CONSTANS Polymorphism Modulates Flowering Time and Maturity in Soybean

Mohammad Abdul Awal Khan†1, Shouwei Zhang†1, Reza Mohammad Emon1,2†, Fulu Chen†1, Wenwen Song†1, Tingting Wu†1, Shan Yuan†1, Cunxiang Wu†1, Wensheng Hou†1, Shi Sun†1, Yongfu Fu†1, Bingjun Jiang*1 and Tianfu Han*1

1 MARA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, 2 Plant Breeding Division, Bangladesh Institute of Nuclear Agriculture, Mymensingh, Bangladesh

CONSTANS (CO) plays a critical role in the photoperiodic flowering pathway. However, the function of soybean CO orthologs and the molecular mechanisms in regulating flowering remain largely unknown. This study characterized the natural variations in CO family genes and their association with flowering time and maturity in soybeans. A total of 21 soybean CO family genes (GmCOLs) were cloned and sequenced in 128 varieties covering 14 known maturity groups (MG 0000-MG X from earliest to latest maturity). Regarding the whole genomic region involving these genes, GmCOL1, GmCOL3, GmCOL8, GmCOL9, GmCOL10, and GmCOL13 were conserved, and the remaining 15 genes showed genetic variation that was brought about by mutation, namely, all single-nucleotide polymorphisms (SNPs) and insertions-deletions (InDels). In addition, a few genes showed some strong linkage disequilibrium. Point mutations were found in 15 GmCOL genes, which can lead to changes in the potential protein structure. Early flowering and maturation were related to eight genes (GmCOL1/3/4/8/13/15/16/19). For flowering and maturation, 11 genes (GmCOL2/5/6/14/20/22/23/24/25/26/28) expressed divergent physiognomy. Haplotype analysis indicated that the haplotypes of GmCOL5-Hap2, GmCOL13-Hap2/3, and GmCOL28-Hap2 were associated with flowering dates and soybean maturity. This study helps address the role of GmCOL family genes in adapting to diverse environments, particularly when it is necessary to regulate soybean flowering dates and maturity.

Keywords: soybean, GmCOL orthologue, natural variation, flowering time, maturity group

INTRODUCTION

Plants can adapt to different environmental conditions in response to various day lengths (photoperiods). In the photoperiodic flowering pathway, CONSTANS (CO), which is a B-box-containing transcription factor (Robson et al., 2001; Gangappa and Botto, 2014), plays a key role (Putterill et al., 1995; Suarez-Lopez et al., 2001). CO also possesses a CONSTANS, CONSTANS-LIKE, TIMING OF CAB1 (CCT) domain at its C-terminus involved in DNA binding (Strayer et al., 2000; Robson et al., 2001). CYCLING DOF FACTOR (CDF) family proteins bind to the CO promoter to repress its transcription in the morning (Imaizumi et al., 2005; Fornara et al., 2009). FLAVIN-BINDING, KELCH REPEAT, and F-BOX1 (FKF1) interact with GIGANTEA (GI) to degrade the CDF1 protein, which results in the elevation of CO mRNA (Sawa et al., 2007). CO acts in the phloem of leaf vascular tissues to activate FLOWERING LOCUS T (FT)
expression, which causes flowering under linkage disequilibrium (LD) conditions (Takada and Goto, 2003; An et al., 2004). SUPPRESSOR OF PHYTOCHROME A-105 (SPA) proteins interact with the CCT domain of CO (Laubinger et al., 2006). CO is ubiquitinated by CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) and degraded by the 26S proteasome (Jiang et al., 2019; Liu L. J. et al., 2008). HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES1 (HOS1) interacts with CO and participates in the degradation of CO mediated by red light (Lazarov et al., 2012, 2015). Nucleoporin96 (Nup96) interacts with HOS1 to gate CO protein levels (Cheng et al., 2020). CO protein is stable under the light in the evening but degraded in the morning or in the dark under LD conditions (Valverde et al., 2004). All of the abovementioned comments are based on studies in Arabidopsis.

As a short-day plant (SDP) distributed over a vast range of latitudes, soybean is characterized as having 14 maturity groups from MG 0000 to MG X (Liu et al., 2017; Jiang et al., 2019). There are hundreds of genetic loci associated with flowering time and maturity in soybean, among which E1 (Xia et al., 2012), E2 (Watanabe et al., 2011), E3 (Watanabe et al., 2009), E4 (Liu B. et al., 2008), E6 (Fang C. et al., 2021), E9 (Kong et al., 2014), E10 (Samanfar et al., 2017), E11 (Lu et al., 2019), J (Lu et al., 2017; Yue et al., 2017) and Tof11/GmPrr37/GmPrr3b (Li et al., 2020; Lu et al., 2020; Wang et al., 2020), GmTof16 (Dong et al., 2021b), GmFUL (Dong et al., 2021a; Sun et al., 2021; Yue et al., 2021), and GmLUX (Bu et al., 2021; Fang X. et al., 2021) are molecularly identified. There are 28 CO orthologs in soybean (Fan et al., 2014), but the functions of most of the CO orthologs remain uncharacterized. It has been shown that GmCOL1a and GmCOL1b can fully complement the late-flowering phenotype of the col mutant in Arabidopsis (Wu et al., 2014). In contrast, GmCOL1a and GmCOL1b repressed flowering in soybean under long-day conditions (Cao et al., 2015; Wu et al., 2019).

Natural variation is associated with photoperiodic flowering and adaptation in different species (Alonso-Blanco et al., 2005; Balasubramanian et al., 2006; Rosas et al., 2014; Li et al., 2017; Lu et al., 2017; Bao et al., 2019; Jiang et al., 2019; Wu et al., 2020). In this study, 128 soybean varieties were selected and planted, covering all 14 maturity groups from MG 0000 to MG X with a continuous distribution in maturity groups (Jiang et al., 2019). Due to the short duration of the project, 21 of the 28 soybean COL genes (Fan et al., 2014) were sequenced, and their sequence polymorphisms were analyzed. Furthermore, we studied the haplotypes of these soybean GmCOL genes to discover their natural variations in association with flowering time and maturity. These results suggested that some natural variations in 21 soybean GmCOL genes were associated with flowering date and maturity.

**MATERIALS AND METHODS**

**Plant Materials**

Soybean varieties covering all 14 maturity groups (MG 0000-MG X) were assessed in this study (Jiang et al., 2019). In this study, 64 Chinese and Russian soybean varieties and 64 North American maturity group standard varieties were included, covering MG 0000-MG X, for a total of 128 accessions analyzed (Table 1).

**Investigation of Flowering and Maturity Dates**

The soybeans were planted in soil in 10-L pots and grown under natural conditions in Haidian District, Beijing (39.95°N, 116.32°E) on May 27, 2015, and May 18, 2016 (Jiang et al., 2019). After emergence (VE), seedlings of similar size were selected so that each pot contained five uniform plants. Each variety was planted in three pots. The four developmental stages of soybean, namely, VE, R1, R7, and R8 (Fehr and Caviness, 1977), were investigated, and the average value of three pots for each variety was used for statistical analysis.

**Table 1 | Soybean varieties and maturity groups.**

| Maturity group | North American varieties | Chinese-Russian varieties |
|---------------|--------------------------|--------------------------|
| MG 0000       | Maple Presto, OAC Vision, Rassvet, Jug 30, Mageva |斯塔/75, Hujiao07-2479, Hujiao07-2123, Dongdong 36, Paula, R-4, Dongdong 41, Lingbei 8 |
| MG 000        | Canatto, Maple Ridge, Daksoy, McCall, Agassiz |Mengdou 32, Beidou 16, Dongdong 44, Mengdou 11, Jiangmoudou 1, Heihe 18, Heihe 43, Heihe 27, Beidou 37, Dengke 1, Fengshou 12, Dongdong 4, Hefeng 25 |
| MG 9          | Trail, Chico, Barnes, Norpro, Dawson | |
| MG I          | Haroson, Kato, Parker, Granite, NE1900 |Heihe 16, Taxingheidou, Heihe 26, Sunion 14 |
| MG II         | Holt, Olympus, Century 84, IL1, LN92-7369 |Jilin 20, Yongchengzhuhoudou, Xiangchunhou 24, Tiefeng 19 |
| MG III        | A thaw, LN98-5699, KS3494, IL2, Williams 82 |Zhonghuang 13, Zhonghuang 30, Tiefeng 33, Zhonghuang 39, Xudou 9, Tiefeng 31, Jindou 19, Huachun 6, Huichun 9 |
| MG IV         | Flyer, Omaha, Calhoun, OF461, UA-4986 |Zheng 92116, Guandou 2, Jindou 39, Shanning 16, Houzimao |
| MG V          | Nathan, Hollady, Hutcheson, R01-3474F, TN04-5321 |Shandong 14, Dian 86-4, Dianrou 7 |
| MG VI         | Desha, Musen, D95-6271, G01-PR16, Boggs |Zhongdou 38, Wujiyuxiahegou, Suxianhou 19, Nannong 493-1 |
| MG VII        | Stonewall, Benning, Santee, Haggard |Huangfengwo, Tngshangbophuang, Hengyangbay.iqiu, Nanxiadou 25 |
| MG VIII       | Motte, Dowling, Crockett, Prichard, IAC-8, CIGRAS-51, CIGRAS-06 |Aijiaqingso, Pingguochuangdou, Nandou 12, Shangaraoyuqingsi, Lanzxadaoyingdou, Gudou 1, Nandou 17, Jialedaoyingdou, Guixia 3 |
| MG IX         | Jupiter, Alamo, FT-15, UFV-3 |Zigongdongdou |
| MG X          | I.C. 192 |
DNA Isolation, PCR, and Sequencing
Genomic DNA was extracted from fresh trifoliolate leaves using the standard cetyltrimethylammonium bromide (CTAB) method (Jiang et al., 2019). To amplify the 21 soybean GmCOL genes, 36 PCR primer pairs were used. GmCOL9/15/16/20/28 was amplified with 2 primer pairs: GmCOL4 with 11 primer pairs and GmCOL12/3/5/6/8/10/13/14/19/22/23/24/25/26 with 1 primer pair. The sequences of these primers and the template designation are listed in Supplementary Table 1. Target regions were amplified with the high-fidelity polymerase of KOD-Plus-Neo and KOD-FX. Their reaction conditions were 94°C for 2 min (98°C for 10 s, 68°C for 3 min), 35 cycles, and 68°C for 10 min. The PCR products were directly sequenced using the Sanger method (TSINGKE Biological Technology Company, BGI and Omega Genetics, Beijing, China). The polymorphic site information of the 21 soybean GmCOL family genes is listed in Supplementary Table 2.

Data Mining and Sequence Analysis
The annotated soybean COL gene sequence was downloaded from the Joint Genome Institute (JGI) Phytozome website. The protein sequences of the annotated Arabidopsis genes (TAIR9 release) were downloaded from the Arabidopsis Information Resource (TAIR) website. The CO-like gene sequence data were also collected from the USDA-ARS Soybean Genetics and Genomics Database (SoyBase Database). The reference sequence of the soybean variety Williams 82 was obtained from the Phytozome (version 12.0) database as a reference design for the amplification and sequencing primers of the COL genes.

Sequencing was performed using an ABI3730 sequencer. Multiple sequence alignment, editing, and stitching were carried out using ClustalW in MEGA5 with default parameters. Single-nucleotide polymorphism (SNP) analysis was carried out using TASSEL5 software. The haplotype analysis of the sequencing results was conducted by DNAsp to determine whether the protein coding was affected. LD was also evaluated using TASSEL5 software. The statistical analysis was performed using "R" software.

RESULTS

Diversity of Soybean Varieties in Flowering Time and Maturity
The results of the growth traits are presented in Table 2. The varieties utilized in this experiment exhibited significant diversity in flowering time and maturity (Table 2). In 2015, 19 varieties, namely, MC119, MC63, MC69, MC70, MC54, MC121, MC122, MC123, MC72, MC124, MC125, MC126, MC48, MC64, MC65, MC66, MC67, MC68, and MC71, flowered but failed to reach R7. In 2016, 50 varieties did not reach R7 (Table 2). In 2015, the days to R1 from emergence (VE) ranged from 21 (MC82, MG 0) to 121 days (MC72, MG VIII) with a span of 100 days. In 2016, the span time was 96 days between MG 0000 (MC05, 13 days) and MG VIII (MC129, 109 days). In 2016, the days to R7 ranged from 68 (MC05 and MC74, MG 0000) to 144 days (MC42, MG IV), and the days to R8 ranged from 76 (MC05, MG 0000) to 109 days (MC37, MG III). The combination of maturity group information for the days of flowering time from the emergence and the days of beginning maturity from emergence in 2016 is plotted in a box diagram (Figure 1).

GmCOL Gene Sequencing and Mutation Study
A large number of mutated sites were found in GmCOL2, GmCOL5, GmCOL14, GmCOL16, GmCOL20, and GmCOL28 (Figures 2B1,D1 and Supplementary Figures 1G1,I1, 2B1,H1). A majority of the mutation sites were in gene intron regions, but a few were located in protein-coding regions caused by insertions-deletions (InDels) and substitutions. In addition, GmCOL6, GmCOL22, GmCOL23, GmCOL24, GmCOL25, and GmCOL26 showed high mutation frequencies in the protein-coding regions (Supplementary Figures 1, 2). In contrast, GmCOL4, GmCOL15, and GmCOL19 expressed lower mutational occurrences (Supplementary Figures 1A1,H1, 2A1). The exception was GmCOL3, which did not show mutations (Figure 2C) in either the intron or exon region of the entire gene. GmCOL3 is a highly conserved gene with no change in sequence. However, the amino acid sequences encoded by GmCOL11, GmCOL8, GmCOL9, GmCOL10, and GmCOL13 were not affected (Figure 2A and Supplementary Figures 1C1-F1).

Soybean COL Gene Family Exhibited Different Linkage Disequilibrium
Linkage and association mapping were drawn among the SNPs of the gene via TASSEL5. GmCOL5 across the region almost from the starting to the end site presented a strong LD by completing the haplotype block (Figure 3D). GmCOL20, GmCOL24, and GmCOL25 across the region expressed strong LD as a haplotype block (Figures 3N,Q,R), and the polymorphic sites of GmCOL6, GmCOL8, GmCOL10, GmCOL13, and GmCOL19 showed a strong LD, with almost the entire region being a haplotype block (Figures 3E,F,H,I,M). Although GmCOL2, GmCOL16, GmCOL22, GmCOL26, and GmCOL28 had a high-sequence polymorphism, the entire region distributed a few LDs (Figures 3B,L,O,S,T) with the same high-sequence polymorphism of GmCOL4, GmCOL9, and GmCOL23, and quite a few strong LDs were distributed broadly (Figures 3C,G,P). GmCOL1, GmCOL14, and GmCOL15 did not reveal any LD (Figures 3A,J,K). GmCOL3 was the most conserved gene, showing no polymorphic site(s).

Haplotype Analysis and Its Association With Maturity Groups of Soybean Varieties
Based on LD and sequence clustering, haplotype definition was analyzed for the 21 soybean GmCOL gene families in
### TABLE 2 | Phenotypic variation of soybean varieties at Beijing.

| Code | Variety name  | MG  | VE-R1 | Phenotype in 2016 |
|------|---------------|-----|-------|-------------------|
| MC01 | Star 4/75     | 0000| 15.4 ± 1.2 | 70.6 ± 1.3 |
| MC02 | Huliao 07-0479| 0002| 16.4 ± 0.82 | 72.0 ± 1.27 |
| MC03 | Huliao 07-2123| 0000| 15.6 ± 0.50 | 71.3 ± 0.35 |
| MC04 | Dongnong 36   | 0000| 14.13 ± 0.51 | 69.13 ± 0.51 |
| MC05 | Paula          | 0000| 13.66 ± 0.48 | 68.4 ± 0.82 |
| MC06 | R-4           | 0000| 25.86 ± 0.74 | 71.46 ± 0.91 |
| MC07 | Dongnong 41   | 0000| 14.33 ± 0.48 | 69.46 ± 0.51 |
| MC08 | Maple-Presto   | 0000| 15.53 ± 0.51 | 69.2 ± 1.01 |
| MC09 | GCAC-Vision    | 0000| 16.73 ± 0.79 | 69.8 ± 1.52 |
| MC10 | Rassvet        | 0000| 17.33 ± 0.48 | 71.53 ± 0.91 |
| MC11 | Juq-30        | 0000| 23.43 ± 0.48 | 71.2 ± 0.41 |
| MC12 | Magrana        | 0000| 25.73 ± 0.45 | 75.66 ± 0.97 |
| MC13 | Canatto       | 0000| 18.2 ± 0.4 | 72.8 ± 1.8 |
| MC14 | Maple-Ridge    | 0000| 18.4 ± 0.5 | 72.0 ± 1.0 |
| MC15 | Dakosoy       | 0000| 24.3 ± 0.4 | 71.3 ± 0.4 |
| MC16 | McCall        | 0000| 23.6 ± 0.8 | 73.4 ± 1.2 |
| MC17 | Agassiz       | 0000| 25.2 ± 0.4 | 71.6 ± 0.8 |
| MC18 | Mengdou 32    | 0000| 22.6 ± 2.4 | 73.2 ± 1.0 |
| MC19 | Mengdou 18    | 0000| 19.9 ± 0.7 | 73.4 ± 1.5 |
| MC20 | Dongnong 44   | 0000| 24.8 ± 1.0 | 71.3 ± 1.4 |
| MC21 | Mengdou 11    | 0000| 24.3 ± 0.4 | 71.2 ± 1.3 |
| MC22 | Trail         | 0000| 27.3 ± 0.4 | 73.3 ± 0.9 |
| MC23 | Chico         | 0000| 26.4 ± 0.9 | 70.8 ± 1.5 |
| MC24 | Barnes        | 0000| 26.5 ± 0.7 | 68.8 ± 1.3 |
| MC25 | Norpro        | 0000| 25.3 ± 0.9 | 71.4 ± 0.50 |
| MC26 | Dawson        | 0000| 23.6 ± 0.4 | 70.2 ± 1.37 |
| MC27 | Jiangmodou 1  | 0000| 23.4 ± 0.9 | 69.7 ± 1.1 |
| MC28 | Hehei 18      | 0000| 25.2 ± 1.5 | 72.0 ± 1.0 |
| MC29 | Hehei 43      | 0000| 26.4 ± 0.5 | 71.9 ± 1.6 |
| MC30 | Hehei 27      | 0000| 23.6 ± 0.4 | 70.8 ± 1.0 |
| MC31 | Hehei 37      | 0000| 21.5 ± 0.9 | 69.5 ± 0.9 |
| MC32 | Mengdou 11    | 0000| 23.8 ± 0.4 | 69.6 ± 0.4 |
| MC33 | Fengshou 12   | 0000| 24.6 ± 0.9 | 71.0 ± 1.0 |
| MC34 | Dongnong 4    | 0000| 32.2 ± 1.0 | 74.1 ± 0.8 |
| MC35 | Hefeng 25     | 0000| 32.7 ± 0.4 | 75.5 ± 0.9 |
| MC36 | Haroson       | 0000| 25.2 ± 0.45 | 71.2 ± 0.4 |
| MC37 | Kato          | 0000| 27.5 ± 0.9 | 72.2 ± 0.4 |
| MC38 | Parker        | 0000| 25.4 ± 0.9 | 71.4 ± 0.9 |
| MC39 | Granite       | 0000| 25.3 ± 0.9 | 71.3 ± 0.9 |
| MC40 | NE1900        | 0000| 35.2 ± 0.4 | 70.7 ± 1.0 |
| MC41 | Heining 18    | 0000| 28.5 ± 0.9 | 71.8 ± 0.9 |
| MC42 | Taixinghedou  | 0000| 26.6 ± 0.9 | 72.73 ± 1.16 |
| MC43 | Heining 26    | 0000| 29.4 ± 0.5 | 70.8 ± 0.6 |
| MC44 | Suijing 14    | 0000| 27.2 ± 1.5 | 69.4 ± 1.3 |
| MC45 | Holt          | 0000| 30.3 ± 2.4 | 70.8 ± 0.8 |
| MC46 | Olympus       | 0000| 27.7 ± 0.9 | 72.3 ± 0.4 |
| MC47 | Century-84    | 0000| 26.6 ± 0.4 | 71.6 ± 0.8 |
| MC48 | IL1           | 0000| 30.4 ± 5.5 | 70.2 ± 1.7 |
| MC49 | LN92-7369     | 0000| 30.4 ± 1.4 | 76.2 ± 1.0 |
| MC50 | Jin 20        | 0000| 25.7 ± 1.0 | 70.3 ± 0.9 |
| MC51 | Yongchengzhoudou| 0000| 41.3 ± 1.2 | 72.1 ± 1.2 |
| MC52 | Xiangchoudou 24| 0000| 43.8 ± 1.0 | 77.4 ± 0.5 |
| MC53 | Tiefeng 19    | 0000| 24.4 ± 0.5 | 71.2 ± 1.2 |
| MC54 | Aither        | 0000| 30.3 ± 0.9 | 76.3 ± 0.9 |
| MC55 | Zhonghuang 13 | 0000| 43.3 ± 0.4 | 78.4 ± 0.9 |
| MC56 | Zhonghuang 30 | 0000| 37.8 ± 0.9 | 78.3 ± 0.9 |
| MC57 | LN92-5699     | 0000| 32.3 ± 0.9 | 80.2 ± 1.5 |
| MC58 | KS3494        | 0000| 37.4 ± 0.5 | 79.8 ± 0.9 |
| MC59 | IL2           | 0000| 42.1 ± 0.8 | 105.2 ± 2.3 |
| MC60 | Williams 82   | 0000| 36.8 ± 0.3 | 107.3 ± 0.8 |

(Continued)
| Code | Variety name | MG | VE-R1   | VE-R7   | VE-R8   |
|------|--------------|----|---------|---------|---------|
| MC97 | Tiefeng 33   | III| 29.5 ± 0.5 | 78.7 ± 0.4 | 106.8 ± 1.2 |
| MC99 | Zhongdou 39  | III| 44.5 ± 0.5 | 77.4 ± 0.5 | 105.2 ± 1.8 |
| MC100| Xudou 9      | III| 37.9 ± 0.7 | 78.0 ± 1.0 | 106.4 ± 1.5 |
| MC101| Tiefeng 31   | III| 24.2 ± 1.1 | 72.2 ± 1.3 | 105.2 ± 1.1 |
| MC102| Jindou 19    | III| 25.8 ± 1.3 | 77.4 ± 1.5 | 104.8 ± 1.5 |
| MC103| Huachun 6    | III| 31.4 ± 0.5 | 77.3 ± 0.4 | 105.6 ± 0.9 |
| MC104| Haidou 9     | III| 34.3 ± 0.4 | ND       | ND       |
| MC40 | Flyer        | IV | 35.5 ± 1.4 | 143.0 ± 1.4 | ND       |
| MC41 | Omaha        | IV | 38.0 ± 1.0 | 143.4 ± 1.4 | ND       |
| MC42 | Cathoun      | IV | 37.3 ± 0.4 | 144.4 ± 0.5 | ND       |
| MC43 | CF461        | IV | 31.8 ± 1.0 | 143.7 ± 0.4 | ND       |
| MC44 | LA-4805      | IV | 61.2 ± 0.4 | ND       | ND       |
| MC105| Zheng 92116  | IV | 37.2 ± 0.4 | ND       | ND       |
| MC106| Guandou 2    | IV | 35.8 ± 1.0 | ND       | ND       |
| MC107| Jindou 39    | IV | 36.2 ± 1.3 | 142.9 ± 1.0 | ND       |
| MC108| Shangdong 14 | IV | 44.7 ± 0.4 | ND       | ND       |
| MC111| Dian 86-4    | V  | 43.6 ± 1.5 | ND       | ND       |
| MC113| Dian dou 7   | V  | 88.3 ± 0.9 | ND       | ND       |
| MC50 | Desha        | VI | 48.0 ± 0.8 | ND       | ND       |
| MC51 | Musen        | VI | 80.0 ± 1.0 | ND       | ND       |
| MC52 | D95-6271     | VI | 82.3 ± 1.5 | ND       | ND       |
| MC53 | G01-PR16     | VI | 81.6 ± 0.9 | ND       | ND       |
| MC54 | Mogi         | VI | 79.8 ± 1.5 | ND       | ND       |
| MC112| Zhongdou 38  | VI | 83.6 ± 2.5 | ND       | ND       |
| MC114| Wuhuasuydeliang | VI | 86.3 ± 0.9 | ND       | ND       |
| MC108| Suxiandou 19 | VI | 85.8 ± 1.2 | ND       | ND       |
| MC116| Jiangang 49-3 | VI | 91.2 ± 0.4 | ND       | ND       |
| MC55 | Stone wall   | VI | 81.4 ± 0.5 | ND       | ND       |
| MC57 | Sanle        | VII| 82.6 ± 1.8 | ND       | ND       |
| MC58 | Hapoo        | VII| 94.6 ± 1.2 | ND       | ND       |
| MC56 | Benning      | VII| 93.3 ± 0.9 | ND       | ND       |
| MC117| Huagfengwu   | VII| 93.7 ± 0.4 | ND       | ND       |
| MC118| Tongshaniopshuang | VII | 93.2 ± 1.2 | ND       | ND       |
| MC60 | Motle        | VIII| 93.3 ± 0.9 | ND       | ND       |
| MC61 | Dowling      | VIII| 95.8 ± 1.5 | ND       | ND       |
| MC62 | Crockett     | VIII| 65.4 ± 0.2 | 142.2 ± 0.5 | 158.0 ± 1.0 |
| MC128| Hengyangbaoyueqing | VII | 96.9 ± 1.0 | ND       | ND       |
| MC119| Nanxian 25   | VII| 92.6 ± 0.9 | ND       | ND       |
| MO63 | Pichard      | VIII| 94.8 ± 1.0 | ND       | ND       |
| MO69 | CIGRAS-51    | VIII| 81.4 ± 0.5 | ND       | ND       |
| MO70 | CIGRAS-06    | VIII| 81.5 ± 0.9 | ND       | ND       |
| MO54 | Aljaqing     | VIII| 94.3 ± 0.4 | ND       | ND       |
| MC121| Pingguoiaxie | VII| 92.9 ± 1.0 | ND       | ND       |
| MC122| Nanxian 12   | VIII| 91.3 ± 0.4 | ND       | ND       |
| MC123| Shangraochangzi | VIII | 92.8 ± 1.0 | ND       | ND       |
| MC72 | Zgongdanding | VIII| 73.4 ± 1.5 | ND       | ND       |
| MC124| Lanyingdongxi | VIII| 91.6 ± 0.5 | ND       | ND       |
| MC125| Quidu 1      | VIII| 94.3 ± 0.4 | ND       | ND       |
| MC126| Nanxian 17   | VIII| 91.6 ± 0.4 | ND       | ND       |
| MC48 | Jiangleiding | VIII| 99.5 ± 0.5 | ND       | ND       |
| MC129| Guina 3      | VIII| 109.2 ± 1.0 | 118.4 ± 0.2 | ND       |
| MO64 | Jupiter      | IX | 96.3 ± 0.9 | ND       | ND       |
| MO65 | Alamo        | IX | 96.0 ± 1.3 | ND       | ND       |
| MO66 | FT-15        | IX | 94.2 ± 1.3 | ND       | ND       |
| MO67 | UFV-3        | IX | 96.2 ± 1.0 | ND       | ND       |
| MO88 | IAC-8        | IX | 90.0 ± 1.0 | ND       | ND       |
| MC71 | I.C.-192     | X  | 80.4 ± 1.5 | ND       | ND       |

ND, Not available data; VE, Variation after emergence.
the 128 varieties investigated. A total of 21 haplotypes of the GmCOL gene family variation types and polymorphic sites used for composing haplotypes are listed in Figure 2 and Supplementary Tables 3, 4 (Supplementary Figures 1, 2), respectively. Supplementary Table 5 shows 21 GmCOL protein groups associated with haplotypes.

GmCOL2-Hap2 was dominant in the 41 accessions covering all maturity groups, except MG X. GmCOL2-Hap4 showed early flowering compared with the other haplotype groups (Figures 2B3, B4, Table 2, and Supplementary Table 3). GmCOL4-Hap1 accounted for 99 varieties, with all maturity groups excluding MG X (Supplementary Figure 1A2). GmCOL4-Hap1 had an earlier flowering date in 2016. GmCOL5-Hap2 was present in 56 accessions, covering all maturity groups (Figure 2D1). GmCOL6-Hap1 was the most abundant in 70 varieties from different maturity groups, including MG 0000-MG V and MG VIII-MG X (Supplementary Figure 1B1). Among the 11 haplotypes, GmCOL9-Hap1 was
dominant in 67 varieties from maturity groups MG 0000-MG X (Supplementary Figures 1D1,D2). GmCOL9-Hap3 and GmCOL9-Hap4 were distributed in a few varieties, and the flowering time for both haplotypes was early. GmCOL10-Hap1 was the most abundantly expressed in 91 varieties from all maturity groups except MG X (Supplementary Figures 1E1,E2). GmCOL16-Hap1 was rich in accessions distributed in the maturity groups MG 0000-MG VIII (Supplementary Figures 1F1,H1). Among all the haplotypes of GmCOL28, GmCOL8-Hap1 was the most common, accounting for 59 accessions, and was distributed in the maturity groups MG 0000, MG 00, MG 0, MG I-MG VIII, and MG X (Supplementary Figures 2H1,H2). In this experiment, no haplotype variants were found in GmCOL1 and GmCOL3 (Figures 2A,C). The Hap1 series of haplotypes of GmCOL8/13/14/15/19/22/23/24/25/26 was the most abundant in the varieties and was distributed in all 14 maturity groups (Supplementary Figures 1C1,F1,G1,H1, 2A1,C1,D1,E1,F1,G1).

### Haplotypes Associated With Flowering Time

The haplotype groups of 21 gene families are presented in Table 3. An analysis of variance (ANOVA) was performed to elucidate the natural variations in flowering time (VE-R1) of the GmCOL gene families based on haplotype groups. In this analysis, eight genes (GmCOL2/5/9/13/15/16/25/28) showed significant results from the first emergence to the first flowering date (VE-R1) in both years, and three genes [GmCOL8 (2015), GmCOL22 (2015), and GmCOL26 (2016)] exhibited significant results in a single year (Supplementary Table 6). The remaining genes (GmCOL4/6/10/14/19/20/23/24) did not show significant results in these years. Notably, GmCOL1 and GmCOL3 were the most conserved, and no polymorphisms in the coding region were observed.

Among the 11 haplotype groups of GmCOL2, 4 were analyzed, and GmCOL2-Hap3 was significantly different from every other haplotype group (Figure 4A). The GmCOL2-Hap1/2/4 haplotype groups, which appeared in the above haplotype group, were not significantly different from each other. Analysis of GmCOL5
It has been shown that GmCOL1a (Glyma08g28370) and GmCOL1b (Glyma18g51320) in soybean are the closest Arabidopsis COL2 orthologs to CO (Thakare et al., 2010). The Arabidopsis COL2 orthologs to CO (Thakare et al., 2010).

It has been shown that GmCOL1a (Glyma08g28370) and GmCOL1b (Glyma18g51320) in soybean are the closest Arabidopsis COL2 orthologs to CO (Thakare et al., 2010). The

DISCUSSION

Selected Soybean Varieties Showed Diversity in Flowering Time and Maturity

The varieties analyzed in this study showed the diversity of flowering time and maturity, which indicated that the flowering time (VE-R1) and the days from emergence to physiological maturity (VE-R7) were related to the variety trait and natural environment (Jiang et al., 2019). According to the photoperiod responses, many soybean varieties have evolved to adapt to a broad range of growing areas in China (Wang et al., 2016).

The flowering time of the selected varieties ranged from 21 to 121 days and 13.66 to 109.2 days, and the maturity time ranged from 61.2 to 150.7 days and 68.4 to 144.4 days, respectively, for two consecutive years (2015 and 2016) (Jiang et al., 2019), indicating high diversity (Table 2). However, Paula showed the shortest VE-R1 in the 2nd year, with the maturity group MG 0000. Paula is recognized as a high-latitude cold region (HCR) soybean variety and HCR soybean, which typically matures early (Jia et al., 2014). This observation suggested that the maturity group MG 0000 was relatively stable for days to start flowering from emergence.

GmCOL Orthologs Showed Divergence in Sequence Polymorphism

It has been shown that GmCOL1a (Glyma08g28370) and GmCOL1b (Glyma18g51320) in soybean are the closest Arabidopsis COL2 orthologs to CO (Thakare et al., 2010). The

Figure 4H. For GmCOL28 Hap2 (S1066), the original nucleotide (T\textsuperscript{263}) was mutated to C\textsuperscript{263} in the coding sequence, resulting in a missense mutation in the amino acid sequence (V\textsuperscript{88} to A\textsuperscript{88}), which was located in the first B-box, and may affect the protein and protein interaction. Thus, the results revealed that Hap2 of GmCOL28 was closely associated with flowering time.

showed that in both years two major haplotypes showed significant differences from each other, and Hap2 was related to late flowering compared with Hap1 (Figure 4B). In terms of GmCOL5 Hap2 (S999), the original nucleotide (A\textsuperscript{499}) was mutated to G\textsuperscript{499} in the coding sequence, resulting in a missense mutation in the amino acid sequence (K\textsuperscript{167} to E\textsuperscript{167}). K\textsuperscript{167} was located between the B-box and the CCT domain. In comparison,
transgenic soybean line overexpressing GmCOL1a flowered late under long-day or natural conditions (Cao et al., 2015). There is a single-nucleotide substitution at the GmCOL1b CCT domain in the gmcol1b mutant, leading to the mutagenesis of conserved threonine at amino acid 314 into isoleucine, which results in early flowering of the gmcol1b mutant (Cao et al., 2015). In this study, the varieties that were taken for the observation of natural variations regarding flowering time and maturation showed divergent characteristics in flowering time (Jiang et al., 2019). Allelic variation can have an effect on flowering time (Irwin et al., 2016). In this study, the insertion and single-nucleotide substitution (InDel918, InDel919, InDel920, s958, s1138, s1139, s1218, and s1260) found in the first exon in GmCOL2 led to missense mutations, which might have an effect on flowering time and maturity. The other genes in the GmCOL family showed various types of point mutations (Figure 2 and Supplementary Figures 1–5). The loss of functions may have an impact on gene function, and the presence of these mutations suggested that polymorphism could be the main cause of the diversity of soybean flowering time.

**Polymorphism of GmCOL Family Genes Was Associated With Flowering Time and Maturity**

Haplotype-based analyses have been successfully carried out in different crops, such as maize (Zea mays L.) (Weber et al., 2009; Van Inghelandt et al., 2012; Lipka et al., 2013), rice (Oryza sativa L.) (Lestari et al., 2011; Yonemaru et al., 2012, 2014; Shao et al., 2013; Wu et al., 2020), and soybean (Choi et al., 2007; Li et al., 2009; Langewisch et al., 2014; Patil et al., 2016; Jiang et al., 2019). In this study, some polymorphic sites were identified and utilized by conducting a haplotype analysis of 21 soybean GmCOL gene families. Based on LD analysis and polymorphic sites, tagging haplotypes composed of some SNPs and InDels indicated that there were 7...
polymorphic sites in GmCOL2 and just 3 polymorphic sites in GmCOL10 and 46 and 10 polymorphic sites in GmCOL2 and GmCOL10, respectively (Supplementary Table 4). In addition, some haplotypes associated with flowering date and maturity (such as GmCOL2-Hap1, Hap2, and Hap4; GmCOL9-Hap1, Hap2, Hap3, and Hap4; GmCOL13-Hap1, Hap2, and Hap3; GmCOL15-Hap1, Hap2, Hap3, and Hap4; GmCOL16-Hap1, Hap2, Hap3, Hap4, and Hap5; and GmCOL25-Hap1, Hap2, Hap3, Hap4, and Hap5) were distributed in varieties with different maturity groups. The LD differences among the 21 soybean COL gene families may reflect selection pressure and, to some extent, the process of natural selection and domestication.

CONCLUSION

In summary, 21 GmCOL genes exhibited natural divergence in association with flowering and growth periods. Significant changes were found in 21 COL genome sequences, among which GmCOL1, GmCOL3, GmCOL8, GmCOL9, GmCOL10, and

FIGURE 4 | Haplotypes of GmCOL genes in relation to flowering time (2015–2016). The relationship between flowering time and haplotype of GmCOL2 (A), GmCOL5 (B), GmCOL9 (C), GmCOL13 (D), GmCOL15 (E), GmCOL16 (F), GmCOL25 (G), and GmCOL28 (H).
GmCOL13 were conserved, but the sequences of the remaining COL genes showed a wide range of changes that could alter their function. In this study, it was noticed that additional polymorphisms linked to the 21 GmCOL gene families in trait-controlling regions might have a significant impact on flowering time and maturity.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/SUPPLEMENTARY MATERIAL.

AUTHOR CONTRIBUTIONS

MA conducted the major phenotypic and genotypic works. SZ, RE, WS, TW, and BJ joined the experiments. SY, CW, WH, SS, and YF provided the materials and technical supports. MA, RE, and BJ prepared the initial draft of the manuscript. FC rewrote the manuscript. FC, BJ, MA, RE, and TH revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Key R&D Program of China (2021YFF1001203 and 2017YFD0101400), China Agriculture Research System (CARS-04), and CAAS Agricultural Science and Technology Innovation Project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.817544/full#supplementary-material

REFERENCES

Alonso-Blanco, C., Mendez-Vigo, B., and Koornneef, M. (2005). From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. Int. J. Dev. Biol. 49, 717–732. doi: 10.1387/ijdb.051994ca
An, H., Roussot, C., Suarez-Lopez, P., Corbesier, L., Vincent, C., Pineiro, M., et al. (2004). CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis. Development 131, 3615–3626. doi: 10.1242/dev.01231

Balasubramanian, S., Sureshkumar, S., Agrawal, M., Michael, T. P., Wessinger, C., Malof, J. N., et al. (2006). The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. Nat. Genet. 38, 711–715. doi: 10.1038/ng1818
Bao, S., Hua, C., Huang, G., Cheng, P., Gong, X., Shen, L., et al. (2019). Molecular Basis of Natural Variation in Photoperiodic Flowering Responses. Dev. Cell 50, 90–101e103. doi: 10.1016/j.devcel.2019.05.018
Bu, T., Lu, S., Wang, K., Dong, L., Li, S., Xie, Q., et al. (2021). A critical role of the soybean evening complex in the control of photoperiod sensitivity and flowering time in the main haplotypes of GmCOL6 in 2015. (B4) Flowering time (VE-R1) of the main haplotypes of GmCOL6 in 2016. (C2) Haplotypes of GmCOL8. (C2) Haplotypes of GmCOL13 in different maturity groups. (C3) Flowering time (VE-R1) of the main haplotypes of GmCOL8 in 2015. (C4) Flowering time (VE-R1) of the main haplotypes of GmCOL9 in 2016. (D1) Haplotypes of GmCOL1 (1) and GmCOL9 (2). (D2) Haplotypes of GmCOL9 in different maturity groups. (D3) Flowering time (VE-R1) of the main haplotypes of GmCOL9 in 2015. (D4) Flowering time (VE-R1) of the main haplotypes of GmCOL9 in 2016. (E1) Haplotypes of GmCOL10. (E2) Haplotypes distribution of GmCOL10 in different maturity groups. (E3) Flowering time (VE-R1) of the main haplotypes of GmCOL10 in 2015. (E4) Flowering time (VE-R1) of the main haplotypes of GmCOL10 in 2016. (F1) Haplotypes of GmCOL13. (F2) Haplotypes distribution of GmCOL13 in different maturity groups. (F3) Flowering time (VE-R1) of the main haplotypes of GmCOL13 in 2015. (F4) Flowering time (VE-R1) of the main haplotypes of GmCOL13 in 2016. (G1) Haplotypes of GmCOL14. (G2) Haplotypes distribution of GmCOL14 in different maturity groups. (G3) Flowering time (VE-R1) of the main haplotypes of GmCOL14 in 2015. (G4) Flowering time (VE-R1) of the main haplotypes of GmCOL14 in 2016. (H1) Haplotypes of GmCOL15. (H2) Haplotypes distribution of GmCOL15 in different maturity groups. (H3) Flowering time (VE-R1) of the main haplotypes of GmCOL15 in 2015. (H4) Flowering time (VE-R1) of the main haplotypes of GmCOL15 in 2016. (I1) Haplotypes of GmCOL16 (1), GmCOL16 (2), GmCOL16 (3), and GmCOL16 (4). (I2) Haplotypes distribution of GmCOL16 in different maturity groups. (I3) Flowering time (VE-R1) of the main haplotypes of GmCOL16 in 2015. (I4) Flowering time (VE-R1) of the main haplotypes of GmCOL16 in 2016.

Supplementary Figure 1 | Haplotype analysis, distribution in different maturity groups, and flowering time (VE-R1) in the main haplotypes of approximately eight GmCOL family genes. (A1) Haplotypes of GmCOL19. (A2) Haplotypes distribution of GmCOL19 in different maturity groups. (A3) Flowering time (VE-R1) of the main haplotypes of GmCOL19 in 2015. (A4) Flowering time (VE-R1) of the main haplotypes of GmCOL19 in 2016. (B1) Haplotypes of GmCOL20 (1), GmCOL20 (2), and GmCOL20 (3). (B2) Haplotypes distribution of GmCOL20 in different maturity groups. (B3) Flowering time (VE-R1) of the main haplotypes of GmCOL20 in 2015. (B4) Flowering time (VE-R1) of the main haplotypes of GmCOL20 in 2016. (C1) Haplotypes of GmCOL22 (1) and GmCOL22 (2). (C2) Haplotypes distribution of GmCOL22 in different maturity groups. (C3) Flowering time (VE-R1) of the main haplotypes of GmCOL22 in 2015. (C4) Flowering time (VE-R1) of the main haplotypes of GmCOL22 in 2016. (D1) Haplotypes of GmCOL23 (1) and GmCOL23 (2). (D2) Haplotypes distribution of GmCOL23 in different maturity groups. (D3) Flowering time (VE-R1) of the main haplotypes of GmCOL23 in 2015. (D4) Flowering time (VE-R1) of the main haplotypes of GmCOL23 in 2016. (E1) Haplotypes of GmCOL24. (E2) Haplotypes distribution of GmCOL24 in different maturity groups. (E3) Flowering time (VE-R1) of the main haplotypes of GmCOL24 in 2015. (E4) Flowering time (VE-R1) of the main haplotypes of GmCOL24 in 2016. (F1) Haplotypes of GmCOL25. (F2) Haplotypes distribution of GmCOL25 in different maturity groups. (F3) Flowering time (VE-R1) of the main haplotypes of GmCOL25 in 2015. (F4) Flowering time (VE-R1) of the main haplotypes of GmCOL25 in 2016. (G1) Haplotypes of GmCOL26. (G2) Haplotypes distribution of GmCOL26 in different maturity groups. (G3) Flowering time (VE-R1) of the main haplotypes of GmCOL26 in 2015. (G4) Flowering time (VE-R1) of the main haplotypes of GmCOL26 in 2016. (H1) Haplotypes of GmCOL28 (1) and GmCOL28 (2). (H2) Haplotypes distribution of GmCOL28 in different maturity groups. (H3) Flowering time (VE-R1) of the main haplotypes of GmCOL28 in 2015. (H4) Flowering time (VE-R1) of the main haplotypes of GmCOL28 in 2016.
Soybean CONSTANS Polymorphism Regulates Development

Lazarro, A., Mouriz, A., Pineiro, M., and Jarillo, J. A. (2015). Red Light-Mediated Degradation of CONSTANS by the E3 Ubiquitin Ligase HOS1 Regulates Photoperiodic Flowering in Arabidopsis. Plant Cell 27, 2437–2454. doi: 10.1105/tpc.15.00529

Lazarro, A., Valverde, F., Pineiro, M., and Jarillo, J. A. (2012). The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. Plant Cell 24, 982–999. doi: 10.1105/tpc.110.081885

Lestari, P., Lee, G., Ham, T. H., Refinier, M. O., and Piao, R. (2011). Single nucleotide polymorphisms and haplotype diversity in rice sucrose synthase 3. J. Hered. 102, 735–746. doi: 10.1037/jhered084

Li, C., Li, Y. H., Li, Y., Lu, H., Hong, H., Tian, Y., et al. (2020). A Domestication-Associated Gene GmPBRR5B Regulates the Circadian Clock and Flowering Time in Soybean. Mol. Plant 13, 745–759. doi: 10.1016/j.molp.2020.01.014

Li, J., Wang, X., Song, W., Huang, X., Zhou, J., Zeng, H., et al. (2017). Genetic variation of maturity groups and four E genes in the Chinese soybean mini core collection. PLoS One 12:e0172106. doi: 10.1371/journal.pone.0172106

Li, Y. H., Zhang, C., Gao, Z. S., Smulders, M. J. M., Ma, Z. L., Liu, Z. X., et al. (2009). Development of SNP markers and haplotype analysis of the candidate gene for rhg1, which confers resistance to soybean cyst nematode in soybean. Mol. Breeding 24, 63–76. doi: 10.1089/fpls.2019.00401

Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., et al. (2013). Genome-wide association study and pathway-level analysis of tocopherol levels in maize grain. G3 3, 1287–1299. doi: 10.1534/g3.113.106148

Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K., and Abe, J. (2008). Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. Genetics 180, 995–1007. doi: 10.1534/genetics.108.092742

Liu, L. J., Zhang, Y. C., Li, Q. H., Sang, Y., Mao, J., Lian, H. L., et al. (2008). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. Plant Cell 20, 292–306. doi: 10.1105/tpc.107.057281

Liu, X., Wu, J. A., Ren, H., Qi, Y., Li, C., Mao, J., et al. (2017). Genetic variation of world soybean maturity date and geographic distribution of maturity groups. Breed. Sci. 67, 221–232. doi: 10.1270/jsbbs.16167

Lu, S., Dong, L., Fang, C., Liu, S., Kong, L., Cheng, Q., et al. (2020). Stepwise selection on homeologous PRR genes controlling flowering and maturity during soybean domestication. Nat. Genet. 52, 428–436. doi: 10.1038/s41588-020-0604-7

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. Nat. Genet. 49, 773–779. doi: 10.1038/ng.3819

Patil, G., Do, T., Vuong, T. D., Vaillyidon, B., Lee, J. D., Chaudhary, J., et al. (2016). Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. Sci. Rep. 6:19199. doi: 10.1038/srep19199

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Robson, F., Costa, M. M., Hepworth, S. R., Vizir, I., Pineiro, M., Reeves, P. H., et al. (2001). Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J. 28, 619–631. doi: 10.1046/j.1365-313x.2001.01163.x

Rosas, U., Mei, Y., Xie, Q., Banta, J. A., Zhou, R. W., Seufferheld, G., et al. (2014). Variation in Arabidopsis flowering time associated with cis-regulatory variation in CONSTANS. Nat. Commun. 5:3651. doi: 10.1038/ncomms4651

Samanfar, B., Molnar, S. J., Charette, M., Schoenrock, A., Dehne, F., Golshani, A., et al. (2017). Mapping and identification of a potential candidate gene for a novel maturity locus. E10, in soybean Theor. Appl. Genet. 130, 377–390. doi: 10.1007/s00122-016-2819-7

Sawa, M., Ninsinow, D. A., Kay, S. A., and Imaizumi, T. (2007). A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics 176, 685–696. doi: 10.1534/genetics.107.070821

Shao, G., Tang, S., Chen, M., Wei, X., He, J., Luo, J., et al. (2013). Haplotype network analysis of Chinese soybean cultivars sequencing soybean genomes. PLoS One 9:e94150. doi: 10.1371/journal.pone.0094150

Tang, S., Lan, H., Zhang, Z., Wang, Q., Chen, K., Liu, T., et al. (2020). Nup96 and HOS1 Are Mutually Stabilized and Gate CONSTANS Protein Level, Conferring Long-Day Photoperiodic Flowering Regulation in Arabidopsis. Plant Cell 32, 374–391. doi: 10.1105/tpc.19.00661
Strayer, C., Oyama, T., Schultz, T. F., Raman, R., Somers, D. E., Mas, P., et al. (2000). Cloning of the Arabidopsis clock gene TOCl, an autoregulatory response regulator homolog. *Science* 289, 768–771. doi: 10.1126/science.289.5480.768

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., Coupland, G., et al. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410, 1116–1120. doi: 10.1038/35074138

Sun, J., Wang, M., Zhao, C., Liu, T., Liu, Z., Fan, Y., et al. (2021). GmFULc Is Induced by Short Days in Soybean and May Accelerate Flowering in Transgenic Arabidopsis thaliana. *Int. J. Mol. Sci.* 22:10333. doi: 10.3390/ijms221910333

Takada, S., and Goto, K. (2003). Terminal flower2, an Arabidopsis homolog of heterochromatin protein1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. *Plant Cell* 15, 2856–2865. doi: 10.1105/tpc.016345

Thakare, D., Kumudini, S., and Dinkins, R. D. (2010). Expression of flowering-time genes in soybean E1 near-isogenic lines under short and long day conditions. *Planta* 231, 951–963. doi: 10.1007/s00425-010-1100-6

Valverde, F., Mouradov, A., Sophie, W., Ravenscroft, D., Samach, A., Coupland, G., et al. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003–1006. doi: 10.1126/science.1091761

Van Inghelandt, D., Melchinger, A. E., Martinant, J. P., and Stich, B. (2012). Genome-wide association mapping of flowering time and northern corn leaf blight (Setosphaeria turcica) resistance in a vast commercial maize germplasm set. *BMC Plant Biol.* 12:56. doi: 10.1186/1471-2229-12-56

Wang, F. F., Nan, H. Y., Chen, L. Y., Fang, C., Zhang, H. Y., Su, T., et al. (2019). A new dominant locus, E11, controls early flowering time and maturity in soybean. *Mol. Breeding* 39:70.

Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., et al. (2020). Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* 18, 1869–1881. doi: 10.1111/pbi.13346

Wang, X. B., Liu, Z. X., Yang, C. Y., Xu, R., Lu, W. G., Zhang, L. F., et al. (2016). Stability of growth periods traits for soybean cultivars across multiple locations. *J. Integr. Agr.* 15, 963–972. doi: 10.1006/jiaf.1996.1152-2

Watanabe, S., Hideshima, R., Xia, Z., Tsukobura, Y., Sato, S., Nakamoto, Y., et al. (2009). Map-based cloning of the gene associated with the soybean maturity locus E1 that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. U S A* 106, E2155–E2164. doi: 10.1073/pnas.1117982109

Yonemaru, J., Ebana, K., and Yano, M. (2014). HapRIce, an SNP haplotype database and a web tool for rice. *Plant Cell Physiol.* 55, 96. doi: 10.1093/pcp/pcv188

Yonemaru, J., Yamamoto, T., Ebana, K., Yamamoto, E., Nagasaki, H., Shibaya, T., et al. (2012). Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. *PLoS One* 7:e32982. doi: 10.1371/journal.pone.0032982

Yue, Y., Liu, N., Jiang, B., Li, M., Wang, H., Jiang, Z., et al. (2017). A Single Nucleotide Deletion in J Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaptation of Tropic Soybean. *Mol. Plant* 10, 656–658. doi: 10.1016/j.molp.2016.12.004

Yue, Y. L., Sun, S., Li, J. W., Yu, H. D., Wu, H. X., Sun, B. Q., et al. (2021). GmFULA improves soybean yield by enhancing carbon assimilation without altering flowering time or maturity. *Plant Cell Rep.* 40, 1875–1888. doi: 10.1007/s00299-021-02752-y

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Awal Khan, Zhang, Emon, Chen, Song, Wu, Yuan, Wu, Hou, Sun, Fu, Jiang and Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.