GmMDE genes bridge the maturity gene E1 and florigens in photoperiodic regulation of flowering in soybean

Hong Zhai,1,2 Zhao Wan,1,2 Shuang Jiao,3 Jingwen Zhou,1 Kun Xu,1 Haiyang Nan,1 Yingxiang Liu,3 Shangshang Xiong,1,2 Rong Fan,3 Jinlong Zhu,1 Wenting Jiang,3 Tian Pang,3 Xiao Luo,4 Hongyan Wu,1 Guang Yang,5 Xi Bai,3 Fanjiang Kong6 and Zhengjun Xia1,*†

1 Key Laboratory of Soybean Molecular Design Breeding (2013DP173244), Northeast Institute of Geography and Agroecology, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Harbin 150081, China
2 College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
3 College of Life Science, Northeast Agricultural University, Harbin 150030, China
4 Institute of Advanced Agricultural Sciences, Peking University, Beijing, China
5 Cultivation and Crop Tillage Institute, Heilongjiang Academy of Agricultural Sciences, Harbin, China
6 Innovative Center of Molecular Genetics and Evolution, School of Life Sciences, Guangzhou University, Guangzhou, China

*Author for correspondence: xiazhj@neigaehrb.ac.cn (Z.X.).
†Senior author.

H. Z. designed and performed most of experiments, analyzed the data, wrote the manuscript. Z.W., S.J, J.W.Z., K.X., Y.X.L., R.F., J.L.Z., W.T.J., T.P., X.L., H.Y.W., and G.Y. performed partial experiments. F.J.K. and H.Y.N. contributed transgenic soybean overexpressing GmFT2a/GmFT5a. X.B. and T.P. generated transgenic soybean overexpressing GmMDE06. Z.J.X. revised and approved the final manuscript. The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/general-instructions) is: Hong Zhai (zhaih@iga.ac.cn).

Abstract

Soybean (Glycine max) is highly sensitive to photoperiod, which affects flowering time and plant architecture and thus limits the distribution range of elite soybean cultivars. The major maturity gene E1 confers the most prominent effect on photoperiod sensitivity, but its downstream signaling pathway remains largely unknown. Here, we confirm that the encoded E1 protein is a transcriptional repressor. The expression of seven GmMDE genes (Glycine max MADS-box genes downregulated by E1) was suppressed when E1 was overexpressed and promoted when E1 was knocked out through clustered regularly-interspaced short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9)-mediated mutagenesis. These GmMDEs exhibited similar tissue specificity and expression patterns, including in response to photoperiod, E1 expression, and E1 genotype. E1 repressed GmMDE promoter activity. Results for two GmMDEs showed that E1 epigenetically silences their expression by directly binding to their promoters to increase H3K27me3 levels. The overexpression of GmMDE06 promoted flowering and post-flowering termination of stem growth. The late flowering phenotype of E1-overexpressing soybean lines was reversed by the overexpression of GmMDE06, placing GmMDE06 downstream of E1. The overexpression of GmMDE06 increased the expression of the soybean FLOWERING LOCUS T orthologs GmFT2a and GmFT5a, leading to feedback upregulation of GmMDE, indicating that GmMDE and GmFT2a/GmFT5a form a positive regulatory feedback loop promoting flowering. GmMDE06 also promoted post-flowering termination of stem growth by repressing the expression of the shoot identity gene Dt1. The E1-GmMDEs-GmFT2a/GmFT5a-Dt1 signaling pathway illustrates how soybean responds to photoperiod by modulating flowering time and post-flowering stem termination.

https://doi.org/10.1093/plphys/kiac092
PLANT PHYSIOLOGY 2022: 189: 1021–1036
Introduction

Soybean (Glycine max) is a typical short-day plant that is highly sensitive to photoperiod. Most soybean cultivars will flower when daylength passes below a certain threshold but will remain in a vegetative state when daylength is longer than this threshold. Soybean plant architecture, particularly plant height, is also sensitive to photoperiod, which is of particular importance for field performance and yield. Due to their photoperiod sensitivity, each soybean cultivar is adapted to a narrow range of latitudes for maximal yield, thus limiting a wider distribution of elite soybean cultivars.

Photoperiodism in soybean was first described in 1920. To date, 12 flowering time or maturity loci (E1–E11 and the J locus for “long juvenile period”) have been identified by classical genetic methods (Bernard, 1971; Buzzell, 1971; Mcblain and Bernard, 1987; Ray et al., 1994; Tsubokura et al., 2014). Additionally, E1 is a legume-specific gene (Xia et al., 2012), raising the possibility that information gathered in soybean might reveal a mechanism different from that of the model plants Arabidopsis and rice (Oryza sativa). The characterization of the E1-mediated pathway might lead to a breakthrough in understanding photoperiodism in soybean.

In our previous study, we cloned the E1 locus through map-based cloning and demonstrated that E1 and its homologs from common bean (Phaseolus vulgaris) and barrel clover (Medicago truncatula) can regulate flowering time and plant height (Zhang et al., 2016b). In this study, we extended our previous study and uncovered a hitherto unknown function of the legume-specific gene E1 as a transcriptional repressor in epigenetic silencing. We identified seven MADS-box genes (designated GmMDEs) as the downstream regulators of the E1 pathway. GmMDEs regulated flowering time and stem growth habit by affecting the expression levels of GmFT2a and GmFT5a and the stem growth habit gene Dt1. GmMDEs therefore link the photoperiodic E1 pathway and the Dt1 pathway modulating growth habit. Taken together, we report the molecular dissection of the E1-mediated mechanism of photoperiod sensitivity.

Results

E1 displays transcriptional repression ability

E1 contains a putative bipartite nuclear localization signal and a region distantly related to the B3 domain, sharing weak (21%–27%) amino acid sequence identity with other B3 domains (Xia et al., 2012). The B3 domain has been identified in several transcription factors, prompting us to investigate the transcriptional regulatory activity of E1 using a dual-luciferase reporter (DLR) assay in Arabidopsis protoplasts using a firefly luciferase (LUC) reporter gene under the control of a minimal promoter with the yeast GAL4 regulatory sequences. As shown in Figure 1, protoplasts transiently transfected with GAL4 DNA-binding domain (GAL4BD)-E1
showed attenuated LUC activity compared with that transfected with the GAL4BD. Compared with the GAL4BD-VP16 positive control, GAL4BD-VP16-E1 showed inhibitory effects on VP16-promoted gene expression. These results indicated that E1 had transcriptional repression activity.

**E1 represses the expression of MADS box genes**

To identify E1-regulated genes in flowering time regulation, we compared the gene expression profiles of the wild-type soybean cultivar Kariyutaka and E1-overexpressing transgenic soybean by transcriptome deep sequencing (RNA-seq) analysis. A total of 1,112 genes were downregulated in response to E1 overexpression (Supplemental Data Set 1). Notably, these repressed genes included a set of MADS-box genes (Supplemental Data Set 1), of which we selected seven for further analysis. We designated these MADS-box genes GmMDE for Glycine max MADS-box genes downregulated by E1: GmMDE04 (Glyma.04G159300), GmMDE05 (Glyma.05G018800), GmMDE06 (Glyma.06G205800), GmMDE08 (Glyma.08G250800), GmMDE13 (Glyma.13G052100), GmMDE15 (Glyma.15G060800), and GmMDE16 (Glyma.16G072700).
DE17 (Glyma.17G081200), and GmMDE19 (Glyma.19G034600). We validated the repression of the GmMDE transcripts by E1 overexpression using reverse transcription-quantitative polymerase chain reaction (RT-qPCR) in the wild-type cultivar Kariyutaka and the two E1-overexpressing lines #L1 and #L2 (Figure 2A).

To investigate the regulation of GmMDE genes by E1 in more detail, we used CRISPR/Cas9-mediated mutagenesis to inactivate E1 in the soybean cultivar Tianlong1, which carries a functional E1 allele. Accordingly, we designed two single-guide RNAs (sgRNAs) targeting the E1 coding region (Supplemental Figure S1A). We transformed the resulting sgRNA/Cas9 vector into the soybean cultivar Tianlong1. We obtained one e1 mutant line, named e1-1. The e1-1 mutant line harbored a 243-bp deletion in the E1 coding region between the two sgRNA target sites (Supplemental Figure S1A). The e1-1 mutant flowered significantly earlier than its wild-type parental cultivar Tianlong1 (Supplemental Figures S1, B and C). All seven GmMDE genes were more highly expressed in the e1-1 mutant compared to Tianlong1 (Figure 2B). These results were consistent with E1 repressing GmMDE expression.

For an independent confirmation, we determined GmMDE transcript levels in the E1 near-isogenic lines Harosoy-E1 and Harosoy-e1. Harosoy-E1 carries the functional E1 allele, while Harosoy-e1 carries e1-as, a leaky allele retaining partial E1 function (Xia et al., 2012). Again, the transcript levels for all seven GmMDE genes were lower in Harosoy-E1 relative to Harosoy-e1, supporting the repression of GmMDE expression by E1 (Figure 2C). We also noticed that GmMDE transcript levels increased as a function of plant age (Figure 2C).

A phylogenetic analysis of all soybean MADS-box genes was previously reported (Shu et al., 2013), but did not include GmMDE06, GmMDE13, or GmMDE19. We therefore repeated the phylogenetic analysis with all seven GmMDE represented, along with members of type II MADS-box proteins from soybean and Arabidopsis. The seven GmMDEs clustered within the AP1/FUL/CAL clade, whose members specify the identity of inflorescence and floral meristems (Supplemental Figure S2; Mandel et al., 1992; Bowman et al., 1993; Gu et al., 1998).

GmMDEs respond to photoperiod at the transcript level and exhibit similar tissue specificity. To assess whether the seven GmMDEs are involved in the photoperiod pathway, we measured their transcript levels in leaves from different soybean cultivars with contrasting flowering times and photoperiod sensitivity (Figure 3A) under long-day (LD) and short-day (SD) conditions by RT-qPCR. E1 expression levels also varied in these soybean cultivars (Figure 3B). We previously determined the genotype at the E1 locus (Zhai et al., 2014). Broadly speaking, all GmMDEs were highly induced under SD conditions and repressed under LD conditions in most soybean cultivars (Figure 3, C–I), especially in the photoperiod-sensitive cultivars Zhonghuang13 and Harosoy-E1, which harbor an active E1 allele that is highly expressed (Figure 3B). In the photoperiod-insensitive cultivar Kariyutaka, with the e3 e4
genotype (carrying two loss-of-function alleles for E3 and E4) and almost undetectable E1 expression (Xia et al., 2012), GmMDE transcript levels were comparable between SD and LD conditions, indicating that GmMDE expression is not induced by SD in this background. The soybean cultivars Harosoy-e1, Sidou11, and Gaogeng1 express the partially functional e1-as allele to high levels (Zhai et al., 2014). In these three cultivars, GmMDE06 and GmMDE19 expression was induced by SD conditions (Figure 3, E and I), while that of GmMDE04, GmMDE05, GmMDE08, GmMDE13, and GmMDE17 was not always induced by SD conditions (Figure 3, C, D, F–H). Although e1-as might retain partial E1 function, e1-as function is nevertheless greatly weakened compared to the E1 allele (Xia et al., 2012). Thus, E1 repression of GmMDE04, GmMDE05, GmMDE08, GmMDE13, and GmMDE17 was strongly weakened under LD conditions, and the expressions of these five genes were not SD inductive.

We then explored GmMDE transcript profiles in various tissues in the soybean cultivar Kariyutaka, the E1-overexpressing soybean line #L1 (in the Kariyutaka background), the near-isogenic lines Harosoy-E1 and Harosoy-e1, and the soybean cultivar HX3 (late flowering, photoperiod-sensitive, and carrying the J locus) (Supplemental Figure S3). The expression pattern of GmMDEs shared commonalities across tissues. We detected extremely high expression levels in flowers, but with no obvious association with the expression level or functional status of the E1 allele present, suggesting that the repression of GmMDE expression by E1 is not active in flowers. In other plant tissues, GmMDE expression levels were relatively high in Kariyutaka (with very low E1 expression) and were relatively low in soybean cultivars expressing a functional E1 allele to high levels (such as Harosoy-E1 and HX3) and in the E1-overexpressing line. Most GmMDE genes exhibited relatively low expression in petioles and a relatively high expression in unifoliolate leaves, fully expanded trifoliate leaves, unexpanded trifoliate leaves, and the apical meristem. In addition, we noticed subtle differences in terms of tissue specificity of individual GmMDE genes.

**E1 binds to DNA in a nonsequence-specific manner in vitro**

E1 is predicted to contain a B3-like domain, based on its low (21%–27%) amino acid sequence identity with known B3 domains (Xia et al., 2012). Members of the plant-specific B3 superfamily bind with specificity to the motifs 5' - TGTCTC-3', 5' - CATGCA-3', or 5' - CACCTG-3' (Suzuki et al., 1997; Ulmasov et al., 1997; Kagaya et al., 1999). We thus examined the promoter sequences within ~2.5-kb upstream
of seven GmMDE genes and identified the putative B3 domain motif 5'-CATGCA-3' or the similar motif 5'-CATG-3' in all promoters. We conducted electrophoretic mobility shift assays (EMSAs) to investigate whether E1 directly binds to GmMDE promoters. Accordingly, we generated biotin-labeled probes that contain the 5'-CATGCA-3' or 5'-CATGCA-3' motif in the promoter regions of four GmMDEs. Soluble recombinant His-tagged E1 (His-E1) bound to four probes in vitro, as evidenced by the observed shift (Supplemental Figure S4A). However, E1 still formed a complex with mutated probes when the 5'-CATGCA-3' or 5'-CATGCA-3' motif was mutated to 5'-TTATT-3'. Since the remaining probe sequences outside the mutated motifs were quite different, we deemed it unlikely that E1 would specifically bind to other sequences outside its presumptive B3. To confirm the identity of the complex, we performed an EMSA supershift using antibodies against the His tag (Supplemental Figure S4B). Indeed, we observed a distinct supershift under these conditions, indicating that the complex contains E1. Based on the above results, we hypothesized that the binding is not sequence specific. The addition of excess unlabelled probes did not prevent formation of the complex (Supplemental Figure S4C). Similarly, E1 bound to a promoter fragment from GsWRKY20 (Supplemental Figure S4A), a soybean flowering gene whose expression is not regulated by E1. We concluded that E1 binds to DNA in a non-sequence-specific manner.

**E1 directly binds to the GmMDE05 and GmMDE06 promoters to repress their expression by increasing H3K27me3 levels**

DNA binding proteins with no sequence specificity can be targeted to specific genomic regions in vivo (Narlikar et al., 2002). To test whether E1 directly represses GmMDE expression in vivo, we performed transient infiltration assays in Nicotiana benthamiana leaves using reporter constructs consisting of the GmMDE05, GmMDE06, and GmMDE17 promoters driving the transcription of the firefly LUC gene. We detected strong LUC activity when pGmMDE05LUC, pGmMDE06LUC, or pGmMDE17LUC constructs were infiltrated into N. benthamiana leaves (Figure 4, A–C). However, co-infiltration of each LUC reporter with the 35S:E1 construct strongly repressed LUC transcription, suggesting that E1 might interact with GmMDE promoters to prevent transcription. We obtained similar results when the B3 domain recognition motifs were deleted from the GmMDE06 promoter, indicating that the repression of GmMDE transcription by E1 does not rely on canonical B3 domain recognition motifs (Supplemental Figure S5).

Nonsequence-specific DNA binding proteins are typically components of multisubunit complexes involved in the regulation of chromatin state (Simon and Tamkun, 2002; Bratzel et al., 2010), raising the possibility that E1 might modulate histone methylation to regulate GmMDE transcription. We analyzed tri-methylation on lysine27 of histone H3 (H3K27me3) levels at GmMDE promoters, as H3K27me3 leads to the repression of target gene expression.

We selected the GmMDE05 and GmMDE06 promoters for H3K27me3 level analysis. We generated 35S:cMYC-E1 transgenic soybean lines (named WEM) in the Williams82 background. The WEM lines showed a clear late flowering time phenotype (Supplemental Figure S6, A and B) compared to the parental cultivar Williams82, indicating that the cMYC-E1 protein is functional. Immunoblot analysis using anti-cMYC antibody confirmed the accumulation of the cMYC-E1 protein (Supplemental Figure S6C). We then performed chromatin immunoprecipitation qPCR (ChIP-qPCR) with an anti-trimethyl-Histone H3 (Lys27) antibody in both WEM and Williams82. Because the binding sequences of E1 are not clear, we analyzed H3K27me3 levels over most of the GmMDE05 and GmMDE06 promoters (Figure 4D; Supplemental Figure S7A). As shown in Figure 4E, we detected higher H3K27me3 levels in many sites along the GmMDE05 promoter region compared to the IgG negative control, especially near the transcription start sites, including regions 05-P1, 05-P2, 05-P3, and 05-P4, indicating that the GmMDE05 promoter sequence is trimethylated at H3K27. Notably, H3K27me3 levels were significantly higher in these regions in the WEM background than in the wild-type (Figure 4E), indicating that the overexpression of E1 increases H3K27me3 levels at the GmMDE05 promoter. To independently confirm the direct binding of E1 to the GmMDE05 promoter, we performed ChIP-qPCR using an anti-cMYC antibody on the WEM transgenic line and the wild-type. Again, we observed a strong enrichment of E1 binding at two sites, the 05-P3 and 05-P4 regions within the GmMDE05 promoter, specifically in WEM plants but not in the wild-type (Figure 4F). Importantly, these E1 binding regions overlapped with the regions exhibiting elevated H3K27me3 levels. In light of the repression of GmMDE05 expression imposed by E1 (Figure 2), we concluded that E1 may directly bind to the GmMDE05 promoter region to raise the levels of H3K27mes and thus repress GmMDE05 expression.

We repeated the same analysis of H3K27me3 levels and ChIP-qPCR for most of the GmMDE06 promoter region (Supplemental Figure S7A). Compared to the IgG negative control, we observed extremely high H3K27me3 levels in many sites along the GmMDE06 promoter region, such as 06-P1, 06-P2, 06-P3, 06-P4, and 06-P5, indicating that the GmMDE06 promoter is also trimethylated at H3K27 (Supplemental Figure S7B). Moreover, we detected higher H3K27me3 levels in WEM seedlings than in the wild-type near the transcription start site, including regions 06-P1, 06-P2, 06-P3, and 06-P5 (Supplemental Figure S7B), indicating that E1 raises H3K27me3 levels at the GmMDE06 promoter. When conducting ChIP-qPCR, we also observed enrichment of E1 at the 06-P1 region (Supplemental Figure S7C), allowing us to conclude that E1 binds to the 06-P1 region of the GmMDE06 promoter in vivo. However, the rise in H3K27me3 levels upon the overexpression of E1 at the
promoter region was less pronounced at the GmMDE06 promoter (about 1.4- to 1.8-fold) than at the GmMDE05 promoter (about 2.4- to 3.5-fold), suggesting that the regulation of GmMDE06 transcription by E1 might be weaker than that of GmMDE05.

Epigenetic modifications, such as histone methylation and demethylation, acetylation and deacetylation, can affect chromatin accessibility. The tightening or loosening of genomic DNA will indeed prevent or allow transcription factors to access their cognate cis-acting elements in promoters to repress or initiate transcription. In this study, E1 appeared to repress the promoter activity of GmMDE05 using transient expression assays in N. benthamiana leaves. Since E1 epigenetically silences GmMDE transcription, T-DNA could
integrate into the plant genomic DNA in transient expression assays so that E1 could regulate the histone methylation. Previous studies had confirmed that T-DNA could integrate stably into chromosome and be transferred to the next generation in transient expression assay by Agroinfiltration (Kopertekh and Schiemann, 2005)

The phenotypic analysis of transgenic soybean overexpressing GmMDE06 shows that GmMDE06 acts downstream of E1

We focused on GmMDE06 to explore the function of GmMDEs in flowering time. Accordingly, we overexpressed GmMDE06 under the control of the CaMV 35S promoter in the soybean cultivar Dongnong50 and selected three lines (06#L2, 06#L8, and 06#L18) with high GmMDE06 expression, as determined by RT-PCR (Figure 5A). We phenotyped T3 lines under field conditions. When grown in LD conditions, all lines showed early flowering compared to their nontransgenic control parent Dongnong50 (Figure 5, B and E). The transgenic lines also reached reproductive stages R2–R8 faster than Dongnong50 (Figure 5, B and E), indicative of the accelerated maturation of these transgenic plants. Moreover, the transgenic plants exhibited determinate stem growth habits, with reduced plant height and decreased node number along the main stems (Figure 5, C, D, F, and G), in contrast to the semi-determinate stem growth habits of the wild-type Dongnong 50. The dwarf phenotype of the transgenic lines was the result of fewer nodes along the main stems with shorter internodes (Figure 5, G and H).

To define the role of GmMDE06 in E1 signaling in more detail, we crossed the GmMDE06 overexpression line with 3SS:E1 plants, which we showed previously flower late in the cultivar Dongnong50 (Zhang et al., 2016b). We tested all F1 plants by RT-qPCR and selected 32 plants with high expression levels for both E1 and GmMDE06 for phenotypic analysis. As shown in Figure 5I, these F1 plants overexpressing both E1 and GmMDE06 flowered earlier than the wild-type and much earlier than E1-overexpressing plants, demonstrating that the overexpression of GmMDE06 largely, but not fully, suppresses the late flowering phenotype of 3SS:E1. We concluded from these results that GmMDE06 acts downstream of E1.

GmMDE06 regulates the expression of GmFT2a, GmFT5a, GmFT1a, and the downstream shoot identity gene Dt1

We showed previously that E1 represses the expression of the flowering activators GmFT2a and GmFT5a, while promoting the expression of the flowering repressors GmFT1a and GmFT4 (Xia et al., 2012; Zhai et al., 2014; Liu et al., 2018). We hypothesized that the regulation of these FT homologs by E1 may be mediated by GmMDEs. We thus measured the expression levels of GmFT2a, GmFT5a, GmFT1a, and GmFT4 in GmMDE06-overexpressing soybean lines: GmFT2a and GmFT5a were more highly expressed in the leaves of soybean overexpressing GmMDE06 than in wild-type leaves, while GmFT1a expression was much lower; the expression of GmFT4 was slightly lower than in wild-type leaves, but not significantly different statistically (P < 0.05; Figure 5, J and M). These results indicated that GmMDE06 regulates the expression of GmFT2a, GmFT5a, and GmFT1a, thus placing GmMDE06 as an important upstream signaling regulator in flowering time.

The stem growth habit of soybean is specified by the shoot identity gene Dt1 (Liu et al., 2010; Tian et al., 2010). Mutations in Dt1 result in the switch of the shoot apical meristem from the vegetative to reproductive state to initiate terminal flowering and thus produce determinate stems. The overexpression of GmMDE06 in soybean induced the termination of stem growth, which prompted us to determine Dt1 expression levels in GmMDE06-overexpressing and wild-type soybean plants. The overexpression of GmMDE06 was associated with a decrease in Dt1 expression in stem tips relative to the wild-type, in agreement with the phenotypic changes from semi-determinate stems to determinate stems seen in the transgenic lines (Figure 5N).

GmMDE06 and GmFT2a/GmFT5a form a positive regulatory feedback loop to promote flowering

Arabidopsis FT interacts with FD at the vegetative shoot apex to activate MADS-box gene expression and initiate flowering (Abe et al., 2005). Previous studies have shown that E1, GmFT2a/GmFT5a, and GmFDL19 are all transcribed in soybean leaves (Xia et al., 2012; Nan et al., 2014), raising the possibility that GmMDEs are also regulated by the flowering activators GmFT2a and GmFT5a. To test this hypothesis, we analyzed the transcript levels of all GmMDEs in the leaves of transgenic lines individually overexpressing GmFT2a (line #2a-1) or GmFT5a (line #5a-1) under the control of the 35S promoter, which were previously produced by Nan et al. (2014). GmMDE expression increased in the lines overexpressing GmFT2a or GmFT5a compared to the nontransgenic control, indicating that GmMDE expression is regulated by GmFT2a and GmFT5a (Figure 6A).

FT-like proteins interact with FD-like proteins to form FT/FD complexes, which bind to the core ACGT cis-elements located at the promoters of AP1-like genes from the MADS-box gene family (Abe et al., 2005). Since the overexpression of GmFT2a and GmFT5a increased the transcript levels of GmMDEs, we suspected that complexes comprising GmFT2a, GmFT5a, and FD-LIKE19 (GmFDL19) might directly bind to the GmMDE promoters. To test this hypothesis, we examined ~2.5 kb of GmMDE promoter regions upstream of the translation start site, which revealed that all GmMDE promoters contain core 5′-ACGT-3′ cis-elements that are recognized by FD family members (Figure 6B). We identified 8 types of ACGT cis-elements: A/G-box (TACGTG), T box (AACGTT), G/T-box (CACGTT), T/A-box (AACGTA), C/T-
Figure 5 Phenotypic analysis of soybean overexpressing GmMDE06 and expression analysis of GmFTs and Dt1. A, Semi-quantitative RT-PCR analysis of GmMDE06 transcript levels in three GmMDE06-overexpressing lines (06#L2, 06#L8, and 06#L18) and wild-type Dongnong 50. TUA5 was used as an internal control. B, Representative photographs of the indicated genotypes at the indicated number of days after sowing under LD conditions. C, D, Photographs of wild-type (Dongnong 50) and GmMDE06-overexpressing plants, with apical stem termination and shorter stature at the R1 and R8 reproductive stages. E, Days to reproductive stages (R1–R8), mean plant height (F), node number of main stem (G), and internode length (H) of GmMDE06-overexpressing and wild-type plants. Values are shown as means in (E–H), and error bars represent standard deviation (n = 40 plants); *P < 0.05; **P < 0.01; ***P < 0.001, as determined by one-tailed Student’s t test. I, The late flowering phenotype of 3SSE1 is suppressed by GmMDE06 overexpression. 35S:E1 soybean (Dongnong50 background) was crossed with 35S:GmMDE06 (Dongnong50 background). Boxplot is made by GraphPad Prism (n = 10 plants). Center line, median; box limits, 25–75th percentiles; whiskers, min to max; points, outliers. Different lowercase letters indicate a significant difference (P < 0.05), as determined by one-tailed Student’s t test. J–M) Relative expression levels of GmFT2a, GmFT5a, GmFT4, and GmFT1a in fully expanded trifoliate leaves, and of Dt1 in stem tips (N) of soybean overexpressing GmMDE06 under LD conditions. For expression analysis of GmFTs, fully expanded trifoliate leaves were collected 4 h after dawn for RT-qPCR. For expression analysis of Dt1, stem tips were collected 4 h after dawn. The wild-type soybean cultivar Dongnong50 was used as a control. Transcript levels were normalized to TUA5; values represent means of three biological replicates; error bars indicate SD.*P < 0.05; **P < 0.01; ***P < 0.001, as determined by one-tailed Student’s t test.
box (GACGTT), C/G-box (GACGTG), G/A-box (CACGTA), and A/C-box (TACGTC) (Figure 6B). To investigate whether GmMDE promoters are direct targets of the GmFT2a/GmFT5a/GmFDL19 complex, we conducted EMSAs with recombinant purified MBP-GmFDL19 (a fusion protein between maltose binding protein and GmFDL19). We designed probes for each GmMDE promoter centered on their ACGT cis-elements, for a total of nine probes (probe 1 for GmMDE06 and probe 1 for GmMDE13 have the same type of ACGT cis-element: an A/G-box). Out of the nine probes tested, four demonstrated binding by MBP-GmFDL19 (Figure 6, C and E), corresponding to three types of ACGT cis-elements: A/G-box (TACGTG), C/G-box (GACGTG), and A/C-box (TACGTC). Binding was effectively competed with unlabeled probes (Figure 6, F and I), but not with unlabeled mutant probes (in which the NACGTN cis-element was mutated to AAAAAA) (Figure 6, F and I), supporting the notion that MBP-GmFDL19 binds to these promoter regions in a sequence-specific manner. Collectively, we showed that GmMDE06 promotes the expression of GmFT2a and GmFT5a, which feeds back to promote the expression of GmMDE06, indicating that GmMDE06 and GmFT2a/GmFT5a form a positive regulatory feedback loop to promote flowering.

Discussion

The E1 protein contains a putative bipartite nuclear localization signal and a region distantly related to B3 domain (Xia et al., 2012). Genes that are highly homologous to E1 were all from legume (Zhang et al., 2016b), so E1 was identified as a legume-specific gene, but homologs of E1 from model legumes have not been well characterized. Based on our previous result of its nuclear localization (Xia et al., 2012), we further proved that E1 protein exerts the transcriptional suppression activity in vivo (Figure 1), so E1 protein acts in transcriptional regulation. But E1 protein binds to DNA in a nonsequence-specific manner in vitro (Supplemental Figure S4). Nonsequence-specific DNA binding proteins are usually found to be components of many multisubunit complexes involved in chromatin regulation (Simon and Tamkun, 2002;
expression of $E_1$ and related proteins.

In this study, our results showed that $E_1$ represses the expression of GmMDEs (Figure 2). As $E_1$ expression is strongly induced under LD conditions (Xia et al., 2012), GmMDEs were characterized by low expression levels (Figure 3). Conversely, under SD conditions, $E_1$ expression is repressed, thus alleviating the suppression of GmMDEs by $E_1$ and leading to their transcriptional activation. The expression of GmMDEs was therefore induced in SD conditions (Figure 3). In the soybean cultivars Zhonghuang13 and Harosoy-$E_1$ with high expression of the functional $E_1$ allele, GmMDE expression responded to photoperiod (Figure 3, C and I). In contrast, the cultivar Kariyutaka showed almost no $E_1$ expression (Xia et al., 2012), which was associated with the photoperiod-insensitive expression of GmMDEs (Figure 3, C and I). The overexpression of GmMDE06 promoted the expression of GmFT2a and GmFT5a in leaves (Figure 5, J and K), and the overexpression of GmFT2a and GmFT5a formed a positive feedback loop onto the expression of GmMDEs, including GmMDE06 (Figure 6A). Under inductive SD conditions, the positive regulatory feedback loop of GmMDE06-GmFT2a/GmFT5a promotes signal amplification perceived in leaves, which contributes to the initiation of flowering. In addition, the seven GmMDE genes displayed an expression pattern that was reminiscent of that of flowering integrator SOC1 (Lee and Lee, 2010; Hyun et al., 2016), including tissue specificity, sensitivity to photoperiod, and association with plant age, hinting that GmMDEs might also be critical flowering regulators. These results above support an important role for GmMDEs as upstream flowering regulators in signal transduction.

When a soybean cultivar is grown southward, its flowering time, as well as its overall plant architecture, particularly plant height, will change as a function of photoperiod, thus potentially greatly influencing yield. $E_1$ exerts the most pronounced effect on photoperiod sensitivity in soybean, which promoted us to explore whether and how $E_1$ might also modulate plant architecture. In this study, we regulated the expression of seven GmMDEs. The overexpression of GmMDE06 promoted flowering and converted the stem from semi-determinate to determinate by suppressing the expression of the major stem growth habit gene $Dt1$ (Figure 5N). $Dt1$ specifies the indeterminate growth habit, which prevents terminal flowering, thus producing taller plants (Liu et al., 2010; Tian et al., 2010; Liu et al., 2016). Previous studies showed that GmFT2a and GmFT5a play major roles in post-flowering stem growth by suppressing $Dt1$ expression (Takeshima et al., 2019). The overexpression of GmMDE06 induced the expression of GmFT2a and GmFT5a (Figure 5, J and K). Thus, $E_1$ might regulate the expression of $Dt1$ through a GmMDEs-GmFT2a/GmFT5a module to control plant height. We propose a model for soybean responses to photoperiod in Figure 7: $Dt1$ is regulated by $E_1$ through the GmMDEs-GmFT2a/GmFT5a module. In LD conditions, the expression of the flowering repressor $E_1$ is induced, while the GmMDE flowering activators are repressed, leading to the repression of GmFT2a/GmFT5a expression and the alleviation of $Dt1$ inhibition, resulting in late flowering and tall plants. Under SDs, the expression of $E_1$ is very low, while the expression of GmMDEs is high, which induces GmFT2a/GmFT5a expression, suppressing $Dt1$ expression and producing earlier flowering of plants with short stature. The E1-GmMDEs-GmFT2a/GmFT5a-Dt1 pathway thus reveals how soybean regulates flowering time and plant height in response to photoperiod.

Conclusion

$E_1$ directly binds to the promoter of GmMDE genes to repress their expression by increasing H3K27me3 levels. GmMDE and GmFT2a/GmFT5a form a positive regulatory feedback loop to promote flowering and post-flowering termination of stem growth. The E1-GmMDEs-GmFT2a/GmFT5a-Dt1 signaling pathway illustrates how soybean responds to photoperiod by modulating flowering time and post-flowering stem termination.
Materials and methods

Plant materials and growth conditions
Soybean (Glycine max) cultivars Kariyutaka, Zhonghuang 13, two E1 near-isogenic lines Harosoy-E1 and Harosoy-e1, Sidou11 and Gaogeng1, E1-overexpressing transgenic soybean, GmFT2a-overexpressing transgenic soybean and GmFT5a-overexpressing transgenic soybean were used. Kariyutaka is an early flowering photoperiod-insensitive cultivar. Zhonghuang 13, Sidou11, Gaogeng1 are late flowering photoperiod-sensitive cultivars. Harosoy-E1 and Harosoy-e1 are E1 near-isogenic lines. Harosoy-E1 carrying the dominant functional E1 allele, is a late flowering photoperiod sensitive line. Harosoy-e1 carrying the recessive e1 allele, is an early flowering photoperiod sensitive line. Allele e1-as, harboring a 1-bp mutation, is a leaky allele, which may retain partial E1 function. Sidou11 and Gaofeng1 carry the e1-as genotype. E1-overexpressing transgenic soybean plants were generated by Xia et al. (2012). GmFT2a-overexpressing transgenic soybean and GmFT5a-overexpressing transgenic soybean plants were generated by Nan et al. (2014). Plants were grown in an artificial climate chamber under either SD conditions (12-h:12-h light/dark) or LD conditions (16-h:8-h light/dark) at 28°C under a light fluency of 200–300 µmol m⁻² S⁻¹. On the 16th day after emergence, fully expanded trifoliolate leaves were sampled 2 h after dawn from three individual plants for RT-qPCR analysis.

For tissue specificity analysis, soybean cultivars Kariyutaka, transgenic soybean overexpressing E1 (E1 was overexpressed in Kariyutaka), Harosoy-E1, Harosoy-e1, and HX3 grown under LD conditions were used. Unifoliolate leaves, fully expanded trifoliolate leaves, unexpanded trifoliolate leaves, apical meristems, petioles, and stems were sampled at the V2 stage according to Fehr’s system. Flowers were sampled at the R2 stage according to Fehr’s system. All samples were from three individual plants sampled for RT-qPCR analysis.

Transient expression assays in protoplasts
The firefly luciferase (LUC) gene driven by a mini-35S (TATA box) promoter with five copies of the GAL4-binding element was used as reporter construct. The internal control construct consisted of the Renilla (REN) reniformis Luciferase gene driven by the cauliflower mosaic virus (CaMV) 35S promoter. The effector construct GAL4BD was generated by cloning the coding sequence of the GAL4 DNA-binding domain (GAL4BD) into the pRT107 vector by SacI/XbaI digestion (Wei et al., 2009). The positive control (GAL4BD-VP16) was constructed by insertion of the VP16 coding sequence in-frame and downstream of the GAL4BD sequence (Wei et al., 2009).
et al., 2009). The coding sequence of E1 was then cloned in-frame and downstream of GAL4BD and GAL4BD-VP16. The full-length E1 coding sequence was cloned into the EcoRI and Sall sites of GAL4BD to generate the GAL4BD-E1 effector construct; the full-length E1 coding sequence was cloned into EcoRI and KpnI sites of GAL4BD-VP16 to generate GAL4BD-VP16-E1. The resulting effector, reporter, and internal control vectors were co-transfected in appropriate combinations into protoplasts prepared from 2-week-old Arabidopsis seedlings. Protoplast isolation and polyethylene glycol (PEG)-mediated transformation were performed as described previously (Yoo et al., 2007). Firefly LUC and REN activity were measured with a DLR assay kit using a GloMax 20/20 luminometer (Promega). LUC activity was normalized to REN activity (LUC/REN). For each plasmid combination, four independent transformations were performed. Values are reported as mean ± SD of four biological replicates (n = 4 transformations). One-tailed, two-sample t tests were used to generate the P-values. The experiment was repeated twice.

RNA-seq assay
E1-overexpressing line #L1 and the wild-type cultivar “Kariyutaka” were used for RNA-seq. Fully expanded trifoliate leaves were collected 4 h after dawn from 30-d-old seedlings grown under LD conditions. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s instructions. The sequencing cDNA libraries were generated according to NEBNext Ultra Directional RNA Library Prep Kit protocol. cDNA libraries were then constructed for sequencing on the Illumina HiSeq TM2500/MiSeq system with 150-bp paired-end read lengths. Tophat2 was used to map clean reads to the soybean reference genome from Phytozome (https://phytozome.jgi.doe.gov/pz/). Differentially expressed genes (DEGs) were detected with DESeq with the following parameters: fold change 2.00 and adjusted P-value (Q-value) 0.05. DEGs are listed in Supplemental Data Set 1. Sequence data of RNA-seq from this article can be found in the National Genomics Data Center (SRA) database under accession number PRJCA005287.

RT-qPCR and semi-quantitative RT-PCR analysis
Total RNA extraction, RT-qPCR, and semi-quantitative RT-PCR were performed as described previously (Zhai et al., 2014). Raw data were standardized as described previously (Willems et al., 2008). Primers used for RT-qPCR and semi-quantitative RT-PCR analysis are listed in Supplemental Table S1. Three fully independent biological replicates (n = 3 plants) were obtained and subjected to RT-qPCR in technical triplicates.

Phylogenetic analysis of MADS-box genes
All candidate MADS-box genes used for phylogenetic analysis were based on Shu et al. (2013). The protein sequences were obtained from the National Center for Biotechnological Information or Phytozome and were aligned using Clustal X2 (protein weight matrix using the Gonnet Series with a gap penalty of 10.00, a gap length penalty of 0.20, and a delay-divergent cutoff of 30%). The phylogenetic analysis was performed using MEGA (V11.0) with the neighbor-joining method, with the following parameters: Poisson correction, pairwise deletion, and bootstrap (1,000 replicates). Phylogenetic tree was colored using ChiPlot (https://xiaochi.chifei3d.com/static/xiaochiPlot/src/index.html).

Generation of CRISPR/Cas9-mediated mutagenesis of E1 gene
The CRISPR/Cas9 target sequences with a G as the first base were designed using the web tool CRISPR-P (http://crispr.hzau.edu.cn/CRISPR2/). Two target sites for E1 designated E1-SP1 and E1-SP2 were integrated into two sgRNA expression cassettes (each sgRNA being driven by the A. thaliana U6 promoter) and cloned into the WMC012 vector, harboring the Bar gene driven by CaMV 35S promoter as selection marker. The CRISPR/Cas9 expression vector was transformed into Agrobacterium (Agrobacterium tumefaciens) strain EHA105 by electroporation. Soybean cultivar “Tianlong1” was used for soybean transformation using the cotyledonal node transformation system as previously described (Flores et al., 2008). To identify mutations, a PCR amplicon covering the sgRNA target site was generated from genomic DNA with E1-specific primers (Supplemental Table S1) and sequenced.

Generation and phenotypic analysis of transgenic soybean plants
To generate overexpression constructs, the coding sequence of GmMDE06 was PCR amplified using cDNA from Kariyutaka leaves with the primer pair listed in Supplemental Table S1, which contained an Xhol or SpeI site (shown in bold font, respectively). The PCR product was then inserted downstream of the CaMV 35S promoter in the pBAM-MYC vector (harboring a CaMV 35S promoter and the Nos terminator). The resulting construct was then introduced into Agrobacterium strain EHA101 and transformed into the “Dongnong 50” cultivar using the cotyledonal node transformation system as described previously (Flores et al., 2008).

For 35S:MYC-E1 transgenic soybean, the E1 coding sequence was PCR amplified using the primer pair listed in Supplemental Table S1. The PCR product was recombined into pEarleyGate203 using LR Clonase II enzyme mix (Invitrogen, USA). The resulting vector was then inserted downstream of the CaMV 35S promoter in the pBAM-MYC vector (harboring a CaMV 35S promoter and the Nos terminator). The resulting construct was then introduced into Agrobacterium strain EHA101 and transformed into the soybean cultivar “Williams82” using the cotyledonal node transformation system (Flores et al., 2008).

T0, T1, and T2 plants were selected by smearing the leaves with 160 mg L−1 glufosinate. Glufosinate-resistant plants were subjected to immunoblot analysis. T3 plants were used for ChIP assays. Soybean plants were grown in pots under natural sunlight (LD conditions) in the glass greenhouse of...
the Northeast Institute of Geography and Agroecology, Harbin, Heilongjiang (45°70′ N, 126°64′ E). The reproductive stages, R1–R8 of soybean, were defined according to Fehr’s system (Fehr, 1971). About 40 plants per genotype were phenotyped.

EMSA

The full-length E1 coding sequence was cloned into the KpnI and SalI sites of vector pET32a and fused to the His-Tag to generate the His-E1 fusion vector. The full-length GmFDL19 coding sequence was cloned into the PVP13 vector via LR recombination to generate the MBP-GmFDL19 fusion vector. The resulting vectors were transformed into Escherichia coli strain Rosetta (DE3). The His-Tag recombinant protein was affinity-purified using Ni-NTA His-Bind resin (Millipore, Cat. No. 70666). The MBP-GmFDL19 recombinant protein was affinity-purified using amylose resin (Milipore, Cat. No. 70666). EMSA supershift assay was conducted using antibodies against His tag. His tag monoclonal antibody (Thermo Fisher, Cat No MA1-21305) was added at the indicated volume per 15-μL reaction system. For competition with unlabeled probe, the labeling step was omitted. EMSA was performed using a Chemiluminescent EMSA kit (Beyotime, Cat No GS009). Probe sequences are given in Supplemental Table S1.

Transient expression assays in N. benthamiana leaves

Promoter fragments were amplified from the soybean cultivar Dongnong50 genome using the primers listed in Supplemental Table S1: a 2,843-bp promoter fragment upstream of the GmMDE05 translation initiation codon, a 3,371-bp promoter fragment upstream of the GmMDE06 translation initiation codon, and a 3,194-bp promoter fragment upstream of the GmMDE17 translation initiation codon. The PCR amplicons were then cloned into HindIII/BamHI-linearized pGreenII vector by the ClonExpress MultiS One Step Cloning Kit (Vazyme, Cat. No. C113-01). The p35S::E1 vector was previously constructed (Zhang et al., 2016b). To generate the GmMDE06 promoter fragment harboring a deletion of the B3 domain recognition motif, the GmMDE06 promoter was PCR amplified as multiple overlapping ampli-con using the primers listed in Supplemental Table S1 and fused to HindIII/BamHI-linearized pGreenII vector by the ClonExpress MultiS One Step Cloning Kit (Vazyme, Cat. No. C113-01). The resulting constructs were introduced into Agrobacterium strain GV3101. Transformed Agrobacteria were mixed in the appropriate combinations and co-infiltrated into epidermal cells of N. benthamiana leaves. Leaves were infiltrated with each LUC reporter alone: (pGmMDE06::LUC/pGmMDE17::LUC/pGmMDE05::LUC/); with the empty vector and each LUC reporter individually in a 1:1 ratio (pGmMDE06::LUC/pGmMDE17::LUC/pGmMDE05::LUC/); or with the 35S::E1 effector construct and each LUC reporter individually in a 1:1 ratio (pGmMDE06::LUC/pGmMDE17::LUC/pGmMDE05::LUC/). At least five leaves from independent N. benthamiana plants were infiltrated for each experiment. Leaves were sprayed with 100-μM D-luciferin and incubated in the dark for 5 min before luminescence measurements. Images were taken with a low-light cooled charge-coupled device imaging apparatus (Tanon 5200). The experiments were repeated twice.

ChIP

The 35S::MYC-E1 transgenic soybean line (WEM) and the wild-type cultivar Williams 82 were used for ChIP as described previously (Saleh et al., 2008). Briefly, approximately 1.5 g of fully expanded trifoliate leaves from 24-d-old soybean plants was collected 4 h after dawn under LD conditions and crosslinked in 1% (w/v) formaldehyde under a vacuum. Crosslinking was then stopped by adding Glycine (final concentration of 100 mM). The samples were ground to powder in liquid nitrogen and used to isolate nuclei. Anti-trimethyl-Histone H3 (Lys27) (1:250 dilution) rabbit polyclonal antibody (Millipore, Cat. No. 07-449) and mouse anti-cMYC antibody (Thermo Fisher, Cat. No. MA1-980) were used to immunoprecipitate the protein–DNA complexes with Protein A + G magnetic beads (Millipore, Cat No. 16-663). The precipitated DNA was recovered and analyzed by qPCR. Relative fold enrichment was calculated by normalization to the reference gene TUA5. Rabbit IgG (Millipore, Cat No NI01) was used as negative control. Relative fold enrichment of Cons4 (soybean ATP binding cassette transporter gene) was calculated and used as an unrelated control to verify the validity of the ChIP assay. The data are presented as mean ± SD of three biological replicates (n = 3 plants) Primers used for ChIP-qPCR are listed in Supplemental Table S1.

Immunoblot analysis

Total proteins were extracted from the leaves of soybean cultivar Williams and 35S::MYC-E1 transgenic soybean lines in protein extraction buffer (50-mM Tris–HCl, pH 7.5, 120-mM NaCl, 5-mM EDTA, 10-mM DTT, 0.2% [w/v] Triton X-100, 0.1% [w/v] sodium dodecyl sulfate (SDS), and protease inhibitor cocktail). An anti-cMYC antibody (Thermo Fisher, Cat. No. MA1-980) was used at a dilution of 1:2,000. Immunoblotting was performed as described previously (Lu et al., 2017).

Accession numbers

Sequence data from this article can be found in the GenBank/EMBL database under the following accession numbers: GmMDE04 (Glyma.04G159300), GmMDE05 (Glyma.05G018800), GmMDE06 (Glyma.06G205800), GmMDE08 (Glyma.08G250800), GmMDE13 (Glyma.13G052100), GmMDE17 (Glyma.17G081200), GmMDE19 (Glyma.19G034600), E1(Glyma.06G207800), Dtr1(Glyma.19G194300), and GmFDL19(Glyma.19G122800). Sequence data of RNA-seq
from this article can be found in the National Genomics Data Center database under accession number PRJCA005287.

**Supplemental data**

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Generation of the e1-1 mutant by CRISPR/Cas9 in soybean.

**Supplemental Figure S2.** Phylogenetic tree of soybean MADS-box genes and well-known MADS-box genes from Arabidopsis thaliana.

**Supplemental Figure S3.** Tissue-organ expression analysis of GmMDEs in different soybean cultivars.

**Supplemental Figure S4.** E1 binds to DNA in a non-sequennc-specific manner in vitro.

**Supplemental Figure S5.** The repression of GmMDE06 transcription by E1 is not via canonical B3 domain recognition motifs.

**Supplemental Figure S6.** Phenotypes and immunoblot analysis of 35S:cMYC-E1 soybean plants.

**Supplemental Figure S7.** ChiP-qPCR analysis of H3K27me3 levels and cMYC-E1 binding at the GmMDE06 promoter.

**Supplemental Table S1.** Primers used in this study.

**Supplemental Data Set 1.** Differentially expressed genes in fully expanded trifoliolate leaves of E1-overexpressing transgenic soybean.

**Acknowledgments**

We would like to thank Professor Jiang N (Michigan State University, USA) for critical reading of the manuscript and Professor ShouYi Chen from Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and JingBo Jin from Institute of Botany, Chinese Academy of Sciences for kindly technical guidance of this manuscript.

**Funding**

This work was supported by the Key Research Program of the Chinese Academy of Sciences (ZDRW-ZS-2019-2-0201), National Natural Science Foundation of China (31771818 and 31971970), Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24010105-4), Natural Science Foundation of Heilongjiang Province of China (ZDRW-ZS-2019-2-0201), Natural Science Foundation of Heilongjiang Province of China (YQ2021C034), and Hainan Yazhou Bay Seed lab (B21Hj0101).

*Conflict of interest statement.* None declared.

**References**

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science 309: 1052–1056

Bernard RL (1971) 2 Major genes for time of flowering and maturity in soybeans. Crop Science 11: 242

Bonato ER, Vello NA (1999) E-6, a dominant gene conditioning early flowering and maturity in soybeans. Genet Mol Biol 22: 229–232

Bowman JL, Alvarez J, Weigel D, Meyerowitz EM, Smyth DR (1993) Control of flower development in Arabidopsis thaliana by Apetala1 and interacting genes. Development 119: 721–743

Bratzel F, Lopez-Torrejon G, Koch M, Del Pozo JC, Calonje M (2010) Keeping cell identity in Arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. Curr Biol 20: 1853–1859

Bu T, Lu S, Wang K, Dong L, Li S, Xie Q, Xu X, Cheng Q, Chen L, Fang C, et al. (2021) A critical role of the soybean evening complex in the control of photoperiod sensitivity and adaptation. Proc Natl Acad Sci USA 118: 8

Buzzell RI (1971) Inheritance of a soybean flowering response to fluorescent-daylength conditions. Can J Genet Cytol 13: 703

Chi Y, Huang F, Liu H, Yang S, Yu D (2011) An APETALA1-like gene of soybean regulates flowering time and specifies floral organs. J Plant Physiol 168: 2251–2259

Chi YJ, Wang TT, Xu GL, Yang H, Zeng XR, Shen YX, Yu DY, Huang F (2017) GmAGL1, a MADS-box gene from soybean, is involved in floral organ identity and fruit dehiscence. Front Plant Sci 8: 175

Cober ER, Voldeng HD (2001) Low R: FR light quality delays flowering of E7E7 soybean lines. Crop Sci 41: 1823–1826

Dong L, Fang C, Cheng Q, Su T, Kou K, Kong L, Zhang C, Li H, Hou Z, Zhang Y, et al. (2021) Genetic basis and adaptation trajectory of soybean from its temperate origin to tropics. Nat Commun 12: 5445

Fehr WR (1971) Stage of development descriptions for soybeans, Glycine max (L) Merrill. Merrill Crop 11: 929–931

Flores T, Karpova O, Su X, Zeng P, Bilyeu K, Sleper DA, Nguyen HT, Zhang ZJ (2008) Silencing of GmFAD3 gene by siRNA leads to low alpha-linolenic acids (18:3) of fad3-mutant phenotype in soybean [Glycine max (Merr.)]. Transgenic Res 17: 839–850

Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development 125: 1509–1517

Huang F, Chi YJ, Gai JY, Yu DY (2009) Identification of transcription factors predominantly expressed in soybean flowers and characterization of GmSEP1 encoding a SEPELLATA1-like protein. Gene 438: 40–48

Huang F, Xu GL, Chi YJ, Liu HC, Xue Q, Zhao TJ, Gai JY, Yu DY (2014) A soybean MADS-box protein modulates floral organ numbers, petal identity and sterility. BMC Plant Biol 14

Hyun Y, Richter R, Vincent C, Martinez-Galgagos R, Porri A, Coupland G (2016) Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the Arabidopsis shoot meristem. Dev Cell 37: 254–266

Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucleic Acids Res 27: 470–478

Kong FJ, Nan HY, Cao D, Li Y, Wu FF, Wang JL, Lu SJ, Yuan XH, Cober ER, Abe J, et al. (2014) A new dominant gene E9 conditions early flowering and maturity in soybean. Crop Sci 54: 2529–2535

Kopertekh L, Schiemann J (2005) Agroinfiltration as a tool for transient expression of cre recombinase in vivo. Transgenic Res 14: 793–798

Lanzuolo C, Orlando V (2012) Memories from the polycym group proteins. Annu Rev Genet 46: 561–589

Lee J, Lee I (2010) Regulation and function of SOC1, a flowering pathway integrator. J Exp Bot 61: 2247–2254

Li C, Li YH, Li Y, Lu H, Hong H, Tian Y, Li H, Zhao T, Zhou X, Liu J, et al. (2020) A domestication-associated gene GmPRR3B regulates the circadian clock and flowering time in soybean. Mol Plant 13: 745–759
Liu B, Kanazawa A, Matsumura H, Takahashi R, Harada K, Abe J (2008) Genetic redundancy in soybean photoreceptors associated with duplication of the phytochrome A gene. Genetics 180: 995–1007

Liu BH, Watanabe S, Uchiyama T, Kong FJ, Kanazawa A, Xia ZJ, Nagamatsu A, Arai MA, Yamada T, Kitamura K, et al. (2010) The soybean stem growth habit gene Dtc1 is an ortholog of Arabidopsis TERMINAL FLOWER1. Plant Physiol 153: 198–210

Liu W, Jiang B, Ma L, Zhang S, Zhai H, Xu X, Hou W, Xia Z, Wu C, Sun S, et al. (2018) Functional diversification of Flowering LOCUS T homologs in soybean: GmFT1a and GmFT2a/5a have opposite roles in controlling flowering and maturation. New Phytol 217: 1381–1395

Liu YF, Zhang DJ, Ping JQ, Li S, Chen ZX, Ma JX (2016) Innovation of a regulatory mechanism modulating semi-determinate stem growth through artificial selection in soybean. PLoS Genet 12: e1005818

Lu S, Dong L, Fang C, Liu S, Kong L, Cheng Q, Chen L, Su T, Nan H, Zhang D, et al. (2020) Stepwise selection on homoeologous PRR genes controlling flowering and maturity during soybean domestication. Nat Genet 52: 428–436

Lu SJ, Zhao XH, Hu YL, Liu SL, Nan HY, Li XM, Fang C, Cao D, Shi XY, Kong LP, et al. (2017) Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. Nat Genet 49: 773

Mandel MA, GustafsonBrown C, Savidge B, Yanofsky MF (1992) Molecular characterization of the Arabidopsis floral homeotic gene Apetala1. Nature 360: 273–277

Mclaren BA, Bernard RL (1987) A new gene affecting the time of flowering and maturity in soybeans. J Heredity 78: 160–162

Nan HY, Cao D, Zhang DY, Li Y, Liu S, Tsubokura Y, Yamada T, Tsubokura Y, et al. (2010) Map-based cloning of the gene associated with the soybean maturity locus E1. Genetics 182: 1251–1262

Wei W, Huang J, Hao YJ, Zou HF, Wang HW, Zhao JY, Liu XY, Zhang WK, Ma B, Zhang JS, et al. (2009) Soybean GmPHD-type transcription regulators improve stress tolerance in transgenic Arabidopsis plants. PLoS One 4: e7209

Xia Z, Watanabe S, Yamada T, Tsubokura Y, Nakamoto Y, Yamanaka N, Takahashi R, Ishimoto M, Anai T, et al. (2009) Map-based cloning of the gene associated with the soybean maturity locus E2. Genetics 182: 1251–1262

Zhai et al. 1036 | PLANT PHYSIOLOGY 2022: 189; 1021–1036

Teper-Bammolker P, Samach A (2005) The flowering integrator FT regulates SEPALLATA3 and FRUITFULL accumulation in Arabidopsis leaves. Plant Cell 17: 2661–2675

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. Plant Mol Biol 42: 115–149

Theissen G, Saedler H (2001) Plant biology. Floral quartets. Nature 409: 469–471

Tian ZX, Wang XB, Lee R, Li YH, Specht JE, Nelson RL, McClean PE, Qiu LJ, Ma JX (2010) Artificial selection for determinate growth habit in soybean. Nat Protoc Acad Sci USA 107: 8563–8568

Tsubokura Y, Watanabe S, Xia Z, Yamada T, Yamagata H, Kaga A, Katayose Y, Abe J, Ishimoto M, Harada K (2014) Natural variation in the genes responsible for maturity loci E1, E2, E3 and E4 in soybean. Ann Bot 113: 429–441

Ulmov T, Hagen G, Guilfoyle TJ (1997) ARF1, a transcription factor that binds to auxin response elements. Science 276: 1865–1868

Upadhyay AP, Ellis RH, Summerfield RJ, Roberts EH, Qi A (1996) Characterization of photothermal flowering responses in maturity isolines of soybean [Glycine-max (L) Merrill] Cv Clark. Ann Bot 74: 87–96

Watanabe S, Hideshima R, Xia ZJ, Tsubokura Y, Sato S, Nakamoto Y, Yamanaka N, Takahashi R, Ishimoto M, Anai T, et al. (2009) Map-based cloning of the gene associated with the soybean maturity locus E3. Genetics 182: 1251–1262

Xia Z, Watanabe S, Yamada T, Tsubokura Y, Nakashima H, Zhai H, Anai T, Sato S, Yamazaki T, Lu S, et al. (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. Proc Natl Acad Sci USA 109: E2155–2164

Xu JH, Zhang XF, Zhang QZ, Li HY (2010) Overexpression of the GmGAL2 gene accelerates flowering in Arabidopsis. Plant Mol Biol Rep 28: 704–711

Xu ML, Yamagishi N, Zhao C, Takeshima R, Kasai M, Watanabe S, Kanazawa A, Yoshikawa N, Liu BH, Yamada T, et al. (2015) The soybean-specific maturity gene E1 family of floral repressors controls night-break responses through down-regulation of FLOWERING LOCUS T orthologs. Plant Physiol 168: 1735–1746

Yoo SD, Cho HY, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient expression analysis. Nat Protoc 2: 1565–1572

Yuan W, Luo X, Li Z, Yang W, Wang Y, Liu R, Du J, He Y (2016) A cis cold memory element and a trans epigenome reader mediate Polycomb silencing of FLC by vernalization in Arabidopsis. Nat Genet 48: 1527–1534

Zhai H, Lu S, Liang S, Wu H, Zhan X, Liu B, Kong F, Yuan X, Li J, Xia Z (2014) GmFT4, a homolog of FLOWERING LOCUS T, is positively regulated by E1 and functions as a flowering repressor in soybean. PLoS One 9: e89030

Zhang HY, Yan HL, Zhang D, Yu DY (2016a) Ectopic expression of a soybean SVP-like gene in tobacco causes abnormal floral organs and shortens the vegetative phase. Plant Growth Regul 80: 345–353

Zhang XZ, Zhai H, Wang YY, Tian XJ, Zhang YP, Wu HY, Lu SX, Yang G, Li YQ, Wang L, et al. (2016b) Functional conservation and diversification of the soybean maturity gene E1 and its homologs in legumes. Scient Rep 6: 29548