JIPB Journal of Integrative Plant Biology

Research Article
https://doi.org/10.1111/jipb.13668

GmNF-YC4 delays soybean flowering and maturation by directly repressing *GmFT2a* and *GmFT5a* expression[oo]

Yupeng Cai[1,2†], Li Chen[1,2†], Xiaoqian Liu[1,2†], Weiwei Yao[1,2] and Wensheng Hou[1,2*]

1. State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
2. Ministry of Agriculture Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
†These authors contributed equally to this article.
*Correspondence: Wensheng Hou (houwensheng@caas.cn, Dr. Hou is fully responsible for the distributions of all materials associated with this article)

Yupeng Cai

Wensheng Hou

ABSTRACT

Flowering time and growth period are key agronomic traits which directly affect soybean (*Glycine max* (L.) Merr.) adaptation to diverse latitudes and farming systems. The FLOWERING LOCUS T (FT) homologs GmFT2a and GmFT5a integrate multiple flowering regulation pathways and significantly advance flowering and maturity in soybean. Pinpointing the genes responsible for regulating *GmFT2a* and *GmFT5a* will improve our understanding of the molecular mechanisms governing growth period in soybean. In this study, we identified the Nuclear Factor Y-C (NFY-C) protein GmNF-YC4 as a novel flowering suppressor in soybean under long-day (LD) conditions. GmNF-YC4 delays flowering and maturation by directly repressing the expression of *GmFT2a* and *GmFT5a*. In addition, we found that a strong selective sweep event occurred in the chromosomal region harboring the *GmNF-YC4* gene during soybean domestication. The *GmNF-YC4*^{Hap3} allele was mainly found in wild soybean (*Glycine soja* Siebold & Zucc.) and has been eliminated from *G. max* landraces and improved cultivars, which predominantly contain the *GmNF-YC4*^{Hap1} allele. Furthermore, the *Gmnf-yc4* mutants displayed notably accelerated flowering and maturation under LD conditions. These alleles may prove to be valuable genetic resources for enhancing soybean adaptability to higher latitudes.

Keywords: CRISPR/Cas9, early flowering and maturity, GmNF-YC4, regional adaptability, soybean

Cai, Y., Chen, L., Liu, X., Yao, W., and Hou, W. (2024). GmNF-YC4 delays soybean flowering and maturation by directly repressing *GmFT2a* and *GmFT5a* expression. J. Integr. Plant Biol. **00:** 1–15.

INTRODUCTION

Contemporary crops bear the marks of a lengthy history shaped by both deliberate and unintentional human selection. Examining the related genetic and physiological processes involved in adaptation and domestication offers valuable insights into how to improve crops to meet the continuous demand for enhanced crop yield and quality. The distinctions between cultivated crops and their wild predecessors make up a set of modifications known as domestication syndrome (Doebley et al., 2006). These commonly include changes such as the loss of seed dormancy and dispersal mechanisms, diminished branching, enlarged fruit or seed size, shifts in photoperiod sensitivity, and the commencement of early and more uniform flowering and maturation (Olsen and Wendel, 2013; Lu et al., 2020). Among

these, flowering is the crucial transition from vegetative growth to reproductive growth, representing a critical element in the adaptation of crops to different geographical latitudes and climates during the domestication process.

Soybean (*Glycine max* (L.) Merr.) is a crucial source of protein and oil for humans and animals. The domestication of soybean from its wild relative (*G. soja* Siebold & Zucc.) occurred approximately 6,000–9,000 years ago (Kim et al., 2012). Soybean is a typical short-day (SD) plant, and the photoperiodic regulation of its flowering is an important aspect of its interaction with the environment (Bäurle and Dean, 2006). Plants adjusting to long-day (LD) conditions in higher latitudes need to exhibit early flowering and a decreased or eliminated sensitivity to photoperiod. Elucidating the molecular intricacies that govern the transition from vegetative to reproductive phases in extended daylight conditions could guide the further adaptation of soybean to higher latitudes and also offers priceless valuable molecular targets for customizing the development and cultivation of advanced high-yielding soybean varieties (Yue et al., 2021).

FLOWERING LOCUS T (FT), a crucial regulator of flowering, is remarkably conserved across angiosperm species. Its expression in leaves is induced in response to specific photoperiods, along with various environmental and internal signals. Subsequently, FT relocates to the shoot apical meristem (SAM), where it interacts with the bZIP transcription factor FD to promote the expression of floral identity genes such as *APETALA1* (*AP1*), initiating the development of the floral organs (Abe et al., 2005; Wigge et al., 2005). Ten FT homologs have been discovered in soybean, among which GmFT2a (Glyma.16G150700) and GmFT5a (Glyma.16G044100) are the key promoters of flowering and maturity (Kong et al., 2010; Li et al., 2021). GmFT2a exerts a stronger influence on flowering time under SD conditions, whereas GmFT5a takes on a more substantial role under LD conditions (Cai et al., 2020).

Other proteins have been suggested to control the soybean growth period by directly regulating the expression levels of *GmFT2a* and *GmFT5a*. One previous study demonstrated that GmRAV (Related to ABI3/VP1) suppresses the transcriptional activity of *GmFT5a* by directly engaging with its promoter, which extends both the vegetative and reproductive phases by slowing soybean growth (Wang et al., 2021). GmTOE4b (TARGET OF EAT 4b), an AP2/ERF transcription factor, binds directly to the promoters and gene bodies of *GmFT2a* and *GmFT5a*, exerting a direct inhibitory effect on their transcription (Li et al., 2023), while two SUPPRESSOR OF OVEREXPRESSION OF CONSTANS homologs, GmSOC1a and GmSOC1b directly engage with the regulatory sequences of *GmFT5a* and *GmFT2a* to enhance their transcription levels, specifically in leaves (Kou et al., 2022). GmE1La, a homolog of the legume-specific E1 protein, represses flowering and maturity by directly regulating *GmFT2a* and *GmFT5a* expression (Dong et al., 2023), while GmFUL2a (FRUITFULL 2a) binds to the promoters of *GmFT2a* and *GmFT5a* to enhance their transcription and promote flowering under LD conditions (Dong et al., 2022). The identification of factors that directly regulate *GmFT2a* and *GmFT5a* is therefore crucial for full elucidation of the molecular mechanisms underlying the regulation of the soybean growth period.

The Nuclear Factor Y (NF-Y) transcription factors are known for their high conservation across various eukaryote species and usually constitute a trimeric complex composed of three proteins known as NF-YA, NF-YB, and NF-YC. The individual components of NF-Y are increasingly recognized as key regulators in vital plant processes, such as embryogenesis, drought tolerance, the maintenance of meristems in nitrogen-fixing nodules, and the modulation of photoperiod-dependent flowering time. In *Arabidopsis thaliana*, AtNF-YC3, AtNF-YC4, and AtNF-YC9 are required for CONSTANS (CO)-mediated photoperiod-dependent flowering, with CO relying on NF-Y transcription factor complexes to trigger the activation of *FT* during photoperiod-dependent floral initiation (Kumimoto et al., 2010).

In soybean, 21 *GmN*F-YA genes, 32 *GmN*F-YB genes, and 15 *GmN*F-YC genes have been identified (Quach et al., 2015). Most studies of the *GmN*F-Y gene family have focused primarily on their roles in biotic or abiotic stress; for instance, GmNF-YA (Glyma.02G303800) competes with GmHDA13 for interaction with GmFVE to positively regulate salt tolerance in soybean (Lu et al., 2021a). GmNF-YC14, along with GmNF-YA16 and GmNF-YB2, activates GmPYR1-mediated abscisic acid (ABA) signaling, enhancing soybean stress tolerance. The *Gmnf-yc14* mutants generated using clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) were more drought-sensitive than wild-type (WT) plants. Field trials showed that overexpressing *GmNF-YC14* or *GmPYR1* increased yield, grain quality, and stem size, as well as improving soybean adaptation to drought (Yu et al., 2021). *GmNF-YA3* is regulated by the microRNA miR169 and enhances plant tolerance of drought stress (Ni et al., 2013). Similarly, overexpression of *GmNF-YA5* improved drought tolerance in soybean via both ABA-dependent and ABA-independent pathways. Overexpression of *GmNF-YC4-1* (Glyma.06g169600) enhanced soybean resistance or reduced its susceptibility to viruses, bacteria, fungi, aphids, and soybean cyst nematodes, and the overexpression lines produced seeds containing more protein while maintaining healthy growth (Qi et al., 2019). Transgenic soybean plants overexpressing *GmNF-YC4-2* (Glyma.04g196200) exhibited phenotypes similar to those overexpressing *GmNF-YC4-1*, except that the *GmNF-YC4-2*-overexpressing plants also showed early pod filling and senescence (O'Conner et al., 2021). The NF-Y transcription factors are also involved in soybean nitrogen nutrition (Xu et al., 2021; Zhuang et al., 2021; Li et al., 2022) and lipid accumulation (Yang et al., 2019; Lu et al., 2021b). Furthermore, current research into the regulation of flowering time by the *GmN*F-Y gene family in soybean is predominantly confined to heterologous expression studies in Arabidopsis. Expressing *GmNF-YB1* in transgenic Arabidopsis lines induced flowering by upregulating the expression of flowering-related genes. Additionally,

these lines exhibited longer hypocotyls and fewer rosette leaves than the WT or *Atnf-yb1* mutants under both LD and SD conditions (Mallano et al., 2021). Notably, there is a lack of evidence elucidating the molecular mechanisms through which the *GmNF*-Y gene family is involved in the regulation of flowering time in soybean.

In this study, we utilized a yeast one-hybrid (Y1H) screening method to identify a transcription factor that controls flowering time and latitude adaptation in soybean. Our results indicated that GmNF-YC4 (Glyma.06G169600) represses flowering and maturity by directly regulating the transcription of *GmFT2a* and *GmFT5a*. Furthermore, we determined that *GmNF-YC4* underwent strong selection during soybean domestication, with one *GmNF-YC4* allele predominantly found in wild soybeans and another predominantly found in landraces and improved cultivars. The *GmNF-YC4* allele found in cultivated soybeans is associated with early flowering and the allele in wild soybeans is associated with late flowering. These results offer a glimpse into the molecular mechanisms that regulate flowering time in soybeans, providing valuable information for soybean breeding strategies and regional adaptation initiatives.

RESULTS

Identification of transcription factors directly binding to the *GmFT2a* or *GmFT5a* promoter

In pursuit of novel proteins with direct affinity for the *GmFT2a* or *GmFT5a* promoters, we performed a Y1H library screening, deploying partial promoter sequences of the *GmFT2a* and *GmFT5a* promoters as baits. We identified 45 and 41 putative interaction partners for the *GmFT2a* and *GmFT5a* promoters, respectively (Tables S1, S2). Subsequently, we selected five transcription factors from each candidate set for further validation through Y1H experiments. GmNF-YC4 (Glyma.06G169600) and GmNF-YC15 (Glyma.20G232400) directly bound to the *GmFT2a* promoter in yeast, whereas GmPIF4 (Glyma.02G282100), GmBHLH133 (Glyma.07G083500), and GmBHLH137 (Glyma.06G165700) did not (Figure S1A). Furthermore, GmNF-YC4, GmNF-YB6 (Glyma.05G183200), and GmNF-YB12 (Glyma.08G141000) demonstrated direct affinity with the *GmFT5a* promoter, whereas GmTCP21 (Glyma.10G285900) and GmNF-YB24 (Glyma.15G118800) did not (Figure S1B). Because GmNF-YC4 directly bound to the promoters of both *GmFT2a* and *GmFT5a*, we selected it for more in-depth analysis.

GmNF-YC4 suppresses flowering and maturation in soybean under LD conditions

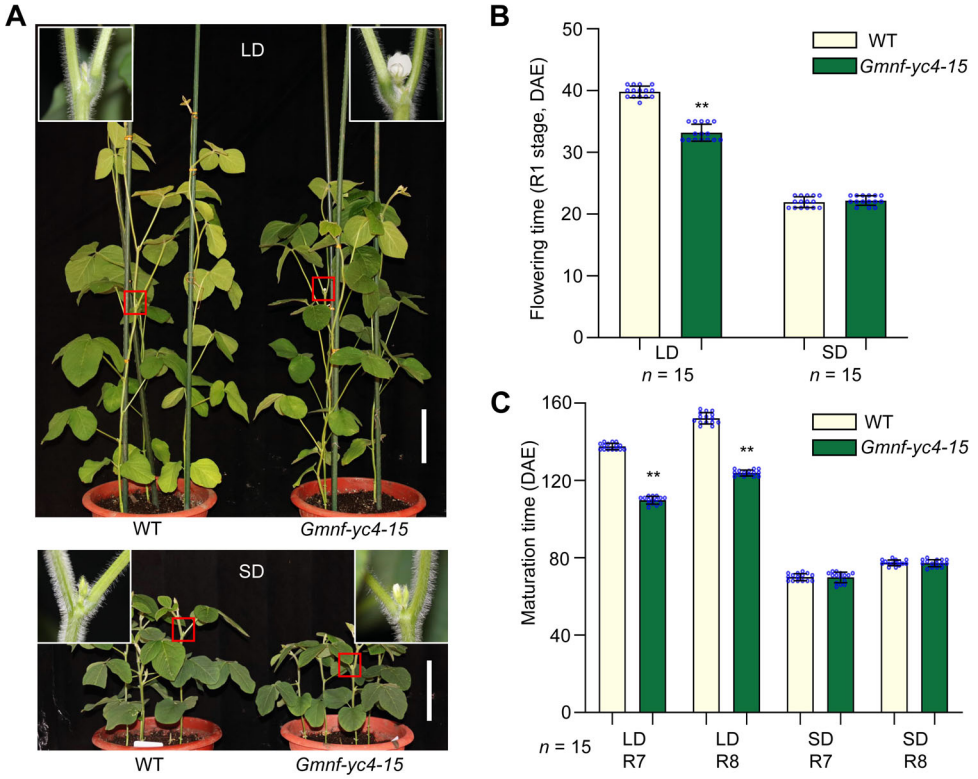
To test the effect of GmNF-YC4 on flowering, we created homozygous *Gmnf-yc4* mutants using a CRISPR/Cas9 system. We observed two categories of mutations (1- and 5-bp deletions), which we denote as *Gmnf-yc4-15* and *Gmnf-yc4-18*, respectively (Figure S2A). The introduction of mutations at the target sites resulted in the premature termination of translation and truncated proteins (Figure S2B). Furthermore, compared with the WT plants, the *Gmnf-yc4-15* and *Gmnf-yc4-18* mutants had significantly reduced expression levels of *GmNF-YC4* (Figure S2C). Transgene-free homozygous *Gmnf-yc4* mutants were identified in the progeny of the mutant lines and were used for subsequent experiments.

To test whether GmNF-YC4 is involved in the regulation of photoperiodic flowering, we grew WT plants and *Gmnf-yc4* mutants under LD (16 h light/30°C and 8 h dark/22°C) and SD (12 h light/30°C and 12 h dark/22°C) photoperiodic conditions. Under SD conditions, the flowering time (R1 stage; one flower at any node) of the *Gmnf-yc4* mutants was almost the same as for the WT plants (22.2 ± 0.8 DAE (d after emergence) for *Gmnf-yc4-15* mutants, 22.1 ± 1.3 DAE for *Gmnf-yc4-18* mutants, and 21.9 ± 0.9 DAE for WT). By contrast, under LD conditions, the *Gmnf-yc4* mutants flowered about 7 d earlier than the WT plants (33.2 ± 1.4 DAE for *Gmnf-yc4-15* mutants, 33.4 ± 1.5 DAE for *Gmnf-yc4-18* mutants, and 39.8 ± 0.9 DAE for WT) (Figures 1, S3).

Next, we analyzed the maturity levels of the materials by observing the time at which the plants reached the R7 stage (one normal pod at any node on the main stem has reached its mature color) and R8 stage (95% of pods have reached their mature color). The *Gmnf-yc4* mutants reached R7 notably faster (about 28 d earlier) than the WT plants (109.8 ± 1.9 DAE for *Gmnf-yc4-15* mutants, 110.1 ± 2.7 DAE for *Gmnf-yc4-18* mutants, and 137.6 ± 1.7 DAE for WT). Similarly, the *Gmnf-yc4* mutants reached R8 significantly earlier (also by about 28 d) than the WT plants (123.9 ± 1.5 DAE for *Gmnf-yc4-15* mutants, 123.9 ± 2.3 DAE for *Gmnf-yc4-18* mutants, and 152.2 ± 2.9 DAE for WT). Under SD conditions, no significant differences were observed between the three genotypes in the timing of either the R7 (69.9 ± 2.7 DAE for *Gmnf-yc4-15* mutants, 70.2 ± 2.8 DAE for *Gmnf-yc4-18* mutants, and 70.0 ± 1.9 DAE for WT) or R8 stages (77.2 ± 1.8 DAE for *Gmnf-yc4-15* mutants, 77.4 ± 2.2 DAE for *Gmnf-yc4-18* mutants, and 77.4 ± 1.4 DAE for WT) (Figures 1, S3). These results provide evidence that GmNF-YC4 acts as an inhibitor of flowering and maturation in soybean under LD conditions. Given the consistent phenotypes observed in the *Gmnf-yc4-15* and *Gmnf-yc4-18* mutants, we selected the *Gmnf-yc4-15* mutant for an in-depth study.

GmNF-YC4 directly inhibits the transcription of *GmFT2a* and *GmFT5a*

GmNF-YC4 is a NF-Y subunit C protein localized to the nucleus (Figure S4A). We quantified the expression levels of *GmNF-YC4* across diverse tissues of soybean, including the root, stem, leaf, flower, and pod, by reverse transcription quantitative polymerase chain reaction (RT-qPCR). *GmNF-YC4* was highly expressed in the leaf, root, and pod tissues, with its highest levels detected in the leaf (Figure S4B). We assessed the transcript levels of *GmNF-YC4* in the leaves of the Jack soybean cultivar over 48-h diurnal cycles of both LD and SD conditions, revealing slightly higher expression levels in SD than LD conditions. In addition, *GmNF-YC4* reached peak expression 4 h after dawn under both LD and SD conditions. By contrast, the



D

Plant	LD-R1 (DAE)	LD-R7 (DAE)	LD-R8 (DAE)	SD-R1 (DAE)	SD-R7 (DAE)	SD-R8 (DAE)
WT	39.8±0.9	137.6±1.7	152.2±2.9	21.9±0.9	70.0±1.9	77.4±1.4
Gmnf-yc4-15	33.2±1.4**	109.8±1.9**	123.9±1.5**	22.2±0.8	69.9±2.7	77.2±1.8

Figure 1. *Gmnf-yc4* mutants showed significant early flowering and maturity phenotype under long-day (LD) conditions
(A) Flowering phenotypes of the wild-type (WT) plants and *Gmnf-yc4-15* mutants under LD (16 h light/8 h dark) and short-day (SD; 12 h light/12 h dark) conditions. Red box, magnified view. Scale bars, 10 cm. **(B–D)** Under LD conditions, the flowering and maturation periods of the *Gmnf-yc4-15* mutants were significantly earlier than the WT plants, while under SD conditions, there was no significant difference. R1 stage, one flower at any node. R7 stage, one normal pod at any node on the main stem has reached its mature color. R8 stage, 95% of the pods have reached their mature color. DAE, d after emergence. Flowering and maturation time of plants are shown as means ± *SD*. *n*, exact numbers of individual plants identified. Asterisks represent statistically significant differences as determined by Student's *t*-test (**$P < 0.01$).

expression level of *GmNF-YC4* in the *Gmnf-yc4-15* mutants remained consistently low (Figure 2A, B).

To investigate whether GmNF-YC4 regulates the growth period by influencing *GmFT2a* and *GmFT5a* expression, we used RT-qPCR to analyze the expression levels of these flowering induction genes in the leaves of WT plants and *Gmnf-yc4-15* mutants in 48-h diurnal cycles under both LD and SD conditions. Both *GmFT2a* and *GmFT5a* were significantly upregulated in the *Gmnf-yc4-15* mutants in comparison with the WT under LD conditions, especially during the peak expression at Zeitgeber time 4 (ZT4) (Figure 2C, E). Under SD conditions, there was no significant difference in the expression levels of *GmFT2a* and *GmFT5a* between the WT and *Gmnf-yc4-15* (Figure 2D, F). These findings indicate that the main mechanism by which GmNF-YC4 controls the growth period in soybean under LD conditions is by downregulating the expression of *GmFT2a* and *GmFT5a*.

We next conducted a series of *in vitro* and *in vivo* experiments to investigate whether *GmFT2a* and *GmFT5a* are direct target genes of GmNF-YC4. In our Y1H assay, GmNF-YC4 demonstrated the capability to directly bind to the promoters of *GmFT2a* and *GmFT5a*. The specific subfragments of the *GmFT2a* and *GmFT5a* promoters, namely $GmFT2a_{pro}$-B (−4,645 to −2,980 bp) and $GmFT5a_{pro}$-C (−1,230 to −42 bp), were adequate to facilitate this interaction in the yeast cells (Figure 3A). We then performed an electrophoretic mobility shift assay (EMSA) using Trigger Factor (TF), a prokaryotic ribosome-associated chaperone protein, as a negative control. His-TF-GmNF-YC4, but not His-TF protein, could bind to the promoters of *GmFT2a* and *GmFT5a*. Moreover, with the addition of

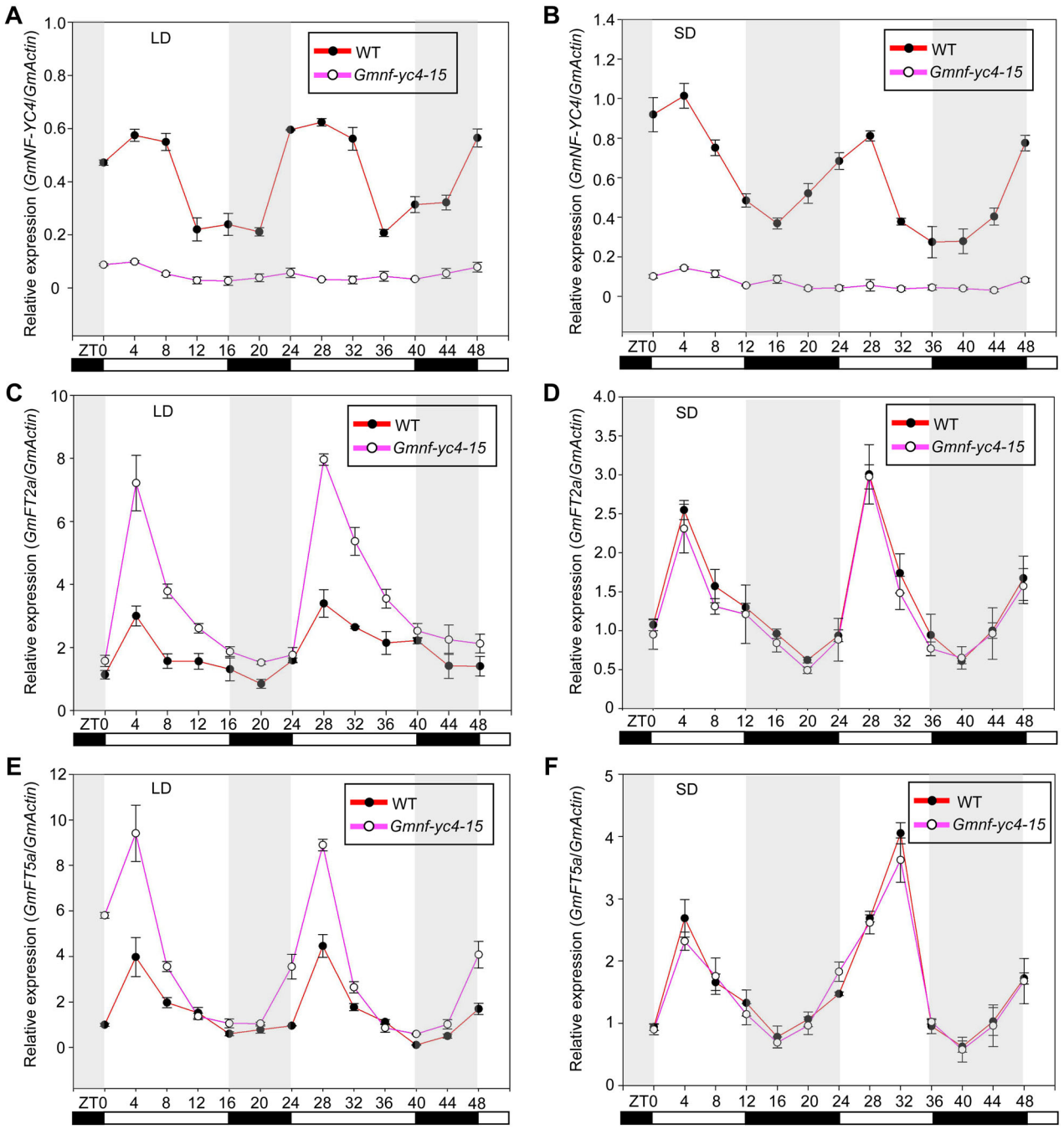


Figure 2. GmNF-YC4 suppresses the transcription levels of *GmFT2a* and *GmFT5a* under long-day (LD) conditions

(A, B) Diurnal expression of *GmNF-YC4* in the wild-type (WT) plants and *Gmnf-yc4-15* mutants under LD (16 h light/8 h dark) and short-day (SD) (12 h light/12 h dark) conditions. All data are given as mean ± *SD* ($n = 3$ plants). Trifoliate leaves were sampled every 4 h at 20 DAE (d after emergence). ZT, Zeitgeber time. **(C, D)** Diurnal expression of *GmFT2a* in the WT plants and *Gmnf-yc4-15* mutants under LD and SD conditions. All data are given as mean ± *SD* ($n = 3$ plants). Trifoliate leaves were sampled every 4 h at 20 DAE. ZT, Zeitgeber time. **(E, F)** Diurnal expression of *GmFT5a* in the WT plants and *Gmnf-yc4-15* mutants under LD and SD conditions. All data are given as mean ± *SD* ($n = 3$ plants). Trifoliate leaves were sampled every 4 h at 20 DAE.

corresponding competitive probes, the brightness of the shifted bands gradually decreased (Figure 3B). These EMSA results indicate that GmNF-YC4 can directly bind to the promoters of *GmFT2a* and *GmFT5a in vitro*. Transient dual-luciferase assays in *Nicotiana benthamiana* leaves showed that GmNF-YC4 represses the expression of $GmFT2a_{pro}$:*LUC* and $GmFT5a_{pro}$:*LUC* (Figure 3C–E). Taken together, our results provide strong evidence that GmNF-YC4 can repress the expression of *GmFT2a* and *GmFT5a* by directly binding to their promoters.

Expression of genes involved in flowering regulation within the shoot apex

We previously reported that GmFT2a and GmFT5a play regulatory roles in modulating the expression of genes associated

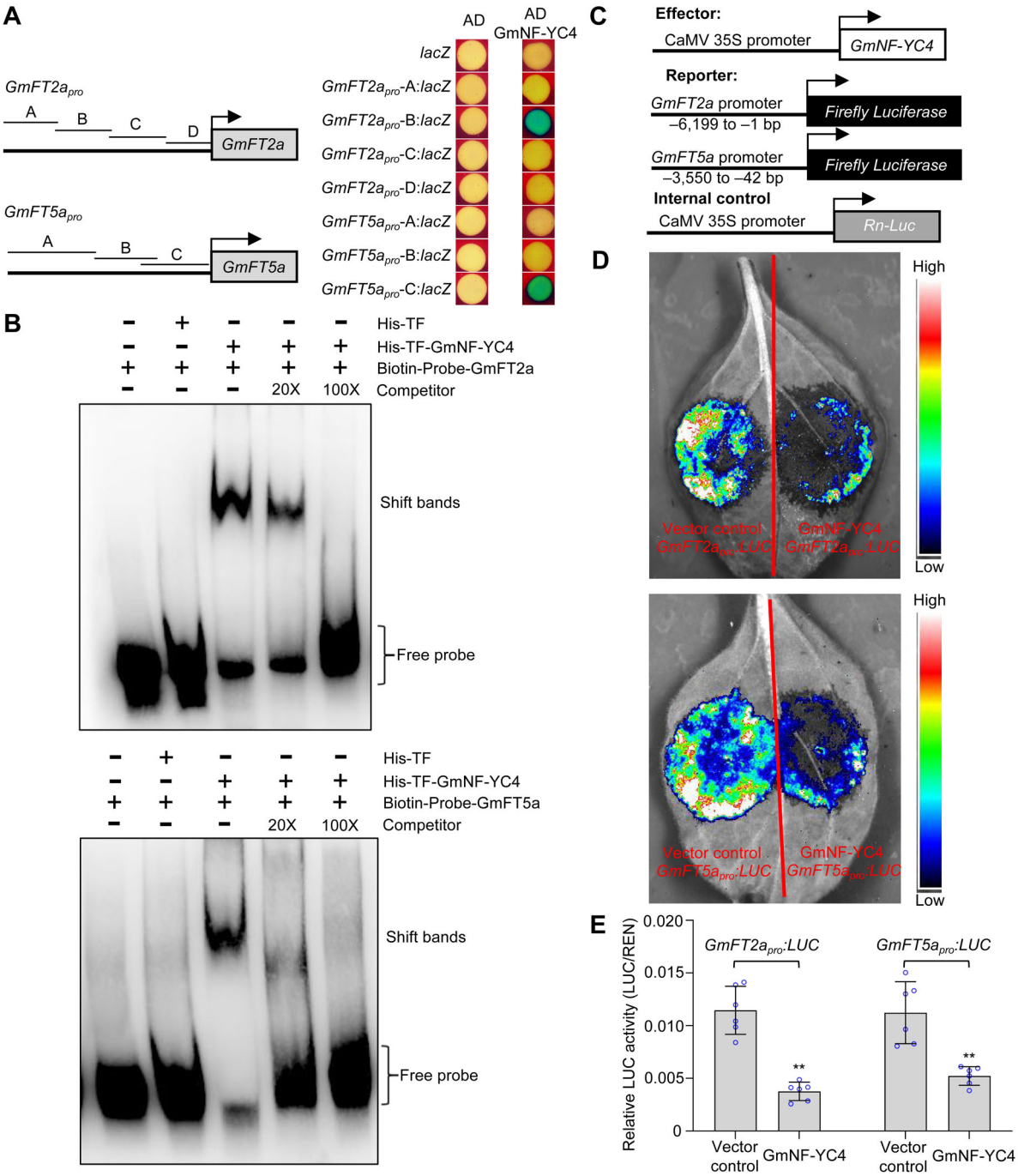


Figure 3. GmNF-YC4 directly inhibits the transcription of *GmFT2a* and *GmFT5a*

(A) Yeast one-hybrid assay showing that GmNF-YC4 bound directly to the promoters of *GmFT2a* and *GmFT5a*. The coding sequence (CDS) of *GmNF-YC4* was cloned into the pB42AD vector. The subfragments of the promoters of *GmFT2a* (*GmFT2a*$_{pro}$-*A*, *B*, *C*, *D*) and *GmFT5a* (*GmFT5a*$_{pro}$-*A*, *B*, *C*) were fused with *lacZ* reporter gene. AD, activation domain. **(B)** Electrophoretic mobility shift assay (EMSA) revealed the direct binding of GmNF-YC4 to the promoters of *GmFT2a* and *GmFT5a in vitro*. Competitor, biotin-unlabeled probes. + and – denote the presence and absence of the corresponding probes or proteins, respectively. 20× and 100× indicate the mole ratios of Competitor versus Biotin-Probe. TF, trigger factor. **(C)** Dual-luciferase assay constructs used to assess *GmFT2a* and *GmFT5a* expression levels regulated by GmNF-YC4. **(D)** Dual-luciferase reporter assay images in *Nicotiana benthamiana* leaves showing that GmNF-YC4 represses the expression of *GmFT2apro:LUC* and *GmFT5apro:LUC*. **(E)** Relative firefly luciferase (LUC) activities are presented as means ± *SD*, $n = 6$. Asterisks above the histogram represent a statistically significant difference determined by Student's *t*-test (**$P < 0.01$).

with flowering within the soybean SAM (Cai et al., 2020). Here, we have demonstrated that GmNF-YC4 can directly bind to the promoters of *GmFT2a* and *GmFT5a*, suppressing their expression. To further explain the molecular mechanisms of GmNF-YC4 in the regulation of flowering, we conducted RT-qPCR experiments to assess the expression levels of several genes involved in flowering within the SAM of WT plants and *Gmnf-yc4* mutants grown under LD conditions. We

focused on several of the flowering-promoting genes that have been characterized in soybean, such as *GmAP1a* (Glyma.16G091300), *GmAP1b* (Glyma.08G269800), *GmAP1c* (Glyma.01G064200), *GmFULa* (Glyma.06G205800), *GmFULb* (Glyma.04G159300), and *GmAGL1* (Glyma.14G027251) (Chi et al., 2011; Nan et al., 2014; Jia et al., 2015; Zeng et al., 2018; Chen et al., 2020). Compared with the WT plants, both *GmFT2a* and *GmFT5a* were significantly upregulated in SAM of the *Gmnf-yc4* mutants (Figure S5). Furthermore, the expression levels of *GmAP1a*, *GmAP1b*, *GmAP1c*, *GmFULa*, *GmFULb*, and *GmAGL1* were all significantly upregulated in the *Gmnf-yc4* mutants. These findings are consistent with the early flowering and maturation phenotype observed in the *Gmnf-yc4* mutants.

Genetic diversity of *GmNF-YC4* in *G. max* and *G. soja*

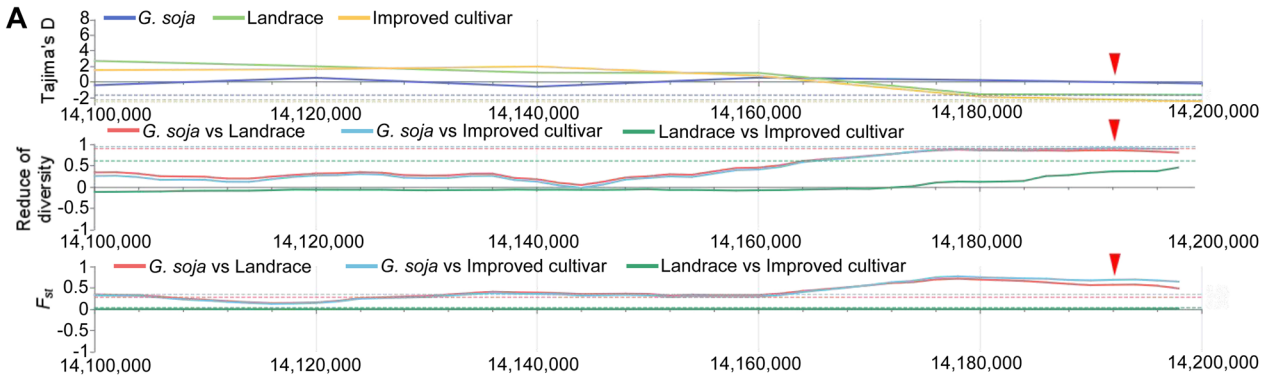
Given the pivotal role of flowering time in soybean domestication (Lu et al., 2020), we next explored whether the *Gm*NF-YC4 gene is linked to the domestication of the *Glycine* genus. We compared the nucleotide diversity across the 100-kb genomic region (Chr06:14,100,000–14,200,000; reference genome Zhonghuang 13) spanning *GmNF-YC4* using previously sequenced accessions and the integrated SoyOmics database (https://ngdc.cncb.ac.cn/soyomics/index) (Liu et al., 2023). Wright's *F* statistic (F_{st}) has been popularly employed in genomic scans looking for recent selection (Wu et al., 2017). We observed notably high F_{st} values between the wild soybean *G. soja* and the landraces (0.57), and between *G. soja* and the improved cultivars (0.69), indicative of significant genetic distinctiveness between these populations (Figure 4A). The reduced diversity between *G. soja* and the landraces (*G. soja* vs. landraces) was 0.86, and that between *G. soja* and the improved cultivars (*G. soja* vs. improved cultivars) was 0.91. Additionally, for the landraces, Tajima's *D* showed a negative value of −1.68, while for the improved cultivars, an even more negative Tajima's *D* value of −2.50 was evident. These results indicate that a strong selective sweep event or linkage with eliminated genes likely occurred in the chromosomal region harboring the *Gm*NF-YC4 gene during the domestication process.

We proceeded to investigate nucleotide polymorphisms within the genomic region (including the 4-kb promoter region) of *GmNF-YC4* across 80 wild soybeans, 851 landraces, and 1,432 improved cultivars. The mapping was conducted against the soybean reference genome Zhonghuang 13 using SoyOmics (https://ngdc.cncb.ac.cn/soyomics/index) (Liu et al., 2023). Single-nucleotide polymorphisms (SNPs) for which less than 10% of lines were missing data and for which the minor allele frequency (MAF) was >5% were included in the analysis. We identified a total of five haplotypes, four of which were present in both *G. soja* and the landraces, and all five of which were detected in the improved cultivars (Figure 4B). In the wild soybeans, Hap3 (77.5%) and Hap4 (15%) were the primary haplotypes, with Hap3 having the highest frequency (Figure 4C). By contrast, the frequencies of Hap3 in the landraces and improved cultivars were 0.2% and 0.35%, respectively, while those of Hap4 were 2.4% and 0.35% in these groups, respectively. For both landraces and improved cultivars, the prevailing haplotypes were Hap1 (the most frequent, at 80.7% and 78.4%, respectively) and Hap2 (16.7% and 20.5%, respectively) (Figure 4C). These findings indicate that *GmNF-YC4* experienced robust selection throughout the domestication process, and that the *GmNF-YC4*^{Hap1} and *GmNF-YC4*^{Hap2} alleles were enriched in the landraces and improved cultivars.

We analyzed the association between the *GmNF-YC4* haplotypes and flowering time (first bloom date) within the 1,238-accession panel in SoyOmics, and found that the *GmNF-YC4*^{Hap1} and *GmNF-YC4*^{Hap2} haplotypes may confer early flowering (Figure 4D). Subsequently, we scrutinized the geographical distribution of the key *GmNF-YC4* alleles within the subset of Chinese accessions (including 38 wild soybeans, 514 landraces, and 1,075 improved cultivars) in the panel described above. We also found that *GmNF-YC4*^{Hap3} was the main allele in *G. soja* but has been almost eliminated from the landraces and improved soybean cultivars, while the *GmNF-YC4*^{Hap1} allele has accumulated and become the predominant allele in the landraces and improved cultivars, followed by *GmNF-YC4*^{Hap2}. *GmNF-YC4*^{Hap2} is mainly found in the northeastern areas of China (Figure 5).

We conducted further research into the molecular mechanisms that determine how the haplotypes of *GmNF-YC4* influence flowering time. The most significant difference in the coding region of the *GmNF-YC4* haplotypes was a change from CCC to GGG (Chr06:14,191,787–14,191,789), resulting in the conversion of proline (Pro) to glycine (Gly) (Figure 4B). Above, we demonstrated that GmNF-YC4 can directly bind to the promoters of *GmFT2a* and *GmFT5a* to inhibit their expression (Figure 3). To investigate whether the change from CCC to GGG would affect the ability of GmNF-YC4 to regulate *GmFT2a* and *GmFT5a*, we conducted dual-luciferase experiments in *N. benthamiana* leaves. Although GmNF-YC4 haplotypes with GGG at Chr06:14,191,787–14,191,789 still had a suppressive effect on the expression levels of *GmFT2a* and GmFT5a, this effect was significantly weaker than that of GmNF-YC4 haplotypes with CCC at Chr06:14,191,787–14,191,789 (Figure 6A–C).

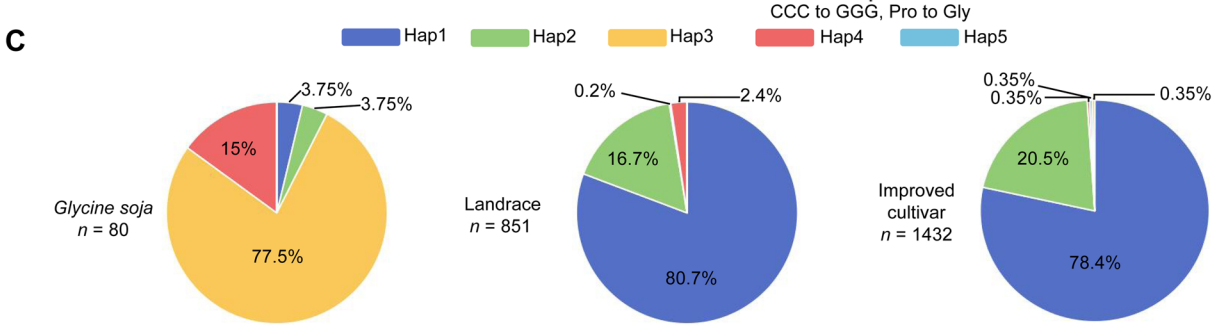
The primary difference between the predominant haplotype in *G. soja* (*GmNF-YC4*^{Hap3}) and the predominant haplotype in *G. max* (*GmNF-YC4*^{Hap1}) lies in their promoter regions (Figure 4B); hence, we cloned these two promoters separately, designating them as *GmNF-YC4*$_{Pro}^{Hap3}$ and *GmNF-YC4*$_{Pro}^{Hap1}$. We examined the expression activities of the two promoters using dual-luciferase experiments, revealing that *GmNF-YC4*$_{Pro}^{Hap1}$ had markedly lower expression than *GmNF-YC4*$_{Pro}^{Hap3}$ (Figure 6D–F). These findings indicate that, in *G. max*, alterations in the *GmNF-YC4* promoter or coding sequence are likely to diminish its inhibitory impact on *GmFT2a* and *GmFT5a*, resulting in earlier flowering than is observed in *G. soja*. These *GmNF-YC4* haplotypes were probably favored and preserved during soybean domestication.



B

Hap_ID	14,187,488	14,188,133	14,188,210	14,188,590	14,188,602	14,188,729	14,188,820	14,191,787	14,191,788	14,191,789	*G. soja*	Landrace	Improved cultivar
Hap1	A	T	C	T	G	T	A	C	C	C	3	687	1123
Hap2	G	C	T	C	A	G	G	G	G	G	3	142	294
Hap3	G	C	C	C	G	T	G	C	C	C	62	2	5
Hap4	G	C	C	C	G	T	A	C	C	C	12	20	5
Hap5	A	T	C	T	G	T	A	G	G	G	0	0	5

CCC to GGG, Pro to Gly



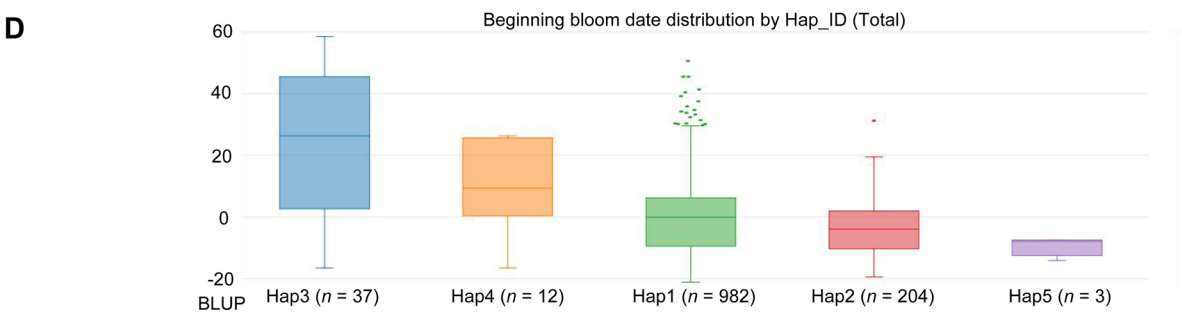


Figure 4. Genetic diversity of *GmNF-YC4* in *Glycine max* and *Glycine soja*

(A) Genetic diversity of *GmNF-YC4* in *G. soja* (wild soybeans) and *G. max* (landraces and improved cultivars). The *Tajima's D*, F_{st} values and nucleic acid diversity were evaluated. The red triangles indicate the location of *GmNF-YC4*. **(B)** Polymorphisms and haplotypes of *GmNF-YC4* in *Glycine soja* and *Glycine max*. The upper inset displays the promoter–exon–intron arrangement of *GmNF-YC4* along with the positions of polymorphisms. Blue arrow represents the 4-kb promoter region of *GmNF-YC4*. Green and gray bars represent the coding sequence (CDS) and 3' untranslated region, respectively. Horizontal line indicates the intron region. Single-nucleotide polymorphisms (SNPs) in the 4-kb promoter and full-length genomic region of *GmNF-YC4* were filtered for minor allele frequency >5%, missing rate <10%, excluding nonfunctional mutations, and removing low association signal SNPs. This yielded 10 high-quality SNPs, classifying accessions into five major haplotypes. **(C)** Proportion of *GmNF-YC4* haplotypes in 80 *G. soja*, 851 landraces and 1,432 improved cultivars. **(D)** Flowering time associated with the five haplotypes at *GmNF-YC4*. The *Y*-axis values represent BLUP (best linear unbiased prediction) values for flowering phenotypes across multiple years and locations, without incorporating estimated intercept terms.

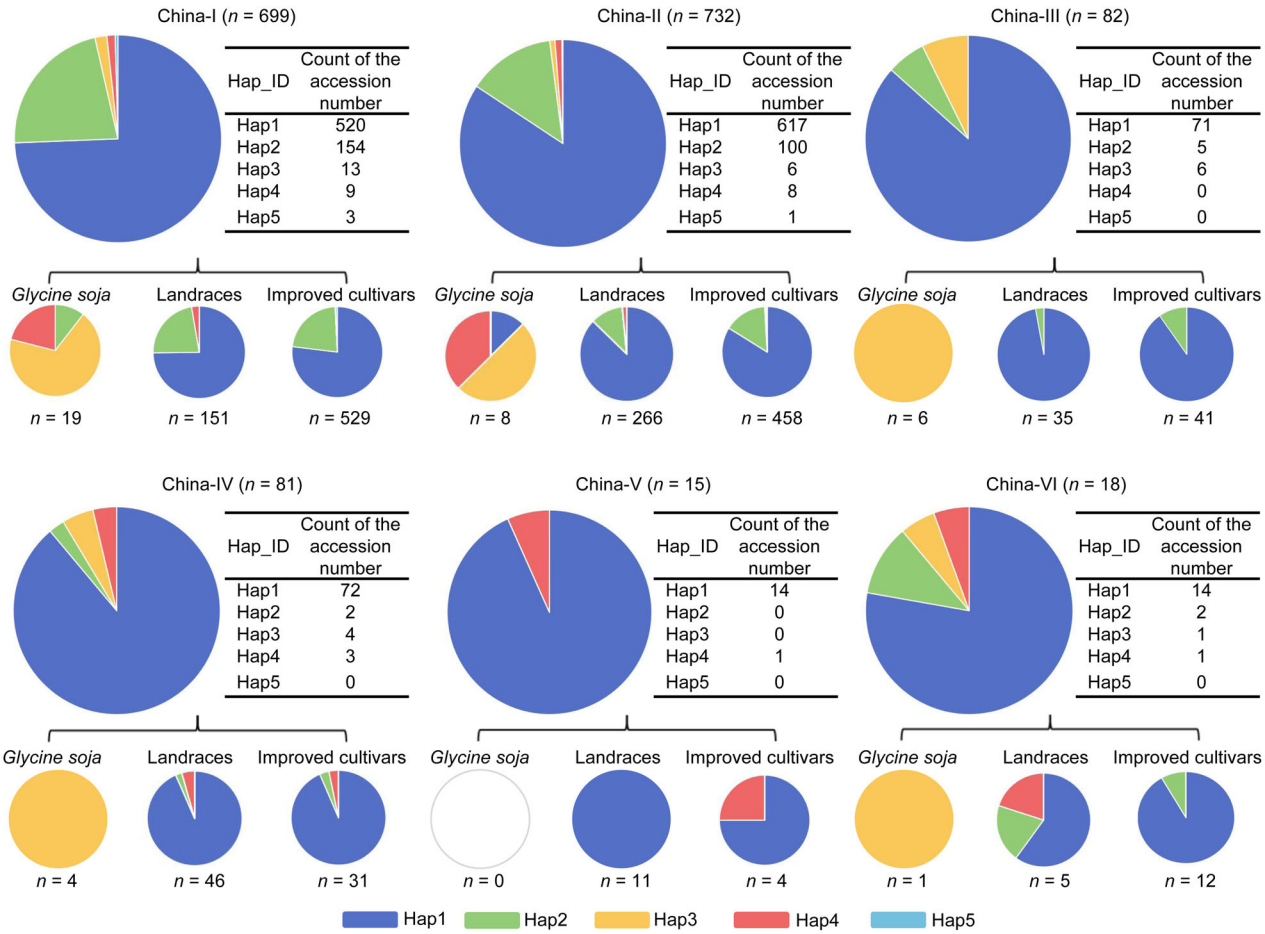


Figure 5. Geographical distribution of *GmNF-YC4* haplotypes in China

The *GmNF-YC4* haplotypes in 38 wild soybeans, 514 landraces, and 1,075 improved cultivars across China were analyzed. China-I–China-VI, eco-regions of soybeans in China defined in a previous study (Liu et al., 2020). The magnitude of each pie chart reflects the number of included accessions in a proportional manner. *n*, count of the accession number of the *GmNF-YC4* haplotypes in each geographic region.

DISCUSSION

Understanding soybean flowering regulation pathways is essential for guiding planting and cultivar selection to optimize the yield and quality of this important crop, which is crucial for sustainable agriculture and food security. In this study, we identified a novel flowering suppressor in soybeans under LD conditions: GmNF-YC4. We showed that GmNF-YC4 inhibits flowering and maturation by directly suppressing the expression of the key flowering genes *GmFT2a* and *GmFT5a*. Additionally, we detected a significant evolutionary event, a selective sweep, in the chromosomal region containing the *Gm*NF-YC4 gene during soybean domestication. The *GmNF-YC4*^{Hap3} variant is prevalent in wild soybean (*G. soja*) but has been phased out in cultivated varieties, whereas the *GmNF-YC4*^{Hap1} variant has become dominant in cultivated soybeans (*G. max*). Furthermore, *Gmnf-yc4* mutants exhibited accelerated flowering and maturation under LD conditions, suggesting their potential utility in breeding to enhance soybean adaptability to higher latitudes.

The *NF-Y* gene family members are involved in flowering regulation through pathways such as photoperiod, aging pathway, gibberellic acid signaling, and vernalization in various plant species, including *Arabidopsis thaliana* (Luo et al., 2018), *N. benthamiana* (Hackenberg et al., 2012), wheat (*Triticum aestivum* L.) (Li et al., 2011), rice (*Oryza sativa* L.) (Hwang et al., 2016), *Chrysanthemum morifolium* (Wei et al., 2017), and Chinese pine (*Pinus tabuliformis*) (Liu et al., 2022). Previous reports have mainly focused on the participation of *GmNF-Y* family genes in the soybean responses to biotic and abiotic stresses (Qi et al., 2019; O'Conner et al., 2021; Yu et al., 2021). Although soybean *Gm*NF-Y genes have been heterologously expressed in Arabidopsis to elucidate their role in the regulation of flowering, there is a lack of reported research utilizing soybean endogenous gene mutants for this purpose. In the present study, we used CRISPR/Cas9 technology to create homozygous *Gmnf-yc4* mutants. Phenotypic evaluations of flowering time and maturity were performed on both WT and *Gmnf-yc4* mutants under LD and SD conditions. We found that the *Gmnf-yc4* mutants displayed markedly earlier flowering and maturity under LD conditions, with no significant difference from WT in the phenotype observed under SD conditions (Figures 1, S3). The *Gmnf-yc4* mutants

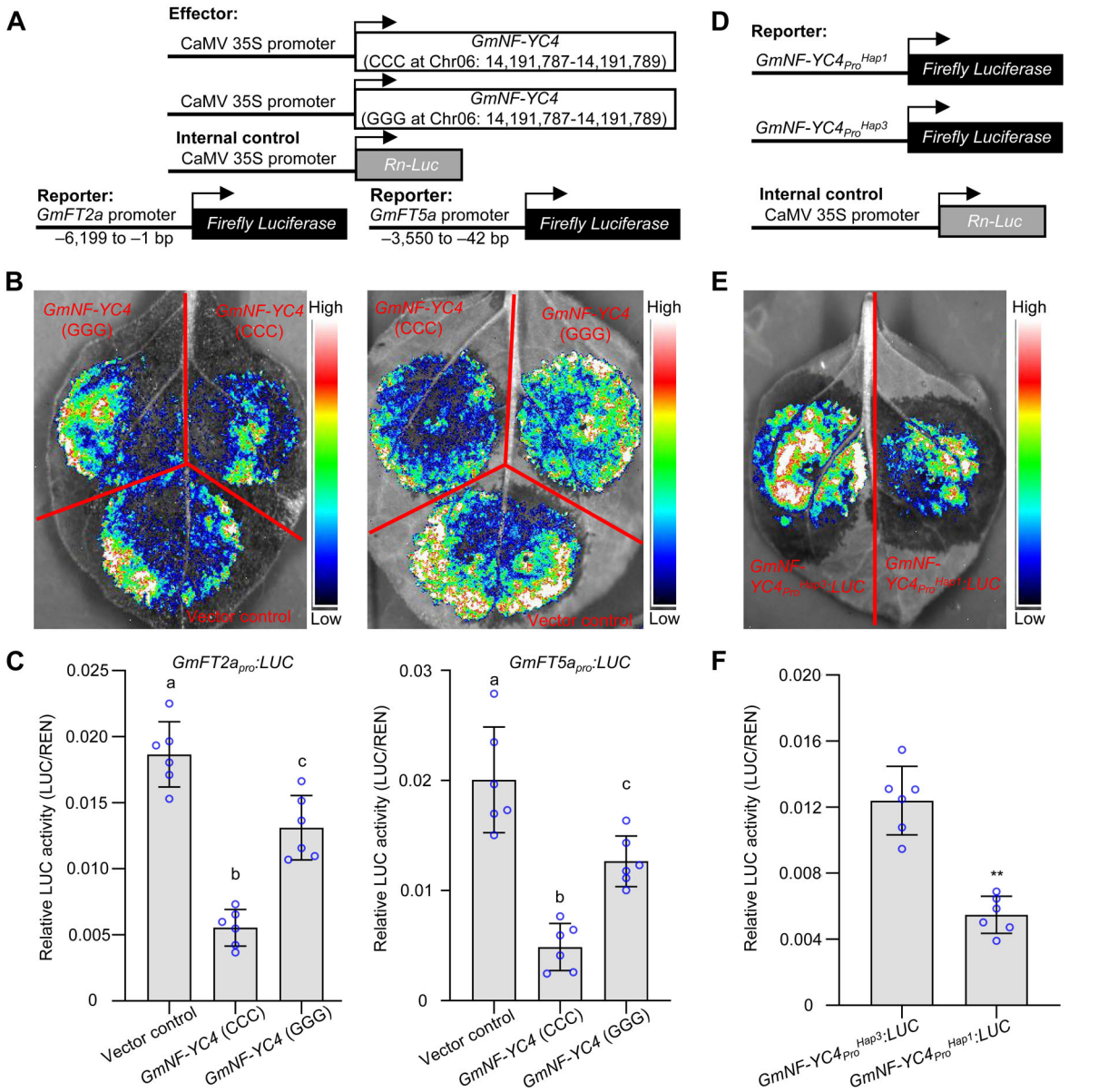


Figure 6. Molecular mechanisms that determine how haplotypes of *GmNF-YC4* influence flowering time

(A) Dual-luciferase assay (DLR) constructs used to assess *GmFT2a* and *GmFT5a* expression levels regulated by two *GmNF-YC4* haplotypes (GGG or CCC at Chr06: 14,191,787-14,191,789). **(B)** DLR assay images in *Nicotiana benthamiana* leaves showing how the change from CCC to GGG (Chr06: 14,191,787–14,191,789) in *GmNF-YC4* affects *GmFT2a* and *GmFT5a* expression. **(C)** Relative firefly luciferase (LUC) activities are presented as means ± *SD*, $n = 6$. Lowercase letters above the histogram represent statistically significant differences ($P < 0.01$) determined by one-way analysis of variance with Tukey's *post hoc* analysis. **(D)** Constructs used for detecting the expression activities of two promoters, $GmNF\text{-}YC4_{Pro}^{Hap3}$ and $GmNF\text{-}YC4_{Pro}^{Hap1}$. **(E)** Fluorescent images showing expression activities of the two promoters. **(F)** Relative LUC activities are presented as means ± *SD*, $n = 6$. Asterisks above the histogram represent a statistically significant difference determined by Student's *t*-test (**$P < 0.01$).

provide new materials and perspectives for studying the regulation of soybean flowering by *GmNF-YC4*. Moreover, the molecular mechanisms by which *GmNF-Y* family genes participate in regulating soybean flowering time and maturity have not previously been elucidated. Here, we utilized Y1H, EMSA, and dual-luciferase methods to confirm that GmNF-YC4 binds directly to the *GmFT2a* and *GmFT5a* promoters, leading to the suppression of their expression (Figure 3). This modulation subsequently impacts the expression of flowering-related genes, including *GmAP1a/b/c*, *GmFULa/b*, and *GmAGL1*, thereby facilitating flowering (Figure 7).

Current research into the genetic networks regulating soybean flowering predominantly emphasizes the crucial roles of *GmFT2a* and *GmFT5a* as central key integrators of various pathways regulating flowering (Lin et al., 2021). While the indirect relationships between upstream regulatory genes and *GmFT2a* or *GmFT5a* have been extensively studied, certain upstream regulators, such as GmRAV, GmTOE4b,

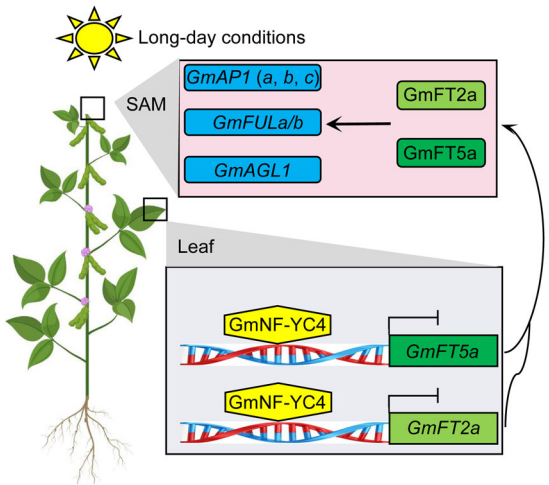


Figure 7. Model summarizing the mechanism of GmNF-YC4 action under long-day (LD) conditions

GmNF-YC4 suppressed flowering and maturation by directly binding to the promoters of *GmFT2a* and *GmFT5a* to suppress their transcription under LD conditions. The mutation in GmNF-YC4 led to the release of expression levels for *GmFT2a* and *GmFT5a*, subsequently upregulating the expression of flowering-promoting genes, including *GmAP1* (*a*, *b*, *c*), *GmFULa/b* and *GmAGL1*, thereby facilitating the process of flowering.

GmE1La, GmFUL2a, GmSOC1a, and GmSOC1b, have also been shown to directly modulate their expression (Wang et al., 2021; Dong et al., 2022, 2023; Kou et al., 2022; Li et al., 2023). Despite these insights, there are still many gaps in our understanding of the direct regulation of the soybean growth period by *GmFT2a* and *GmFT5a*. In future studies, we will explore whether GmNF-YC4 interacts with the six reported activators or repressors of *GmFT2a* and *GmFT5a*, and whether they collectively regulate the expression of these genes, enabling further elucidation of the molecular mechanisms underlying flowering in soybean.

Studying crop domestication—the process by which wild species are transformed into cultivated crops through artificial selection—provides a distinct opportunity to investigate the genetic mechanisms that drive the evolution of adaptive traits. Throughout domestication and diversification, the rigorous selection pressures imposed by humans resulted in rapid, targeted modifications across a range of agriculturally significant traits. Soybean (*G. max*) was domesticated from its wild progenitor (*G. soja*) around 6,000–9,000 years ago (Kim et al., 2012) in a process that involved changes to several key traits, including seed coat color, pod shattering, grain size, plant architecture, flowering time, and rate of maturity (Liu et al., 2020; Lu et al., 2020; Lyu et al., 2023). Flowering, a significant phase marking the shift from vegetative to reproductive growth, holds crucial importance in determining plant adaptation and productivity, and is greatly influenced by photoperiod. Thus, selecting plants adapted to specific latitudinal photoperiods emerged as a pivotal focus in the domestication and diversification of soybean (Lu et al., 2020).

In this study, we found that a strong selective sweep event or linkage with eliminated genes likely occurred in the chromosomal region harboring the *GmNF-YC4* gene during the domestication process. We therefore examined nucleotide variations in the vicinity of *GmNF-YC4*, including its 4-kb promoter region. We identified five haplotypes, of which four were present in *G. soja* and the landraces, and all five were detected in the improved cultivars (Figure 4B). The *GmNF-YC4*^{Hap3} allele was predominant in *G. soja*, but has been eliminated from the landraces and improved cultivars, while the *GmNF-YC4*^{Hap1} allele has accumulated to become the predominant allele in *G. max*. We further investigated how the different haplotypes of *GmNF-YC4* affect flowering time. The key difference in the coding region of these haplotypes is a three-nucleotide change (CCC to GGG, Chr06:14,191,787–14,191,789), resulting in the conversion of Pro to Gly. We demonstrated that GmNF-YC4 can directly bind to the promoters of *GmFT2a* and *GmFT5a* to inhibit their expression, and dual-luciferase experiments conducted in *N. benthamiana* leaves revealed that although GmNF-YC4 (GGG) still suppressed *GmFT2a* and *GmFT5a* expression, the inhibitory effects were reduced compared with GmNF-YC4 (CCC) (Figure 6A–C). Additionally, we found that the main difference between the predominant haplotypes in *G. soja* (*GmNF-YC4*^{Hap3}) and *G. max* (*GmNF-YC4*^{Hap1}) lies in their promoter regions. Cloning and testing these promoters separately revealed that *GmNF-YC4-Pro*^{Hap1} had lower expression activity than *GmNF-YC4-Pro*^{Hap3} (Figure 6D–F). These findings suggest that changes in the *GmNF-YC4* promoter or coding sequence in *G. max* may reduce its inhibition of *GmFT2a* and *GmFT5a*, resulting in earlier flowering compared with *G. soja*. The weak alleles of *GmNF-YC4* identified in this study, which are associated with the early flowering and maturity phenotype, offer valuable genetic resources for enhancing soybean adaptation to, and productivity in, high-latitude environments.

MATERIALS AND METHODS

Plant materials and growth conditions

The CRISPR/Cas9 system was used to mutate the soybean endogenous gene *GmNF-YC4*. A target site within the first exon of *GmNF-YC4* was identified using the CRISPR-P web tool (Lei et al., 2014). An oligonucleotide pair corresponding to the single-guide RNA (sgRNA) was synthesized by Tsingke (Beijing) and then annealed to create dimers. Subsequently, the dimerized oligonucleotide was incorporated into the CRISPR/Cas9 expression vector as previously described (Cai et al., 2018). Following this, the vector was introduced into *Agrobacterium tumefaciens* strain GV3101 via electroporation and used to transform the soybean variety Jack, as previously described (Chen et al., 2018).

The plants were grown in growth chambers under either a LD (16 h light/30°C and 8 h dark/22°C) or a SD (12 h light/30°C and 12 h dark/22°C) photoperiod. The lighting conditions were

characterized by a red:blue quantum ratio of 5.03:1 and a red:far-red quantum ratio of 3.26:1.

Subcellular localization

Agrobacterium tumefaciens GV3101 strain harboring the plant binary construct Gateway-YFP-GmNF-YC4 was infiltrated into the fully expanded young leaves of *N. benthamiana*. The plants were then cultivated under a 16 h light/8 h dark photoperiod for 3 d. The yellow fluorescent protein (YFP) fluorescence was visualized using a laser scanning confocal microscope (LSM510 Meta; Carl Zeiss, Oberkochen, Germany).

Quantitative real-time PCR

Total RNA was extracted from the various tissues using the FastPure Universal Plant Total RNA Isolation Kit (Vazyme, Nanjing, China). A 1 μg aliquot of total RNA served as the template to synthesize first-strand complementary DNA (cDNA), which was performed using the HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) from Vazyme. The RT-qPCR was conducted using the ChamQ SYBR qPCR Master Mix (Low ROX Premixed; Vazyme) and the Applied Biosystems QuantStudio 7 Flex Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer's guidelines. The $2^{-\Delta\Delta C_t}$ method was employed to determine the relative expression levels of the target genes, which were normalized against the expression of *GmActin* serving as the internal control.

Yeast one-hybrid library screening and validation assay

Total RNA was extracted from a composite sample of leaves and shoot tips of the soybean variety Jack after 21 d of cultivation under SD or LD conditions. The resulting RNA was used to construct a Y1H cDNA library, which was performed by Ngene (Guangzhou, China). The promoter fragments of *GmFT2a* (−3,208 to −3,109 bp) and *GmFT5a* (−241 to −42 bp) were cloned into the pAbAi vector (Takara Bio, Kusatsu, Japan) between the *Hin*dIII and *Sal*I sites and then integrated into the genome of the Y1HGold yeast strain as the baits. Subsequently, the cDNA library was screened using the Matchmaker Gold Yeast One-Hybrid Library Screening System (Takara Bio), in accordance with the manufacturer's instructions. Individual yeast colonies were chosen and characterized through DNA sequencing.

The full-length coding sequence (CDS) fragments of *GmNF-YC4*, *GmNF-YC15*, *GmPIF4*, *GmbHLH133*, *GmbHLH137*, *GmTCP21*, *GmNF-YB24*, *GmNF-YB6*, and *GmNF-YB12* were amplified by PCR and cloned into the *Eco*RI/*Xho*I sites of the pB42AD vector (Takara Bio). The promoter fragments of *GmFT2a* (−6,199 to −4,600 bp, −4,645 to −2,980 bp, −3,017 to −1,335 bp, −1,359 to −1 bp, and −6,199 to −1 bp) and *GmFT5a* (−3,550 to −2,019 bp, −2,042 to −956 bp, −1,230 to −42 bp, and −3,550 to −42 bp) were amplified using PCR and cloned into the *Kpn*I/*Xho*I sites of the pLacZ2U vector (Takara Bio). In the Y1H assay, a variety of pB42AD and pLacZ2u recombination constructs were introduced simultaneously into the yeast strain EGY48. These transformed cells were cultivated on a minimal synthetic defined medium supplemented with a dropout mix lacking uracil and tryptophan (−Ura/−Trp). Additionally, the medium was supplemented with X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) for blue color development. This blue coloration indicates successful interactions between the introduced constructs in the yeast cells.

Purification of recombinant proteins

The complete CDS of *GmNF-YC4* was amplified via PCR and subsequently inserted into the *Kpn*I/*Bam*HI sites of the pCold-TF vector. After the integrity of the resulting vectors was confirmed through sequencing, they were introduced into *Escherichia coli* strain BL21. To induce protein expression, the bacterial culture containing the *His-TF-GmNF-YC4* construct was treated with 0.8 mmol/L isopropyl β-D-thiogalactoside (IPTG) and incubated overnight at 16°C. Following the incubation, the bacterial cells were harvested by centrifugation at 3,783×*g* for 5 min at 4°C. The TF-tagged protein was purified using Ni-NTA agarose (Qiagen, Hilden, Germany). Simultaneously, employing the same methodology, the His-TF protein was purified and designated as the control group for subsequent experimental purposes.

Electrophoretic mobility shift assay

The 79-bp subfragment (−3,207 to −3,129 bp upstream of ATG) of the *GmFT2a* promoter and 119-bp subfragment (−201 to −83 bp upstream of ATG) of the *GmFT5a* promoter were used as probes. Biotin labeling of the double-stranded probes was achieved using the EMSA Probe Biotin Labeling Kit (Beyotime Biotech, Shanghai, China). The EMSA was conducted as previously described (Cai et al., 2023), following the guidelines provided in the LightShift EMSA Optimization and Control Kit (Thermo Fisher Scientific). In brief, the mixture containing 2 μg protein and 50 ng of the corresponding probe were mixed with 20 μL binding buffer that comprised 1× binding buffer, 0.05 μg/μL of Poly (dI-dC), 2.5% (v/v) glycerol, 5 mmol/L $MgCl_2$, and 0.05% (v/v) Nonidet P-40 and then incubated at 25°C for 60 min. Subsequently, the DNA–protein complex was separated on native polyacrylamide gels containing 5% (w/v) acrylamide and 2.5% (v/v) glycerol. To prevent the gel from overheating, the electrophoresis was conducted in an ice-water bath. Following the electrophoresis, the probes were transferred onto Hybond N+ nylon membranes (Millipore Sigma, Burlington, MA, USA) using an electroblotting technique within an ice-water bath. The transferred complexes were detected in accordance with the protocols outlined in the LightShift Chemiluminescent EMSA Kit (Thermo Fisher Scientific).

Transient dual-luciferase assay

The CDS of *GmNF-YC4* was amplified using PCR, then integrated into the pGreenII 62-SK vector at the *Xba*I/*Kpn*I sites, generating the *pGreenII 62-SK-GmNF-YC4* construct.

To produce the reporter constructs, the promoters of *GmFT2a* (−6,199 to −1 bp), *GmFT5a* (−3,550 to −42 bp), and *GmNF-YC4* (−4,000 to −1 bp) were amplified and incorporated into the *Kpn*I/*Xho*I sites of the pGreenII 0800-LUC vector (Hellens et al., 2005). Each of the resulting constructs was separately introduced into *A. tumefaciens* strain GV3101. The transformed Agrobacterium cells were subsequently infiltrated into fully expanded young leaves of *N. benthamiana* using a needleless syringe. Post-infiltration, the plants were cultivated for 3 d under LD conditions. Their luminescence activity was assessed using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) to detect the signals from both firefly luciferase (LUC) and *Renilla* luciferase (REN). The ratio of LUC to REN activity was then calculated, providing a measure of *LUC* expression levels.

Haplotype analysis of *GmNF-YC4* in the soybean population

To investigate the genetic diversity of *GmNF-YC4* in *G. max* and *G. soja*, data on the SNPs within the 4-kb promoter region and full-length genomic region of this gene were obtained for 80 wild soybeans, 851 landraces, and 1,432 improved cultivars from the previously reported SNP and Indel dataset (Liu et al., 2023). The SNP data underwent filtering, requiring a MAF >5% and a missing rate <10%. SNPs resulting in nonfunctional mutations or with low association signals were also removed, leaving 10 high-quality SNPs for further analysis. The polymorphisms were used to classify the accessions into five major haplotypes.

Primer sequences

All primers used in this study are listed in Table S3.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (32201881), Central Public-interest Scientific Institution Basal Research Fund (S2022QH02), and China Postdoctoral Science Foundation (2020M672652).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

W.H. and Y.C. conceived and designed the experiments. Y.C., L.C., and X.L. performed the experiments. Y.C. analyzed the haplotype data. Y.C. wrote the manuscript. L.C. and W.Y. assisted in soybean transformation. All authors read and approved the contents of this paper.

Edited by: Fanjiang Kong, Guangzhou University, China

Received Dec. 6, 2023; **Accepted** Apr. 3, 2024

OO: OnlineOpen

REFERENCES

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

Bäurle, I., and Dean, C. (2006). The timing of developmental transitions in plants. Cell **125:** 655–664.

Cai, Y., Chen, L., Liu, X., Guo, C., Sun, S., Wu, C., Jiang, B., Han, T., and Hou, W. (2018). CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. Plant Biotechnol. J. **16:** 176–185.

Cai, Y., Liu, Y., Fan, Y., Li, X., Yang, M., Xu, D., Wang, H., Deng, X.W., and Li, J. (2023). MYB112 connects light and circadian clock signals to promote hypocotyl elongation in Arabidopsis. Plant Cell **35:** 3485–3503.

Cai, Y., Wang, L., Chen, L., Wu, T., Liu, L., Sun, S., Wu, C., Yao, W., Jiang, B., Yuan, S., et al. (2020). Mutagenesis of GmFT2a and GmFT5a mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. Plant Biotechnol. J. **18:** 298–309.

Chen, L., Cai, Y., Liu, X., Yao, W., Guo, C., Sun, S., Wu, C., Jiang, B., Han, T., and Hou, W. (2018). Improvement of soybean Agrobacterium-mediated transformation efficiency by adding glutamine and asparagine into the culture media. Int. J. Mol. Sci. **19:** 3039.

Chen, L., Nan, H., Kong, L., Yue, L., Yang, H., Zhao, Q., Fang, C., Li, H., Cheng, Q., Lu, S., et al. (2020). Soybean AP1 homologs control flowering time and plant height. J. Integr. Plant Biol. **62:** 1868–1879.

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

Doebley, J.F., Gaut, B.S., and Smith, B.D. (2006). The molecular genetics of crop domestication. Cell **127:** 1309–1321.

Dong, L., Cheng, Q., Fang, C., Kong, L., Yang, H., Hou, Z., Li, Y., Nan, H., Zhang, Y., Chen, Q., et al. (2022). Parallel selection of distinct T of 5 alleles drove the adaptation of cultivated and wild soybean to high latitudes. Mol. Plant **15:** 308–321.

Dong, L., Li, S., Wang, L., Su, T., Zhang, C., Bi, Y., Lai, Y., Kong, L., Wang, F., Pei, X., et al. (2023). The genetic basis of high-latitude adaptation in wild soybean. Curr. Biol. **33:** 252–262.e254.

Hackenberg, D., Keetman, U., and Grimm, B. (2012). Homologous NF-YC2 subunit from Arabidopsis and tobacco is activated by photo-oxidative stress and induces flowering. Int. J. Mol. Sci. **13:** 3458–3477.

Hellens, R.P., Allan, A.C., Friel, E.N., Bolitho, K., Grafton, K., Templeton, M.D., Karunairetnam, S., Gleave, A.P., and Laing, W.A. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. Plant Methods **1:** 13.

Hwang, Y.-H., Kim, S.-K., Lee, K.C., Chung, Y.S., Lee, J.H., and Kim, J.-K. (2016). Functional conservation of rice OsNF-YB/YC and Arabidopsis AtNF-YB/YC proteins in the regulation of flowering time. Plant Cell Rep. **35:** 857–865.

Jia, Z., Jiang, B., Gao, X., Yue, Y., Fei, Z., Sun, H., Wu, C., Sun, S., Hou, W., and Han, T. (2015). GmFULa, a FRUITFULL homolog, functions in the flowering and maturation of soybean. Plant Cell Rep. **34:** 121–132.

Kim, M.Y., Van, K., Kang, Y.J., Kim, K.H., and Lee, S.H. (2012). Tracing soybean domestication history: From nucleotide to genome. Breed. Sci. **61:** 445–452.

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B.M., Watanabe, S., Yamada, T., Tabata, S., Kanazawa, A., Harada, K., et al. (2010). Two coordinately

regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. Plant Physiol. **154:** 1220–1231.

Kou, K., Yang, H., Li, H., Fang, C., Chen, L., Yue, L., Nan, H., Kong, L., Li, X., Wang, F., et al. (2022). A functionally divergent SOC1 homolog improves soybean yield and latitudinal adaptation. Curr. Biol. **32:** 1728–1742.

Kumimoto, R.W., Zhang, Y., Siefers, N., and Holt, 3rd, B.F. (2010). NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in *Arabidopsis thaliana*. Plant J. **63:** 379–391.

Lei, Y., Lu, L., Liu, H.Y., Li, S., Xing, F., and Chen, L.L. (2014). CRISPR-P: A web tool for synthetic single-guide RNA design of CRISPR-system in plants. Mol. Plant **7:** 1494–1496.

Li, C., Distelfeld, A., Comis, A., and Dubcovsky, J. (2011). Wheat flowering repressor VRN2 and promoter CO_2 compete for interactions with NUCLEAR FACTOR-Y complexes. *Plant J*. **67:** 763–773.

Li, H., Du, H., Huang, Z., He, M., Kong, L., Fang, C., Chen, L., Yang, H., Zhang, Y., Liu, B., et al. (2023). The AP2/ERF transcription factor TOE4b regulates photoperiodic flowering and grain yield per plant in soybean. Plant Biotechnol. J. **21:** 1682–1694.

Li, X., Fang, C., Yang, Y., Lv, T., Su, T., Chen, L., Nan, H., Li, S., Zhao, X., Lu, S., et al. (2021). Overcoming the genetic compensation response of soybean florigens to improve adaptation and yield at low latitudes. Curr. Biol. **31:** 3755–3767.

Li, X., Zhou, H., Cheng, L., Ma, N., Cui, B., Wang, W., Zhong, Y., and Liao, H. (2022). Shoot-to-root translocated GmNN1/FT2a triggers nodulation and regulates soybean nitrogen nutrition. PLoS Biol. **20:** e3001739.

Lin, X., Liu, B., Weller, J.L., Abe, J., and Kong, F. (2021). Molecular mechanisms for the photoperiodic regulation of flowering in soybean. J. Integr. Plant Biol. **63:** 981–994.

Liu, H., Guo, Y., Wang, H., Yang, W., Yang, J., Zhang, J., Liu, D., El-Kassaby, Y.A., and Li, W. (2022). Involvement of PtCOL5-PtNF-YC4 in reproductive cone development and gibberellin signaling in Chinese pine. Plant Sci. **323:** 111383.

Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., Zhou, G.A., Zhang, H., Liu, Z., Shi, M., et al. (2020). Pan-genome of wild and cultivated soybeans. Cell **182:** 162–176.e113.

Liu, Y., Zhang, Y., Liu, X., Shen, Y., Tian, D., Yang, X., Liu, S., Ni, L., Zhang, Z., Song, S., et al. (2023). SoyOmics: A deeply integrated database on soybean multi-omics. Mol. Plant **16:** 794–797.

Lu, L., Wei, W., Li, Q.T., Bian, X.H., Lu, X., Hu, Y., Cheng, T., Wang, Z.Y., Jin, M., Tao, J.J., et al. (2021b). A transcriptional regulatory module controls lipid accumulation in soybean. New Phytol. **231:** 661–678.

Lu, L., Wei, W., Tao, J.J., Lu, X., Bian, X.H., Hu, Y., Cheng, T., Yin, C.C., Zhang, W.K., Chen, S.Y., et al. (2021a). Nuclear factor Y subunit GmNFYA competes with GmHDA13 for interaction with GmFVE to positively regulate salt tolerance in soybean. Plant Biotechnol. J. **19:** 2362–2379.

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 homeologous PRR genes controlling flowering and maturity during soybean domestication. Nat. Genet. **52:** 428–436.

Luo, X., Gao, Z., Wang, Y., Chen, Z., Zhang, W., Huang, J., Yu, H., and He, Y. (2018). The NUCLEAR FACTOR-CONSTANS complex antagonizes Polycomb repression to de-repress FLOWERING LOCUS T expression in response to inductive long days in Arabidopsis. Plant J. **95:** 17–29.

Lyu, X., Li, Y.H., Li, Y., Li, D., Han, C., Hong, H., Tian, Y., Han, L., Liu, B., and Qiu, L.J. (2023). The domestication-associated L1 gene encodes a eucomic acid synthase pleiotropically modulating pod pigmentation and shattering in soybean. Mol. Plant **16:** 1178–1191.

Mallano, A.I., Li, W., Tabys, D., Chao, C., Yang, Y., Anwar, S., Almas, H.I., Nisa, Z.U., and Li, Y. (2021). The soybean GmNFY-B1 transcription factor positively regulates flowering in transgenic Arabidopsis. Mol. Biol. Rep. **48:** 1589–1599.

Nan, H., Cao, D., Zhang, D., Li, Y., Lu, S., Tang, L., Yuan, X., Liu, B., and Kong, F. (2014). GmFT2a and GmFT5a redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor GmFDL19 in soybean. PLoS One **9:** e97669.

Ni, Z., Hu, Z., Jiang, Q., and Zhang, H. (2013). GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress. Plant Mol. Biol. **82:** 113–129.

O'Conner, S., Zheng, W., Qi, M., Kandel, Y., Fuller, R., Whitham, S.A., and Li, L. (2021). GmNF-YC4-2 increases protein, exhibits broad disease resistance and expedites maturity in soybean. Int. J. Mol. Sci. **22:** 3586.

Olsen, K.M., and Wendel, J.F. (2013). A bountiful harvest: Genomic insights into crop domestication phenotypes. Annu. Rev. Plant Biol. **64:** 47–70.

Qi, M., Zheng, W., Zhao, X., Hohenstein, J.D., Kandel, Y., O'Conner, S., Wang, Y., Du, C., Nettleton, D., MacIntosh, G.C., et al. (2019). QQS orphan gene and its interactor NF-YC4 reduce susceptibility to pathogens and pests. Plant Biotechnol. J. **17:** 252–263.

Quach, T.N., Nguyen, H.T., Valliyodan, B., Joshi, T., Xu, D., and Nguyen, H.T. (2015). Genome-wide expression analysis of soybean NF-Y genes reveals potential function in development and drought response. Mol. Genet. Genomics **290:** 1095–1115.

Wang, Y., Xu, C., Sun, J., Dong, L., Li, M., Liu, Y., Wang, J., Zhang, X., Li, D., Sun, J., et al. (2021). GmRAV confers ecological adaptation through photoperiod control of flowering time and maturity in soybean. Plant Physiol. **187:** 361–377.

Wei, Q., Ma, C., Xu, Y., Wang, T., Chen, Y., Lü, J., Zhang, L., Jiang, C.-Z., Hong, B., and Gao, J. (2017). Control of chrysanthemum flowering through integration with an aging pathway. Nat. Commun. **8:** 829.

Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J. U., and Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. Science **309:** 1056–1059.

Wu, F., Sedivy, E.J., Price, W.B., Haider, W., and Hanzawa, Y. (2017). Evolutionary trajectories of duplicated FT homologues and their roles in soybean domestication. Plant J. **90:** 941–953.

Xu, H., Li, Y., Zhang, K., Li, M., Fu, S., Tian, Y., Qin, T., Li, X., Zhong, Y., and Liao, H. (2021). miR169c-NFYA-C-ENOD40 modulates nitrogen inhibitory effects in soybean nodulation. New Phytol. **229:** 3377–3392.

Yang, S., Miao, L., He, J., Zhang, K., Li, Y., and Gai, J. (2019). Dynamic transcriptome changes related to oil accumulation in developing soybean seeds. Int. J. Mol. Sci. **20:** 2202.

Yu, T.F., Liu, Y., Fu, J.D., Ma, J., Fang, Z.W., Chen, J., Zheng, L., Lu, Z. W., Zhou, Y.B., Chen, M., et al. (2021). The NF-Y-PYR module integrates the abscisic acid signal pathway to regulate plant stress tolerance. Plant Biotechnol. J. **19:** 2589–2605.

Yue, L., Li, X., Fang, C., Chen, L., Yang, H., Yang, J., Chen, Z., Nan, H., Chen, L., Zhang, Y., et al. (2021). FT5a interferes with the Dt1-AP1 feedback loop to control flowering time and shoot determinacy in soybean. J. Integr. Plant Biol. **63:** 1004–1020.

Zeng, X., Liu, H., Du, H., Wang, S., Yang, W., Chi, Y., Wang, J., Huang, F., and Yu, D. (2018). Soybean MADS-box gene GmAGL1 promotes flowering via the photoperiod pathway. BMC Genomics **19:** 51.

Zhuang, Q., Xue, Y., Yao, Z., Zhu, S., Liang, C., Liao, H., and Tian, J. (2021). Phosphate starvation responsive GmSPX5 mediates nodule growth through interaction with GmNF-YC4 in soybean (*Glycine max*). Plant J. **108:** 1422–1438.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: http://onlinelibrary.wiley.com/doi/10.1111/jipb.13668/suppinfo

Figure S1. Identify and validate genes interacting with *GmFT2a* or *GmFT5a* promoter using yeast one-hybrid (Y1H) screening

Figure S2. The homozygous *Gmnf-yc4* mutants generated by clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9

Figure S3. Phenotypes of the wild-type (WT) plants and *Gmnf-yc4-18* mutants under long-day (LD) and short-day (SD) conditions

Figure S4. Subcellular localization and tissue expression pattern of *GmNF-YC4*

Figure S5. Expression of genes involved in flowering regulation within the shoot apex under long-day (LD) conditions

Table S1. Genes from yeast one-hybrid (Y1H) screening with *GmFT2a* promoter region

Table S2. Genes from yeast one-hybrid (Y1H) screening with *GmFT5a* promoter region

Table S3. Primer sequences used in the study

Scan using WeChat with your smartphone to view JIPB online

Scan with iPhone or iPad to view JIPB on Twitter