespread expansion over 90° of latitude across the North–South axis of the Americas was a remarkable achievement in the pre-Columbian era compared with the spread of other New World crops (i.e., potato and cotton), which remained in narrower latitudinal ranges before the era of modern breeding (15, 16). Understanding how this feat was achieved specifically with maize is an important question in the history of crop domestication.

Transposable elements were first discovered in maize by Barbara McClintock (17) and have since been shown to play important roles in shaping genome evolution and gene regulatory networks of many species (18). Compared with other crops, maize is exceptionally abundant in transposon activity, and nearly 85% of the maize genome consists of transposable elements (19, 20). This abundant transposon activity helps generate substantial genetic diversity upon which selection can act during evolution, as inferred by Barbara McClintock decades ago (21). Although many loci affecting natural variation in flowering time have been detected in maize using various types of mapping populations (6, 13, 22), only two of them, *Vgt1* (23–25) and *ZmCCT* (6, 7, 26), have been well characterized and found to contribute to flowering-time adaptation. Interestingly, both loci are associated with the regulatory roles of transposable elements. Despite this progress, whether transposable elements play a prevalent role in driving maize flowering-time adaptation remains to be further tested and elucidated.

In this study, we report the cloning and characterization of a flowering-time QTL (*qDTA9*) that was previously mapped in a maize-teosinte experimental population (27). We show that *qDTA9* is underlain by a Harbinger-like transposable element located 57 kb upstream of a CCT domain-containing gene, *ZmCCT9*. The Harbinger-like transposable element functions as a repressor of *ZmCCT9* to promote flowering under LD conditions. *ZmCCT9* confers LD-dependent flowering repression by negatively regulating the expression of florigen *ZCN8*. We finally show that the Harbinger-like transposable element arose more recently compared with the CACTA-like transposon insertion at another CCT paralog, *ZmCCT10* (6, 7, 26). Both transposon insertions were targeted by selection and have played crucial roles as maize spread geographically from its tropical origin to higher latitudes.

Results and Discussion

Positional Cloning of *qDTA9*. Using a large population of 866 maize-teosinte BC_2S_3 recombinant inbred lines (RILs), we previously performed quantitative trait locus (QTL) mapping for days to anthesis (DTA) under LD conditions (27) and detected a QTL (*qDTA9*) between markers M114923 and M123776 on chromosome 9 (Fig. 1*A*). To fine-map *qDTA9*, we selected a heterogeneous inbred family (HIF) that was heterozygous only at *qDTA9* (*SI Appendix*, Fig. S1) and used it to generate a large near-isogenic line (NIL) population (n = 5,394). Following a

Significance

Flowering time is a critical determinant of crop adaptation to local environments. As a result of natural and artificial selection, maize has evolved a reduced photoperiod sensitivity to adapt to regions over 90° of latitude in the Americas. Here we show that a distant Harbinger-like transposon acts as a *cis*-regulatory element to repress *ZmCCT9* expression to promote flowering under the long days of higher latitudes. The transposon at *ZmCCT9* and another functional transposon at a second flowering-time gene, *ZmCCT10*, arose sequentially following domestication and were targeted by selection as maize spread from the tropics to higher latitudes. Our results demonstrate that new functional variation created by transposon insertions helped maize to spread over a broad range of latitudes rapidly.

Author contributions: C.H. and F.T. designed research; C.H., H.S., D.X., Q.C., Y.L., X.W., G.X., J.T., C.W., D.L., and L.W. performed research; X.Y., W.J., and J.F.D. contributed new reagents/analytic tools; C.H. and F.T. analyzed data; and C.H. and F.T. wrote the paper.

Reviewers: N.M.S., University of Minnesota; and J.F.W., Iowa State University.

The authors declare no conflict of interest.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. MG233400–MG233921).

[1]To whom correspondence may be addressed. Email: jdoebley@wisc.edu or ft55@cau.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718058115/-/DCSupplemental.

recombinant-derived progeny testing strategy (6, 27, 28), *qDTA9* was delimited to a 2.4-kb noncoding region between markers M115705 and M115707 (Fig. 1 *B* and *C*).

At the same time, a pair of NILs that is homozygous for the maize and teosinte alleles across the *qDTA9* region, designated NIL(maize) and NIL(teosinte), was developed from the HIF for subsequent analysis. Field examination under natural LD conditions showed that NIL(maize) flowered approximately 4 d earlier than NIL(teosinte) (Fig. 1 *D*–*F*), along with reduced plant height and leaf number (*SI Appendix*, Fig. S2). However, no significant differences in flowering time were observed when NIL(maize) and NIL(teosinte) were grown under natural SD conditions (*SI Appendix*, Fig. S3). These results suggest that *qDTA9* is regulated by photoperiod.

Association Mapping for *qDTA9*. To further identify potential causative variants responsible for *qDTA9*, we performed an association analysis by resequencing the 2.4-kb causative region in an association panel containing 513 diverse maize inbred lines that was scored for DTA under LD conditions in Beijing (39.9°N, 116.4°E) (7). In total, 41 variants (SNPs and insertions and deletions) with minor allele frequencies ≥0.05 were identified and tested for association with DTA using a mixed linear model (29). Overall, 13 variants exhibited significant association with

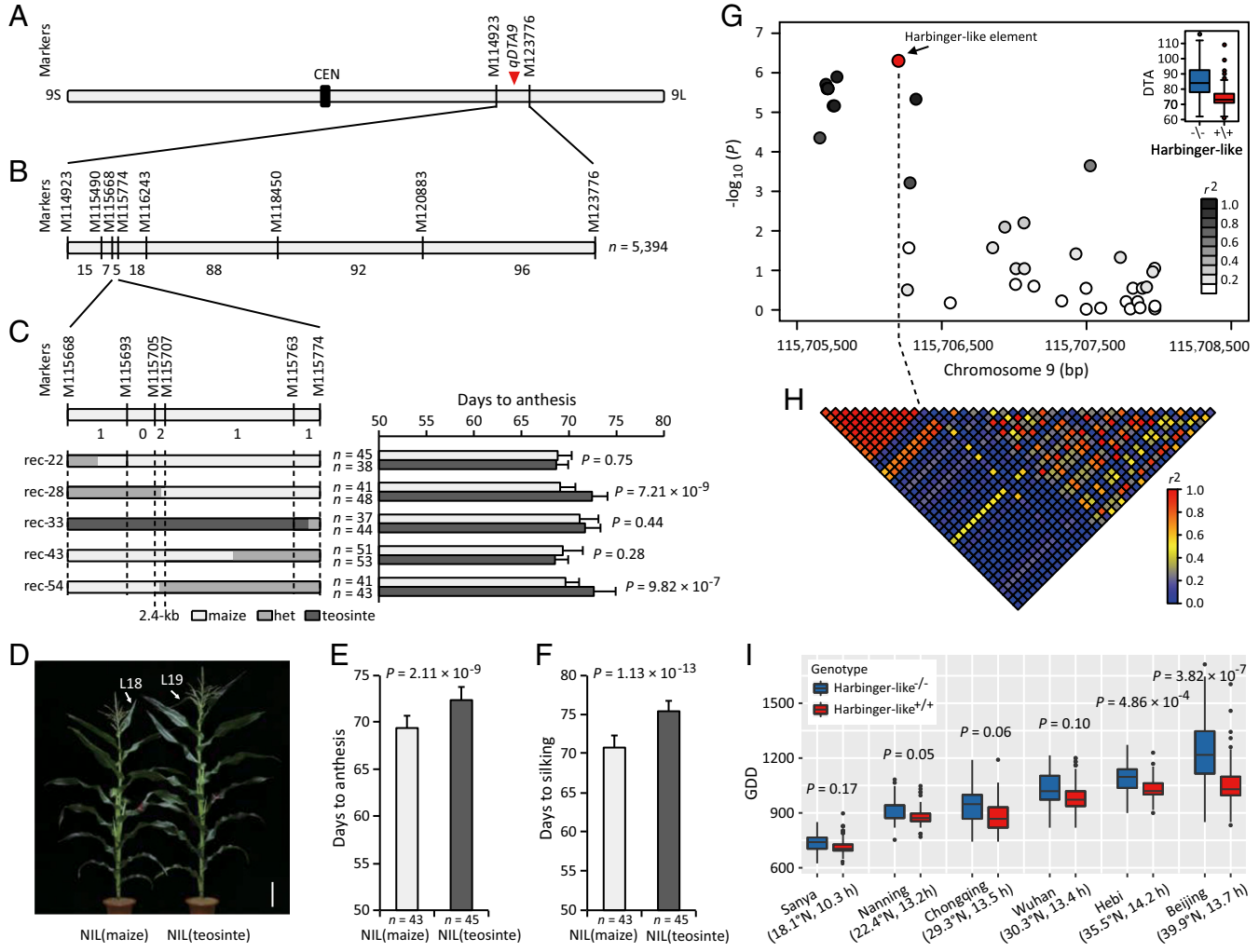

Fig. 1. Fine mapping and association analysis of *qDTA9*. (*A*) Chromosomal location of *qDTA9* on chromosome 9. CEN, centromere. (*B*) Fine mapping of *qDTA9* using a NIL population (n = 5,394) derived from a HIF that was heterozygous only at *qDTA9*. The number of recombinants between the adjacent markers is indicated below the linkage map. (*C*) Progeny testing of recombinants delimited the *qDTA9* locus to a 2.4-kb noncoding region between markers M115705 and M115707. (*Left*) The graphical genotypes of the five critical recombinants. (*Right*) The homozygous progenies (homozygous recombinants and homozygous nonrecombinants) identified within each recombinant-derived F_3 family. Light-gray, gray, and dark-gray bars represent regions homozygous for maize, heterozygous regions, and regions homozygous for teosinte, respectively. Homozygous progenies of each recombinant were phenotyped and compared to determine the genotype of *qDTA9* of the parental recombinant. (*D*) Gross morphologies of NIL(maize) and NIL(teosinte). The uppermost leaf (L) is indicated by a white arrow. (Scale bar, 20 cm.) (*E* and *F*) Phenotypic comparison between NIL(maize) and NIL(teosinte). (*E*) Days to anthesis. (*F*) Days to silking. (*G*) Association analysis of the 2.4-kb causative region for *qDTA9* in a panel of 513 diverse maize inbred lines. The Harbinger-like element showing the most significant association is indicated by a red dot. The intensity of gray shading indicates the level of linkage disequilibrium (r^2) between the Harbinger-like element and other variants identified in the 2.4-kb region. (*H*) Pairwise linkage disequilibrium analysis of variants identified in the 2.4-kb region. The intensity of red shading indicates the level of linkage disequilibrium (r^2) between variants. (*I*) Association testing of the Harbinger-like element with flowering time scored in five other locations at different latitudes. Days to anthesis were converted to growing degree days (GDD) to account for the effect of temperature differences among environments. The latitude and the mean day length of each location are indicated in parentheses. The data in *C*, *E*, and *F* are shown as means ± SD; *P* values were determined by Student's *t* test. Harbinger-like$^{+/+}$ and Harbinger-like$^{-/-}$ in *G* and *I* denote inbred lines containing the Harbinger-like element and inbred lines lacking the Harbinger-like element, respectively.

DTA after Bonferroni multiple test correction ($P \leq 2.44 \times 10^{-4}$) (Fig. 1*G*). Interestingly, a 360-bp Harbinger-like transposable element that is present in 41.3% of maize inbred lines showed the strongest association with DTA ($P = 5.26 \times 10^{-7}$), with the Harbinger-like element promoting flowering (Fig. 1*G*). As expected, the maize parent of the BC_2S_3 population carried the Harbinger-like element, whereas the teosinte parent lacked the Harbinger-like element (*SI Appendix*, Fig. S4 *A* and *B*). The Harbinger-like element is in strong linkage disequilibrium ($r^2 > 0.5$) with all other significant variants, creating a linkage disequilibrium block flanking the Harbinger-like element (Fig. 1*H*). We also tested the association of the Harbinger-like element with flowering time scored in five other locations at diverse latitudes (7). To account for the effect of temperature differences among environments on the observed flowering times, DTA were converted to growing degree days (30). Interestingly, the Harbinger-like element exhibited significant associations with flowering time only at higher latitudes (Fig. 1*I*), potentially indicating a role for the Harbinger-like element in latitudinal adaptation.

According to the B73 reference genome (AGPv2), the Harbinger-like transposable element is located 57 kb upstream of *GRMZM2G004483* (*SI Appendix*, Fig. S4*A*), which encodes a CCT domain protein homologous to the rice photoperiod response regulator *Ghd7* (*SI Appendix*, Fig. S5) (3). *GRMZM2G004483* shows 36% protein sequence identity to *ZmCCT* (*SI Appendix*, Fig. S6), a CCT paralog residing on chromosome 10 that was recently shown to play an important role in the regulation of maize photoperiod flowering (6, 7, 26). To distinguish these two CCT genes, we hereafter refer to them as *ZmCCT9* and *ZmCCT10* based on the chromosomes on which they reside. We speculate that *ZmCCT9* is most likely the target gene regulated by *qDTA9* and that the Harbinger-like element in the 2.4-kb causative region for *qDTA9* is a distant *cis*-regulatory variant of *ZmCCT9* that alters *ZmCCT9* expression to control flowering time.

Knockout of *ZmCCT9* Causes Early Flowering Under Long Days. To validate the function of *ZmCCT9*, we knocked out the endogenous *ZmCCT9* gene in maize inbred line ZC01 using CRISPR/Cas9 technology (31, 32). Two 20-bp sequences in the first and second exons of *ZmCCT9* were selected as target sites for Cas9 cleavage (Fig. 2*A*). PCR and sequencing analysis of 20 independent first-generation (T0) transgenic lines identified three homozygous mutations with deletions in the target sites that truncated the *ZmCCT9* ORF; these are referred to as KO#1, KO#2, and KO#3 (Fig. 2*B* and *SI Appendix*, Fig. S7). These three null mutants were self-pollinated, and the resulting T1 progenies were planted in natural LD conditions to investigate their flowering times. Notably, all three null mutants (KO#1, KO#2, and KO#3) flowered significantly earlier than wild-type plants under LD conditions (Fig. 2 *C*–*E*), strongly indicating that *ZmCCT9* is involved in the regulation of flowering time in maize.

***ZmCCT9* Is Diurnally Regulated and Exhibits Differential Expression Between the Maize Allele and the Teosinte Allele.** To elucidate the cellular localization of ZmCCT9, we introduced the *ZmCCT9-GFP* and *ZmCCT9-EGFP* fusion genes under the control of the

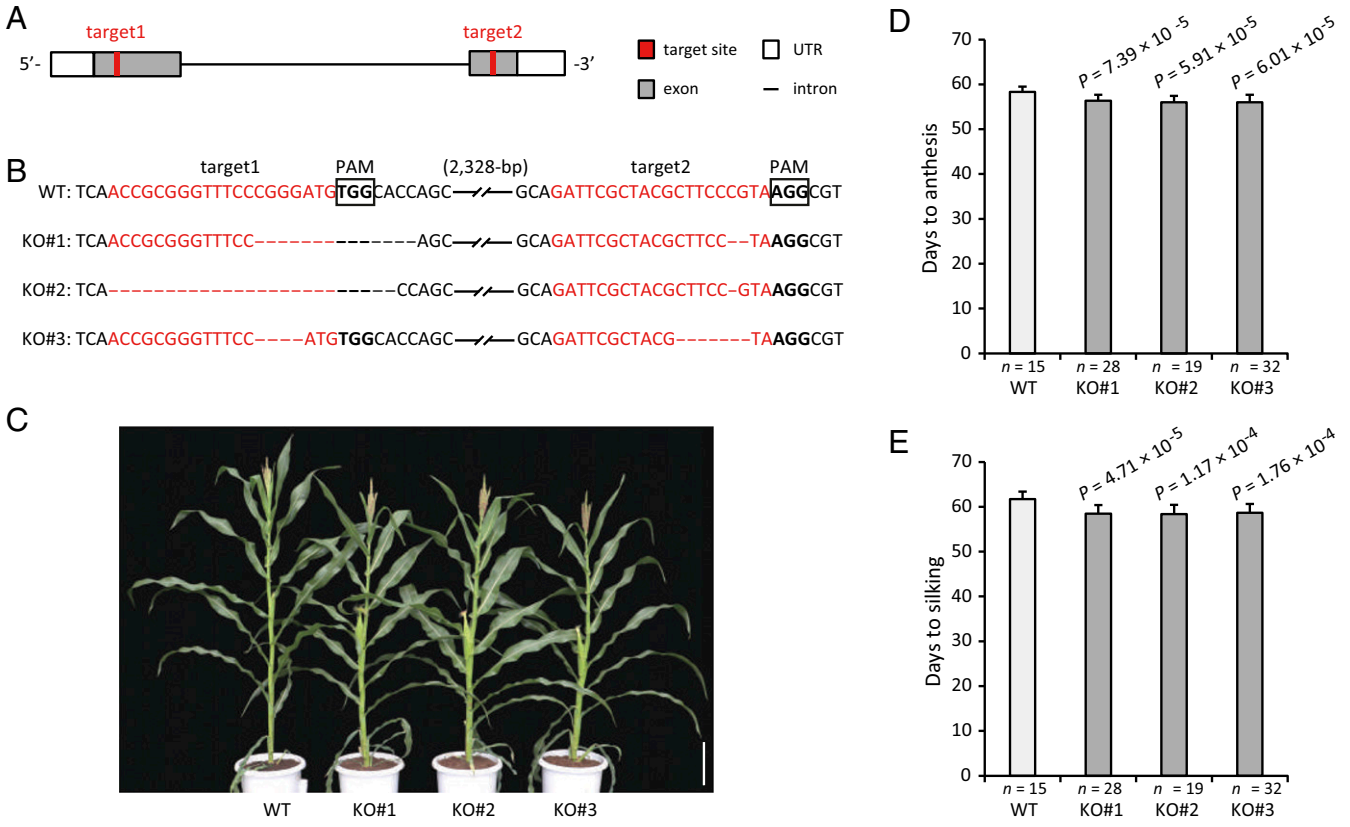

Fig. 2. Knockout of *ZmCCT9* by CRISPR/Cas9 causes early flowering under natural LD conditions. (*A*) Gene structure of *ZmCCT9* and the two target sites (red boxes) in the first and second exons of *ZmCCT9* for CRISPR/Cas9 editing. (*B*) Sequences of three homozygous knockout lines with deletions in target sites that truncated the *ZmCCT9* ORF (KO#1, KO#2, and KO#3). The wild-type sequence is shown at the *Top*. Target sites and protospacer-adjacent motif (PAM) sequences are highlighted in red and boldface fonts, respectively, and deletions are indicated by dashes. The sequence gap length is shown in parentheses. (*C*) Gross morphologies of wild-type and knockout lines. (Scale bar, 20 cm.) (*D* and *E*) Phenotypic comparison of wild-type and three knockout lines under natural LD conditions. (*D*) Days to anthesis. (*E*) Days to silking. The data in *D* and *E* are shown as means ± SD; *P* values were determined by Student's *t* test.

CaMV35S promoter into onion epidermal cells and maize protoplasts, respectively. The results showed that the GFP signal is localized to the nucleus (Fig. 3*A*), supporting a role for ZmCCT9 as a nuclear transcription factor. To determine whether the upstream Harbinger-like element altered *ZmCCT9* expression, we investigated the spatiotemporal expression patterns of *ZmCCT9* in NIL(maize) and NIL(teosinte) under LD conditions. *ZmCCT9* was expressed mainly in mature leaves (Fig. 3*B*) and showed higher expression in NIL(teosinte) than in NIL(maize) at all examined growth stages, with a maximum difference in expression at the V4 stage (Fig. 3*C*). RNA in situ hybridization analysis showed that *ZmCCT9* primarily accumulates in leaf vascular bundles and sclerenchyma fiber cells (Fig. 3*D*).

Genes involved in photoperiod flowering are diurnally regulated (2–4, 7, 33–37). We therefore examined the diurnal expression patterns of *ZmCCT9* in mature leaves of NIL(maize) and NIL(teosinte) under artificial LD and SD conditions. Under LD conditions, *ZmCCT9* transcript levels in NIL(teosinte) showed strong diurnal oscillation, with the highest expression 3 h after dawn; in contrast, the amplitude of *ZmCCT9* expression in

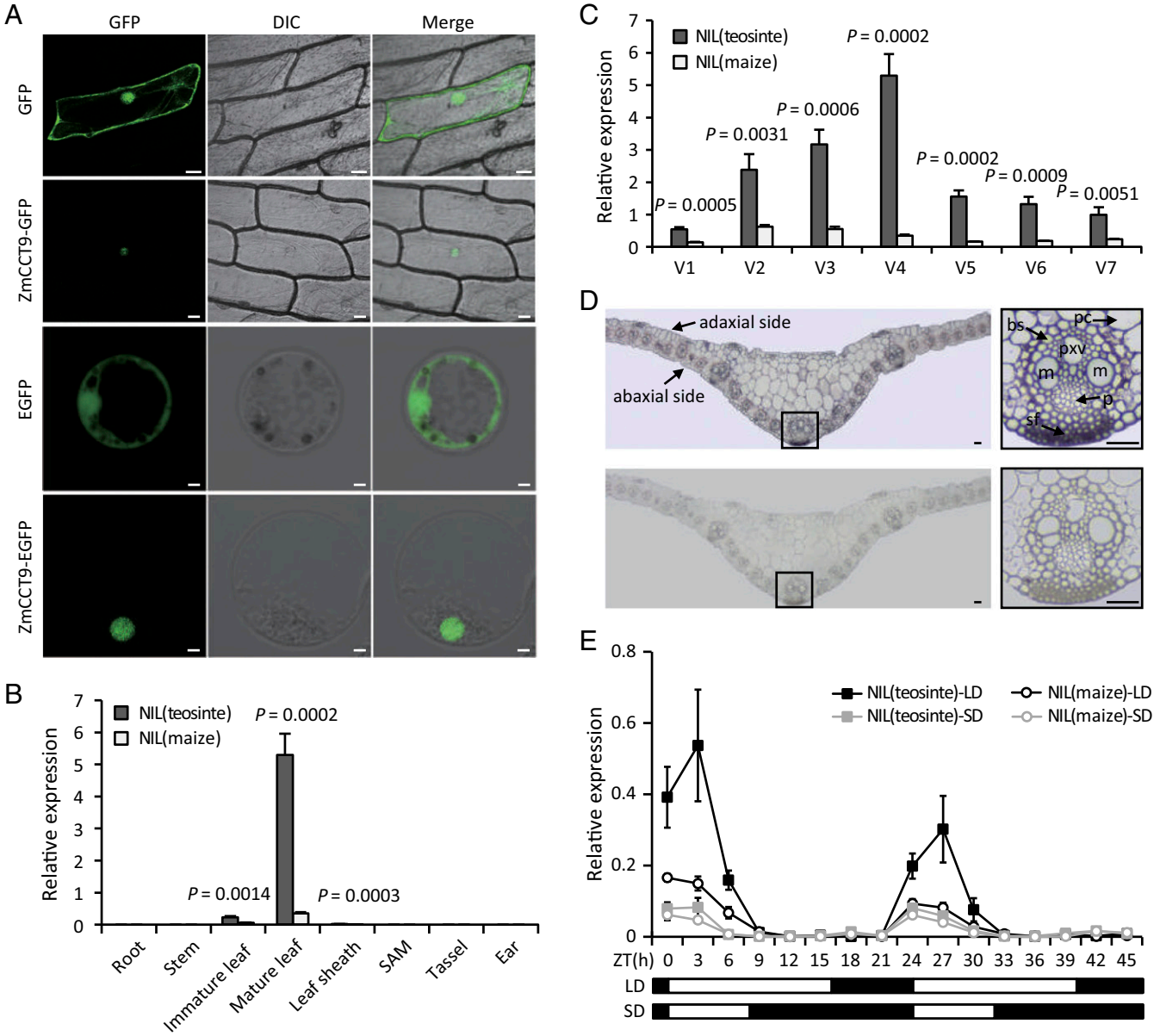

Fig. 3. *ZmCCT9* expression features. (*A*) ZmCCT9 protein is localized in the nucleus in onion epidermal cells (*Top* two rows) (Scale bars, 50 μm.) and maize protoplasts (*Bottom* two rows). (Scale bars, 5 μm.) (*B*) Spatial expression pattern of *ZmCCT9* in NIL(maize) and NIL(teosinte) in various tissues under natural LD conditions. (*C*) Temporal expression pattern of *ZmCCT9* in mature leaves of NIL(maize) and NIL(teosinte) at various growth stages under natural LD conditions. (*D*) RNA in situ hybridization analysis of *ZmCCT9*. (*Top*) The hybridization signal obtained using a *ZmCCT9* DIG-labeled antisense probe in transverse sections of leaf blade tips from NIL(teosinte) at the V4 stage under natural LD conditions. (*Bottom*) The negative control for hybridization using a sense probe. An individual vascular bundle from near the midrib (indicated by the outlined square) is shown on the *Right*. (Scale bars, 50 μm.) bs, bundle sheath; m, metaxylem vessels; p, phloem; pc, parenchyma cells; pxv, protoxylem vessels; sf, sclerenchyma fibers. (*E*) Diurnal expression patterns of *ZmCCT9* in mature leaves of NIL(maize) and NIL(teosinte) under artificial LD and SD conditions. The black bars and white bars indicate the dark period and the light period, respectively. ZT, zeitgeber time. The data in *B*, *C*, and *E* are relative to the control gene *ZmTubulin1* and represent means ± SEM of three biological replicates; *P* values were determined by Student's *t* test.

NIL(maize) was significantly attenuated (Fig. 3*E*). Under SD conditions, *ZmCCT9* in NIL(maize) and NIL(teosinte) showed similar weak diurnal expression patterns (Fig. 3*E*). These results suggest that *ZmCCT9* is diurnally regulated and that the maize allele exhibits a much reduced photoperiod sensitivity relative to the teosinte allele, likely due to the repressive effect of the Harbinger-like element in the 2.4-kb causative region for *qDTA9*.

The Harbinger-Like Element Acts in *Cis* to Repress *ZmCCT9* Expression to Promote Flowering Under Long Days. To test the effect of the Harbinger-like element on gene expression, we cloned the maize and teosinte sequences that differed only at the Harbinger-like element (*SI Appendix*, Fig. S8) into reporter constructs upstream of the minimal promoter of the cauliflower mosaic virus (mpCaMV), the firefly luciferase (*LUC*) ORF, and the nopaline synthase terminator (Fig. 4*A*) and compared their effect on gene expression in maize leaf protoplasts (Fig. 4*B*). The results showed that the maize construct containing the Harbinger-like element significantly repressed luciferase activity relative to the teosinte construct without the Harbinger-like element and the empty construct (Fig. 4*B*). However, no significant differences between the teosinte construct and the empty construct were detected (Fig. 4*B*). These results suggest that the Harbinger-like element represses *LUC* expression and, by inference, *ZmCCT9* expression in vivo.

The importance of *qDTA9* has been observed consistently in previous flowering-time mapping studies (13, 22, 27, 30, 38). *qDTA9* was one of the largest-effect flowering-time QTLs detected under LD conditions in a maize nested association mapping (NAM) population consisting of 5,000 RILs derived from crossing 25 diverse inbred lines to a common parent, B73 (13). We found that the Harbinger-like genotypes across the 26 NAM founders were highly consistent with the QTL segregation patterns across the NAM population (Fig. 4*C*). The effect of *qDTA9* was detected mainly in tropical NAM founders that lack the Harbinger-like element, and these lines consistently exhibited delayed flowering relative to the reference temperate line B73 harboring the Harbinger-like element (Fig. 4*C*). Conversely, NAM founders carrying the Harbinger-like element tended to have no effect at *qDTA9* (Fig. 4*C*).

Cis-acting variants typically affect the transcription of their target genes in an allele-specific manner (6, 23, 39, 40). We measured allele-specific *ZmCCT9* transcript abundance in mature leaves of F_1 crosses between the 25 NAM founders and the reference B73. The results showed that the NAM founders lacking the Harbinger-like element, which are mostly tropical lines, consistently exhibited higher expression of *ZmCCT9* than

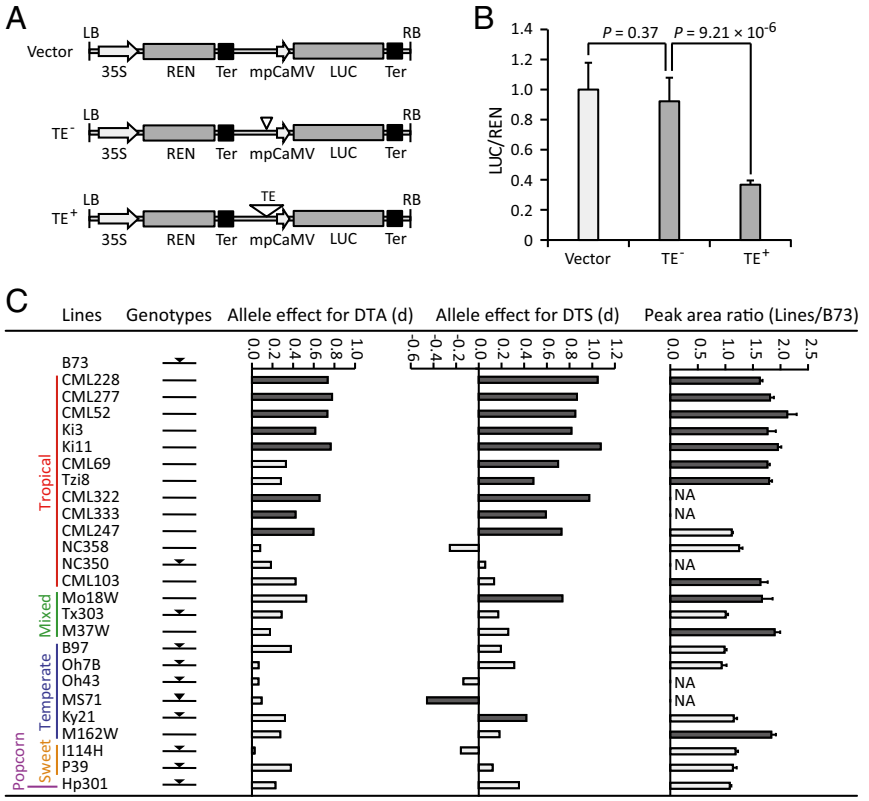

Fig. 4. The Harbinger-like element acts in *cis* to regulate *ZmCCT9* expression. (*A*) Constructs used to test the effect of the Harbinger-like element on gene expression in transient expression assays in maize leaf protoplasts. TE^+ represents the construct with the maize sequence containing the Harbinger-like element cloned upstream of the mpCaMV; TE^- represents the construct with the teosinte sequence lacking the Harbinger-like element cloned upstream of the mpCaMV. The maize and teosinte sequences used in the transient expression assay are 484 bp and 124 bp long, respectively, and they differ only at the Harbinger-like element. (*B*) The maize construct containing the Harbinger-like element significantly repressed luciferase activity relative to the teosinte construct without the Harbinger-like element and the empty construct. The data are normalized with respect to the average values of the empty construct and are shown as means ± SEM; *P* values were determined by Student's *t* test. (*C*) Congruence between the Harbinger-like genotypes and the patterns of QTL segregation and relative *ZmCCT9* expression across NAM founders. The black inverted triangle indicates that the corresponding NAM founder contains the Harbinger-like element. The allele effects for DTA and days to silking (DTS) at *qDTA9* across the 25 NAM founders were obtained from the previous NAM flowering-time mapping study (13). The last column indicates allele-specific *ZmCCT9* transcript abundance in mature leaves of F_1 crosses between the 25 NAM founders and the reference B73. The dark-gray bars indicate significant differences ($P < 0.05$), and the light-gray bars indicate no significant differences.

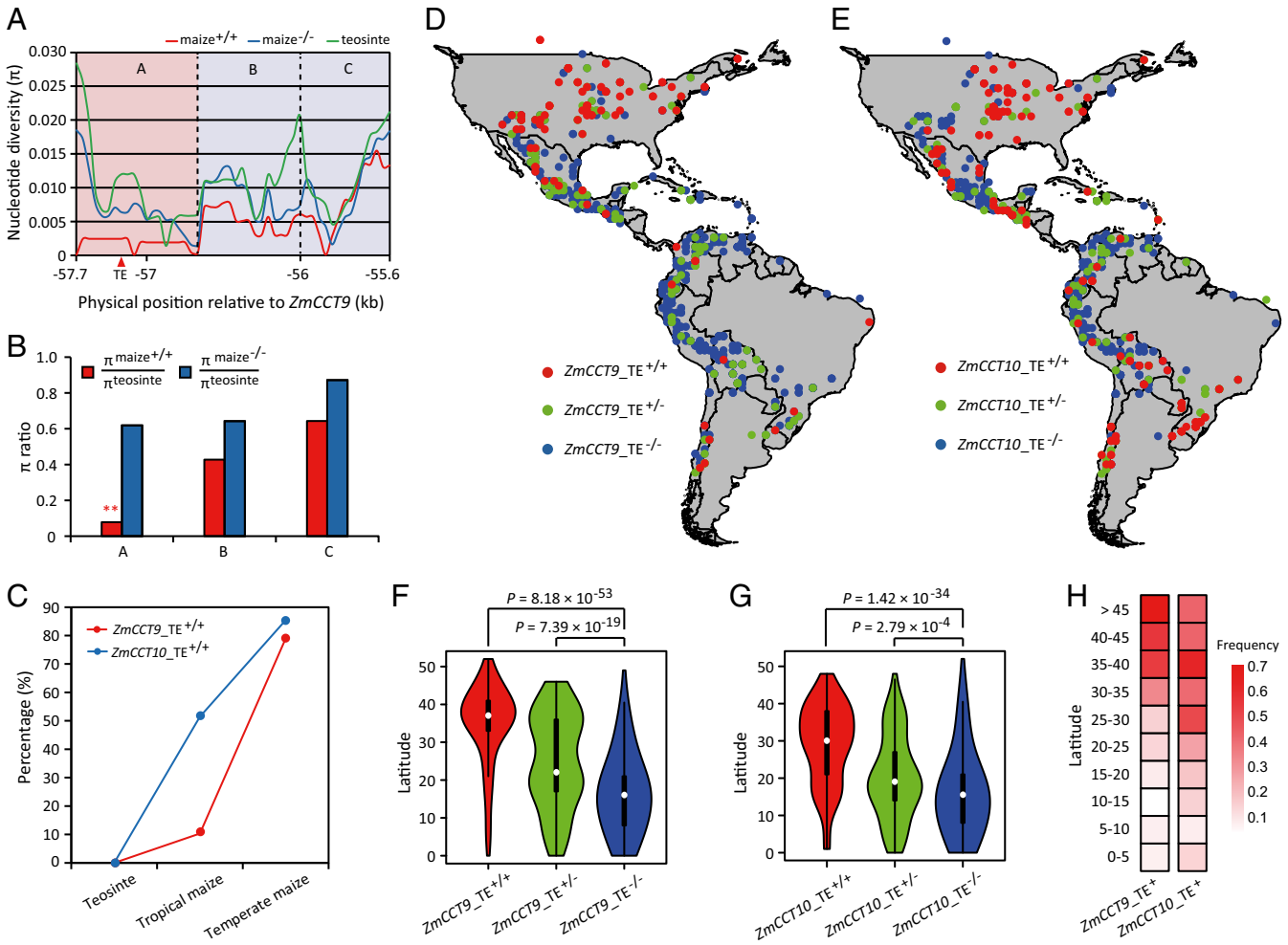

Fig. 5. Evolutionary features of the Harbinger-like transposon insertion at *ZmCCT9*. (*A*) Nucleotide diversity across the 2.4-kb causative region for *qDTA9*. The red, blue, and green lines indicate nucleotide diversity (π) of maize lines carrying the Harbinger-like element (maize$^{+/+}$), maize lines lacking the Harbinger-like element (maize$^{-/-}$), and teosinte accessions, respectively. Base-pair positions are relative to the start codon of *ZmCCT9*. The red triangle indicates the position of the Harbinger-like element. (*B*) The Harbinger-like element shows evidence of selection. In the A region flanking the Harbinger-like element, maize lines carrying the Harbinger-like element retained only 7.8% of the nucleotide diversity in teosinte, a finding that cannot be explained by the neutral domestication bottleneck alone. The double asterisks represent significance difference determined by the coalescent simulations at $P < 0.01$. (*C*) Allele-frequency distribution of the Harbinger-like transposon insertion at *ZmCCT9* and the CACTA-like transposon insertion at *ZmCCT10* in teosinte, temperate maize inbred lines, and tropical maize inbred lines. (*D* and *E*) Geographic distribution of the Harbinger-like transposon insertion at *ZmCCT9* (*D*) and the CACTA-like transposon insertion at *ZmCCT10* (*E*) in 1,008 maize landraces native to the Americas. (*F* and *G*) The Harbinger-like transposon insertion at *ZmCCT9* (*F*) and the CACTA-like transposon insertion at *ZmCCT10* (*G*) predominantly accumulate at higher latitudes. In the violin plots, a violin shape indicates the kernel-density curve, a white node in the center indicates the median, and a black box inside the violin indicates a box-and-whisker plot. *P* values were determined by Student's *t* test. (*H*) Frequency distributions of the Harbinger-like transposon insertion at *ZmCCT9* and the CACTA-like transposon insertion at *ZmCCT10* along latitudinal gradients.

the B73 allele under LD conditions (Fig. 4*C*), which highly corresponds with the QTL segregation patterns across the NAM founders. Taken together, these data further demonstrate that the Harbinger-like element acts in *cis* to repress *ZmCCT9* expression to promote flowering under LD conditions.

***ZmCCT9* Functions Upstream of *ZCN8* to Regulate Flowering Time.** To determine the regulatory relationship of *ZmCCT9* in the maize photoperiod pathway, we examined the expression patterns of several known photoperiod genes (*Gigz1a*, *ZCN8*, *ZmCCT10*, and *CONZ1*) (6, 7, 26, 37, 41, 42) in mature leaves of NIL(maize) and NIL(teosinte) under LD conditions. Interestingly, of these genes, only *ZCN8*, which encodes the maize florigen and functions as a floral activator (37, 42), exhibited significant difference in expression in the two NILs; NIL(maize) showed higher expression of *ZCN8* than NIL(teosinte), indicating that *ZCN8* is negatively regulated by *ZmCCT9* under LD conditions (*SI Appendix*, Fig. S9). Consistent with this result, *ZCN8* expression was significantly elevated in CRISPR/Cas9-engineered null mutants compared with wild-type plants under LD conditions (*SI Appendix*, Fig. S10). Analysis of diurnal expression of these photoperiod-related genes further showed that only *ZCN8* exhibited differential diurnal expression in NIL(maize) and NIL(teosinte) under LD conditions (*SI Appendix*, Fig. S11*A*). However, *ZCN8* showed similar diurnal expression patterns in NIL(maize) and NIL(teosinte) under SD conditions (*SI Appendix*, Fig. S11*B*). Taken together, these results suggest that *ZmCCT9* functions as a critical LD suppressor connecting *ZCN8* expression and the circadian network.

The Harbinger-Like Element Arose After Initial Domestication and Was Targeted by Selection. To examine the evolutionary origin of the Harbinger-like element, we genotyped the Harbinger-like locus in 73 diverse teosinte accessions (*Z. mays* ssp. *parviglumis*) (*SI Appendix*, Table S1). Interestingly, none of the investigated

teosintes carried the Harbinger-like element, indicating that the Harbinger-like transposon insertion is most likely a de novo mutation that occurred after initial domestication. To determine whether selection has acted on the Harbinger-like element, we analyzed the nucleotide diversity across the 2.4-kb region in a panel of 27 diverse maize inbred lines and 19 teosinte (*Z. mays* ssp. *parviglumis*) accessions (Fig. 5*A* and *SI Appendix*, Table S2). Interestingly, in the A region flanking the Harbinger-like element, maize lines carrying the Harbinger-like element retained only 7.8% of the nucleotide diversity of teosinte (Fig. 5*B*). A coalescence simulation incorporating the maize domestication bottleneck (43–45) showed that this severe loss of genetic diversity cannot be explained by the maize domestication bottleneck alone, indicating a strong past selection near the Harbinger-like element. No significant selection signals were detected at regions B and C (Fig. 5*B*). Phylogenetic analysis of the A region clearly separated lines carrying the Harbinger-like element from lines lacking it, such that all teosintes cluster together with maize lines lacking the Harbinger-like element (*SI Appendix*, Fig. S12).

The Two Transposon Insertions at *ZmCCT9* and *ZmCCT10* Arose Sequentially Following Domestication and Exhibited Strong Latitudinal Adaptation. We previously cloned *ZmCCT10*, which controls the largest-effect flowering-time QTL, *qDTA10*, in the same maize-teosinte BC_2S_3 population and the maize NAM population (6). A CACTA-like transposon insertion at the promoter of *ZmCCT10* was found to be the causative variant (7). Interestingly, both the Harbinger-like transposon insertion at *ZmCCT9* and the CACTA-like transposon insertion at *ZmCCT10* are de novo mutations that occurred after initial domestication and show evidence of selection. To reveal how these two transposable elements contributed to maize flowering-time adaptation, we analyzed their allele distribution in the 513 modern maize inbred lines that were used in the association mapping. We found that 51.4% of tropical inbred lines carried the CACTA-like transposon at *ZmCCT10* and that this percentage further increased to 85.4% in temperate inbred lines (Fig. 5*C*). However, the Harbinger-like transposon insertion at *ZmCCT9* remained at a low frequency (10.3%) in tropical inbred lines but increased to 79% in temperate inbred lines (Fig. 5*C*). These contrasting frequency distributions suggest that these two transposons might be associated with distinct patterns of geographic dispersal. To test this hypothesis, we investigated the distribution of these two transposons among 1,008 maize landrace accessions (*SI Appendix*, Table S3) representing the entire pre-Columbian range of maize races native to the Americas (46, 47). As expected, both transposons showed strong associations with latitude, with both predominantly accumulating at higher latitudes (Fig. 5 *D*–*H*). However, closer inspection showed that the two transposons exhibit different frequency distributions along latitudes (Fig. 5*H*). Compared with the Harbinger-like transposon insertion at *ZmCCT9*, which occurs at low frequency at low latitudes, the CACTA-like transposon insertion at *ZmCCT10* occurs at a relatively higher frequency at low latitudes (Fig. 5*H*), suggesting that the CACTA-like transposon insertion at *ZmCCT10* arose earlier than the Harbinger-like transposon insertion at *ZmCCT9*. To verify this, we performed molecular dating using previously described methods (48). The results showed that the CACTA-like transposon insertion and the Harbinger-like transposon insertion arose ∼7,269 and ∼4,645 y B.P., respectively. These estimates suggest that both transposon insertions arose postdomestication and that the Harbinger-like transposon insertion at *ZmCCT9* arose ∼2,624 y later than the CACTA-like transposon insertion at *ZmCCT10*, thereby leading to the differences in the geographic range.

Conclusions

In summary, this work uncovered a distant Harbinger-like transposon acting in *cis* to alter *ZmCCT9* expression to regulate maize flowering time. *ZmCCT9* confers LD-dependent flowering repression by negatively regulating the expression of the florigen *ZCN8*. The two transposon insertions at *ZmCCT10* and *ZmCCT9* arose sequentially after the initial domestication of maize and were targeted by selection as maize spread from its tropical origin to higher latitudes. Additionally, a transposable element insertion at *Vgt1*, a third locus involved in maize adaption to higher latitudes, has also been identified as a candidate causal variant (23–25). These findings further demonstrate the important roles of transposable elements in creating functional variation and driving species evolution (7, 23, 48–53), as inferred by Barbara McClintock decades ago (21). Moreover, our results help explain how the dynamic maize genome with abundant transposon activity enabled maize to "break the rule" that latitude constrains crop diffusion and was spread by its Native American inventors over 90° of latitude from Chile to Canada during the pre-Columbian era. The selection of natural regulatory variation at paralogs with similar biological functions might represent an efficient and prevalent strategy for local adaptation of crops.

Materials and Methods

Plant materials are described in *SI Appendix, SI Materials and Methods*. Details of the methods and experimental procedures are provided in *SI Appendix, SI Materials and Methods*, including fine mapping, association mapping, transgenic analysis, RNA extraction, quantitative real-time PCR analysis, RACE, subcellular localization, RNA in situ hybridization, protoplast transient expression assay, allele-specific expression assay, nucleotide diversity analysis, selection test, molecular dating analysis, and phylogenetic analysis.

ACKNOWLEDGMENTS. We thank Jigang Li (State Key Laboratory of Plant Physiology and Biochemistry, China Agricultural University) and Lubin Tan (Department of Plant Genetics and Breeding, China Agricultural University) for helpful discussions and advice on molecular experiments. This work was supported by grants from the National Key Research and Development Program of China (2016YFD0100404 and 2016YFD0100303); the National Natural Science Foundation of China (31771806 and 91535108); and the Recruitment Program of Global Experts and the Fundamental Research Funds for the Central Universities.

1. Yan L, et al. (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640–1644.
2. Turner A, Beales J, Faure S, Dunford RP, Laurie DA (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310:1031–1034.
3. Xue W, et al. (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet* 40:761–767.
4. Murphy RL, et al. (2011) Coincident light and clock regulation of *pseudoresponse regulator protein 37 (PRR37)* controls photoperiodic flowering in sorghum. *Proc Natl Acad Sci USA* 108:16469–16474.
5. Comadran J, et al. (2012) Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat Genet* 44:1388–1392.
6. Hung HY, et al. (2012) *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc Natl Acad Sci USA* 109:E1913–E1921.
7. Yang Q, et al. (2013) CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc Natl Acad Sci USA* 110:16969–16974.
8. Park SJ, et al. (2014) Optimization of crop productivity in tomato using induced mutations in the florigen pathway. *Nat Genet* 46:1337–1342.
9. Soyk S, et al. (2017) Variation in the flowering gene *SELF PRUNING 5G* promotes day-neutrality and early yield in tomato. *Nat Genet* 49:162–168.
10. Lu S, et al. (2017) Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. *Nat Genet* 49:773–779.
11. Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080–6084.
12. Kuleshov NN (1933) World's diversity of phenotypes of maize. *Agron J* 25:688–700.
13. Buckler ES, et al. (2009) The genetic architecture of maize flowering time. *Science* 325:714–718.
14. Swarts K, et al. (2017) Genomic estimation of complex traits reveals ancient maize adaptation to temperate North America. *Science* 357:512–515.
15. Diamond J (1997) *Guns, Germs, and Steel: The Fates of Human Societies* (Norton, New York).
16. Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature* 418:700–707.
17. McClintock B (1950) The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* 36:344–355.

18. Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8:272–285.
19. Schnable PS, et al. (2009) The B73 maize genome: Complexity, diversity, and dynamics. *Science* 326:1112–1115.
20. Jiao Y, et al. (2017) Improved maize reference genome with single-molecule technologies. *Nature* 546:524–527.
21. McClintock B (1984) The significance of responses of the genome to challenge. *Science* 226:792–801.
22. Chardon F, et al. (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. *Genetics* 168:2169–2185.
23. Salvi S, et al. (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc Natl Acad Sci USA* 104: 11376–11381.
24. Ducrocq S, et al. (2008) Key impact of *Vgt1* on flowering time adaptation in maize: Evidence from association mapping and ecogeographical information. *Genetics* 178: 2433–2437.
25. Castelletti S, Tuberosa R, Pindo M, Salvi S (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL *Vgt1*. *G3 (Bethesda)* 4:805–812.
26. Ducrocq S, et al. (2009) Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. *Genetics* 183: 1555–1563.
27. Li D, et al. (2016) The genetic architecture of leaf number and its genetic relationship to flowering time in maize. *New Phytol* 210:256–268.
28. Xu G, et al. (2017) Complex genetic architecture underlies maize tassel domestication. *New Phytol* 214:852–864.
29. Yu J, et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203–208.
30. Coles ND, McMullen MD, Balint-Kurti PJ, Pratt RC, Holland JB (2010) Genetic control of photoperiod sensitivity in maize revealed by joint multiple population analysis. *Genetics* 184:799–812.
31. Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* 346:1258096.
32. Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V (2015) Editing plant genomes with CRISPR/Cas9. *Curr Opin Biotechnol* 32:76–84.
33. Suárez-López P, et al. (2001) *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature* 410:1116–1120.
34. Kojima S, et al. (2002) *Hd3a*, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol* 43:1096–1105.
35. Valverde F, et al. (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303:1003–1006.
36. Kobayashi Y, Weigel D (2007) Move on up, it's time for change: Mobile signals controlling photoperiod-dependent flowering. *Genes Dev* 21:2371–2384.
37. Meng X, Muszynski MG, Danilevskaya ON (2011) The *FT*-like *ZCN8* gene functions as a floral activator and is involved in photoperiod sensitivity in maize. *Plant Cell* 23: 942–960.
38. Li YX, et al. (2016) Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. *Plant J* 86: 391–402.
39. Clark RM, Wagler TN, Quijada P, Doebley J (2006) A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. *Nat Genet* 38:594–597.
40. Springer NM, Stupar RM (2007) Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. *Plant Cell* 19:2391–2402.
41. Miller TA, Muslin EH, Dorweiler JE (2008) A maize *CONSTANS*-like gene, *conz1*, exhibits distinct diurnal expression patterns in varied photoperiods. *Planta* 227: 1377–1388.
42. Lazakis CM, Coneva V, Colasanti J (2011) *ZCN8* encodes a potential orthologue of *Arabidopsis* FT florigen that integrates both endogenous and photoperiod flowering signals in maize. *J Exp Bot* 62:4833–4842.
43. Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. *Proc Natl Acad Sci USA* 95: 4441–4446.
44. Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: A multilocus investigation of the domestication process in maize. *Mol Biol Evol* 21: 1214–1225.
45. Wright SI, et al. (2005) The effects of artificial selection on the maize genome. *Science* 308:1310–1314.
46. Vigouroux Y, et al. (2008) Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *Am J Bot* 95:1240–1253.
47. van Heerwaarden J, et al. (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc Natl Acad Sci USA* 108:1088–1092.
48. Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43:1160–1163.
49. Bejerano G, et al. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 441:87–90.
50. Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* 319:1527–1530.
51. Naito K, et al. (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* 461:1130–1134.
52. González J, Karasov TL, Messer PW, Petrov DA (2010) Genome-wide patterns of adaptation to temperate environments associated with transposable elements in Drosophila. *PLoS Genet* 6:e1000905.
53. Gray MM, Sutter NB, Ostrander EA, Wayne RK (2010) The *IGF1* small dog haplotype is derived from Middle Eastern grey wolves. *BMC Biol* 8:16.