



Article

Characterization and Activity Analyses of the *FLOWERING LOCUS T* Promoter in *Gossypium Hirsutum*

Na Sang ^{1,2}, Darun Cai ³, Chao Li ⁴, Yuqiang Sun ^{5,*} and Xianzhong Huang ^{2,6,*} 

¹ The Key Laboratory of Oasis Eco-Agriculture, College of Agriculture, Shihezi University, Shihezi 832000, China; sn899097@163.com

² Special Plant Genomics Laboratory, College of Life Sciences, Shihezi University, Shihezi 832000, China

³ State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China; cdrhxz104@163.com

⁴ State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome Editing, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China; chaoyue_014@163.com

⁵ Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, College of Life Sciences, Zhejiang Sci-Tech University, Hangzhou 310016, Zhejiang, China

⁶ College of Agriculture, Anhui Science and Technology University, Fengyang 233100, Anhui, China

* Correspondence: sunyuqiang@zstu.edu.cn (Y.S.); xianzhongh106@163.com (X.H.);

Tel.: +152-5888-1990 (Y.S.); +136-7756-8119 (X.H.)

Received: 20 August 2019; Accepted: 25 September 2019; Published: 26 September 2019



Abstract: Flowering transition is a crucial development process in cotton (*Gossypium hirsutum* L.), and the flowering time is closely correlated with the timing of *FLOWERING LOCUS T* (*FT*) expression. However, the mechanism underlying the coordination of various *cis*-regulatory elements in the *FT* promoter of cotton has not been determined. In this study, a 5.9-kb promoter of *FT* was identified from cotton. A bioinformatics analysis showed that multiple insertion–deletion sites existed in the 5.9-kb promoter. Different expression levels of a reporter gene, and the induction by sequential deletions in *GhFT* promoter, demonstrated that 1.8-kb of the *GhFT* promoter was stronger than 4.2-, 4.8-, and 5.9-kb promoter fragments. The binding sites of the CONSTANS (CO) and NUCLEAR FACTOR Y transcription factors were located within the 1.0-kb sequence upstream of the *FT* transcription start site. A large number of repeat segments were identified in proximal promoter regions (−1.1 to −1.4 kb). A complementation analysis of deletion constructs between 1.0 and 1.8 kb of *G. hirsutum*, *Gossypium arboreum*, and *Gossypium raimondii* *FT* promoters revealed that the 1.0-kb fragment significantly rescued the late-flowering phenotype of the *Arabidopsis* *FT* loss-of-function mutant *ft-10*, whereas the 1.8-kb promoter only slightly rescued the late-flowering phenotype. Furthermore, the conserved CORE motif in the cotton *FT* promoter is an atypical TGTG(N2-3)ATG, but the number of arbitrary bases between TGTG and ATG is uncertain. Thus, the proximal *FT* promoter region might play an important role affecting the activity levels of *FT* promoters in cotton flowering.

Keywords: cotton; *FLOWERING LOCUS T*; promoter analysis; flowering transition; CORE

1. Introduction

Flowering in plants is an important physiological process in the switch from vegetative to reproductive growth, and this process is strictly controlled by a complex regulatory network consisting of environmental cues and plant developmental signals [1]. The regulation of the *FLOWERING LOCUS T* (*FT*) gene in *Arabidopsis thaliana* is understood to involve at least six major pathways controlling

flowering: the photoperiod, vernalization, ambient temperature, age, autonomous, and gibberellin pathways. They converge at the integrator *FT* gene [2,3]. *FT* protein is produced in the vascular tissues of leaves and is translocated to the shoot apical meristem, where it can interact with the basic leucine zipper transcription factor *FD* to form a complex and then activate the expression of downstream genes involved in floral morphogenesis, such as *APETALA1* and *LEAFY* [4–8].

The transcriptional activation of *FT* is directly mediated by *CONSTANS* (*CO*) under long-day (LD) inductive conditions [1]. *CORE1* and *CORE2* are the two *CO* responsive elements in the *FT* promoter that share the consensus sequence TGTG(N2-3)ATG [9]. The *CONSTANS*, *CONSTANS*-like, and *TOC1* domain, which is a conserved motif at the C-terminus of *CO*, can interact with nuclear factor Y (*NF-Y*) transcription factors [9–11]. Cao et al. (2014) proposed a photoperiodic flowering recruitment model: *NF-Y* complexes bind to *CCAAT* distal to the *FT* promoter and recruit and stabilize the binding of *CO* to *CORE1* and *CORE2* proximal to the *FT* promoter to facilitate the transcriptional activation of *FT* [12]. Upland cotton (*Gossypium hirsutum* L.) is an allotetraploid species, which was formed by hybridization of the A-genome ancestor *Gossypium arboreum* (AA) with D-genome ancestor *Gossypium raimondii* (DD), followed by chromosomal doubling [13]. Homeologous gene pairs are included in the A and D subgenomes. Our group previously identified 42 *CO*-like (*COL*) homologues in the *G. hirsutum* genome, and a phylogenetic analysis classified them into three groups [14]. Group I contains 14 *GhCOLs*, which are clustered with *AtCO* and rice (*Oryza sativa*) *Hd1* [15]. One homoeologous pair, *GhCOL1-A* and *GhCOL1-D*, in Group I have the greatest sequence similarities with *AtCO* and *Hd1*. Furthermore, the overexpression of *GhCOL1* can fully rescue the late-flowering phenotype of the *co-2* mutant, suggesting that the *GhCOL1* homologues act as flowering inducers in *G. hirsutum* [14].

To clarify the complex relationships between activation and suppression signals that regulate flowering through *FT*, Adrian et al. (2010) studied the conserved sequence of the *FT* promoter in *Arabidopsis* in detail. First, the 7.0-kb *FT* promoter sequences of Columbia (Col), *Landsberg erecta* (Ler), *Arabidopsis lyrata*, *Brassica rapa*, and *Arabis alpine* were compared. Three conserved sequence regions were located upstream, in the middle and downstream of the 7.0-kb *FT* promoter and are known as Blocks A, B, and C, respectively. Further investigations of the *FT* promoter (1.0-, 4.0-, 5.7-, and 8.1-kb segments) showed that Block A, located at the proximal end, and Block C, located at the distal end at approximately 5.0-kb, played important roles in the activity of the *FT* promoter mediated by *CO*. Furthermore, the overexpression of *CO* resulted in changes in the chromatin state, such as decreased LIKE HETEROCHROMATIN PROTEIN1 (*LHP1*) binding and increased H3K9K14 acetylation. These changes could be the results of the upregulation of *FT* expression, rather than a prerequisite for *FT* activation [16]. Furthermore, a number of other studies on the interactions between *cis*-acting elements on the *FT* promoter and corresponding transcription factors have also been conducted in *Arabidopsis* [17–20].

The function of the *FT* promoter has been studied in other plants, such as *FT2a* in soybean (*Glycine max*) [21], *Hd3a* in rice [22], and *FT* in wheat (*Triticum aestivum*) [23]. Our group identified a cotton *FT* homologous gene, *GhFT*, from *G. hirsutum* and demonstrated that it was predominantly expressed in stamens and sepals and had a relatively higher expression level during the initiation stage of fiber development. The ectopic overexpression of *GhFT* promoted flowering, lateral shoot growth, and leaf morphology, indicating that cotton *FT* regulates shoot architecture by advancing determinate growth [24,25]. Similar results have been reported by other groups [26–28]. However, details of the *FT* promoter in cotton, such as its activity, function, and sequence, remain unknown and require further study. Here, we analyzed the activity of the 5.9-kb promoter of *GhFT* using a series of deletion constructs and found that the *GhFT* promoter had an activity length of approximately 1.8-kb upstream of the transcription start site. Further research is warranted to examine the 1.8-kb promoter in *G. hirsutum*, *G. arboreum*, and *G. raimondii*; therefore, we have constructed a truncated 1.0-kb promoter, and the *GUS* activity induced by which was stronger and was more efficient in rescuing the late-flowering phenotype of *ft-10* than the 1.8-kb promoter.

2. Results

2.1. Bioinformatics Analysis of the FT Promoters in Cotton

To identify candidate sequences for regulatory motifs, the sequences from 5.9-kb upstream of *FT* in *G. hirsutum*, *G. arboreum*, and *G. raimondii* were analyzed. The pairwise alignment revealed five Indel sites among them (Figure 1). *Indels E, D, and C* are located approximately 4.9-, 4.2-, and 1.6-kb upstream of the *FT* start codon, respectively, and do not exist in *G. raimondii* (Figure 1A–C). *Indel B* is located approximately 0.9-kb upstream of the *FT* initiation codon and does not exist in *G. arboreum* or the A subgenome promoters of *G. hirsutum*. Furthermore, repeat sequences (RSs) are identified approximately 1.0-kb and 1.2-kb upstream of ATG, but *G. arboreum* and the A subgenomes of *G. hirsutum* lack these RSs (Figure 1D). *Indel A*, closest to the *FT* start codon, have been deleted in the D subgenome promoter of *G. hirsutum* (Figure 1E).

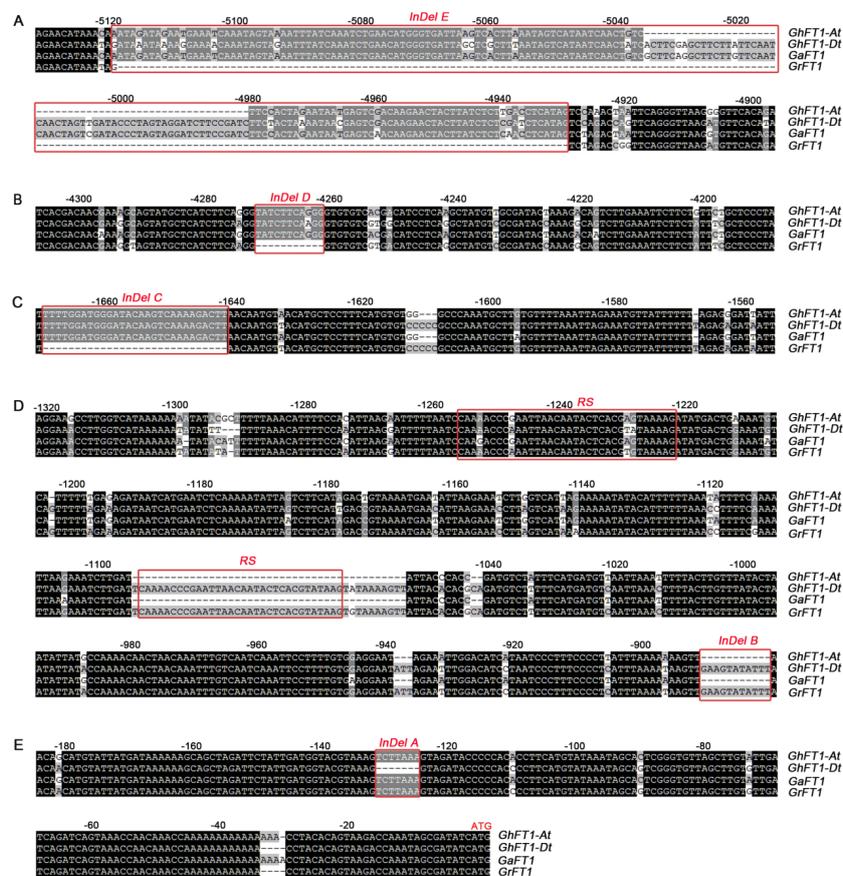


Figure 1. Multiple sequence alignment of the *FT* promoter among *G. hirsutum*, *G. raimondii* and *G. arboreum*. (A) Sequence alignment of the distal *FT* promoter (−4900 to −5120 bp). One *Indel* sequence was identified and called *Indel E* (−4930 to −5120 bp). (B) Sequence alignment of the *FT* promoter (−4190 to −4305 bp). One *Indel* sequence was identified and named *Indel D* (−4260 to −4270 bp). (C) Sequence alignment of the *FT* promoter (−1560 to −1670 bp). One *Indel* sequence was identified and named *Indel C* (−1640 to −1670 bp). (D) Sequence alignment of the *FT* promoter (−4190 to −4305 bp). One *Indel* sequence was identified and named *Indel B* (−879 to −890 bp). One repetitive sequence was also found (−1220 to −1260 bp and −1055 to −1095 bp). (E) Sequence alignment of a proximal *FT* promoter region (−1 to −185 bp). One *Indel* sequence was identified and named *Indel A* (−879 to −890 bp).

2.2. Histochemical Activity Analysis of GhFT Promoter

To assess the activity of the *GhFT* promoter, *GhFT* promoter deletion constructs (1.0-, 1.5-, 1.8-, 4.2-, 4.8-, and 5.9-kb *FT* promoter fragments) were independently fused with GUS (Figure 2A). These

constructs were stably transformed into *A. thaliana* Col-0 plants using *Agrobacterium*-mediated genetic transformation. Histochemical GUS staining was performed in various tissues, such as seedlings, cotyledons, leaves, flowers, and siliques (Figure 2B). In seedlings and cotyledons, the GUS activities were induced by all the promoter fragments, and the 1.0-, 1.5-, and 1.8-kb fragments' activities were higher than those of the 4.2-, 4.8-, and 5.9-kb fragments. The GUS activities induced by the 1.0-, 1.5-, and 1.8-kb fragments were higher in leaves, flowers, and siliques, in which did not induced by the 4.2-, 4.8-, and 5.9-kb fragments. We speculate that the 4.2-, 4.8-, and 5.9-kb promoter fragments may have the ability to induce transcription, but there are some upstream repressing sequences between 1.8- and 5.9-kb, which needs to be confirmed in the future. These results indicated that the activity regulator of the *GhFT* promoter in *G. hirsutum* was approximately 1.8-kb upstream of the *FT* start codon, but it did not exclude the effects of other distal sequences on *GhFT* expression.

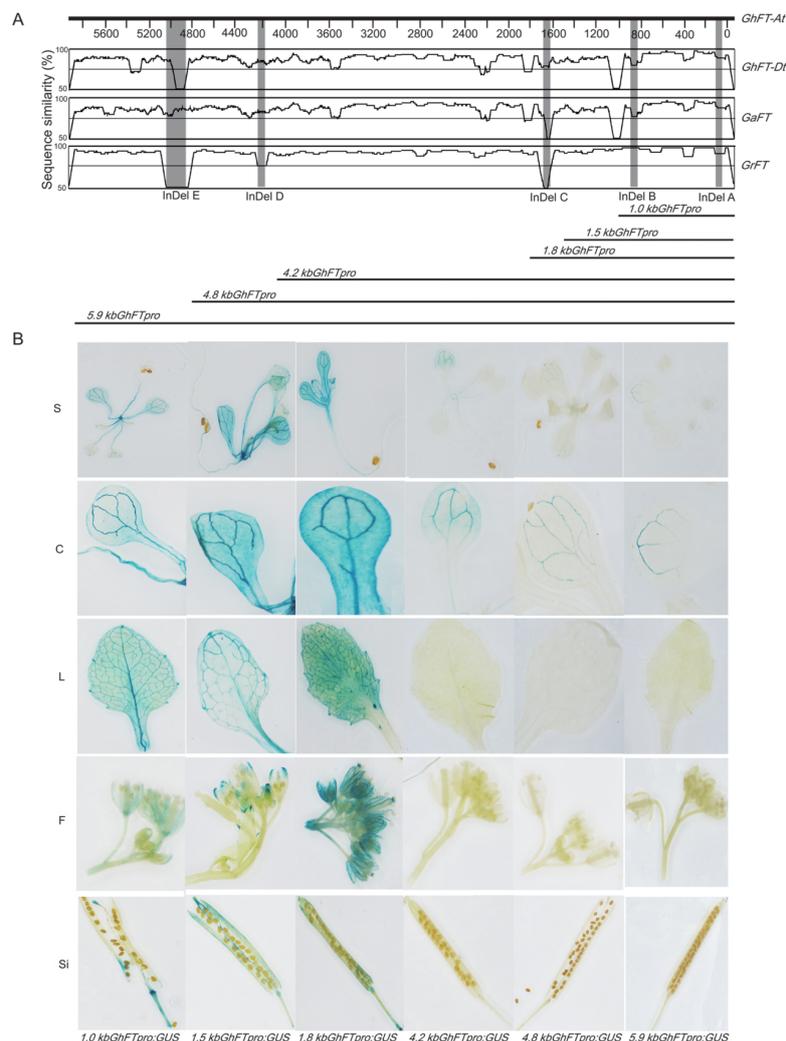


Figure 2. Histochemical activity analyses of the *GhFT* promoter. **(A)** Alignment of the same promoter region of the 5.9-kb *FT* promoter from *G. hirsutum*, *G. arboreum*, and *G. raimondii* using mVISTA. Graphical shows base-pair identity in a range of 50–100%. Light-gray areas showed different Indels. Promoter constructs used for analyses are depicted as black boxes. A transient expression assay was performed using a *GUS* gene under the control of *FT* promoter fragments of 1.0-, 1.5-, 1.8-, 4.2-, 4.8-, and 5.9-kb in length. **(B)** *GUS* activity analysis of the *GhFT* promoter in transgenic plants of different deletions and different tissues. S, C, L, F, and Si represent seedlings, cotyledons, leaves, flowers, and siliques, respectively.

2.3. Sequence Analyses of the 1.8-kb FT Promoters in *G. hirsutum*, *G. arboreum*, and *G. raimondii*

To analyze the 1.8-kb promoter of *FT*, we first performed a phylogenetic analysis of *G. hirsutum*, *G. arboreum*, and *G. raimondii*. The constructed phylogenetic tree indicated that the 1.8-kb cotton *FT* promoters could be clearly divided into two types of the A and D subgenomes on the basis of the source of the promoter genome (Figure 3).

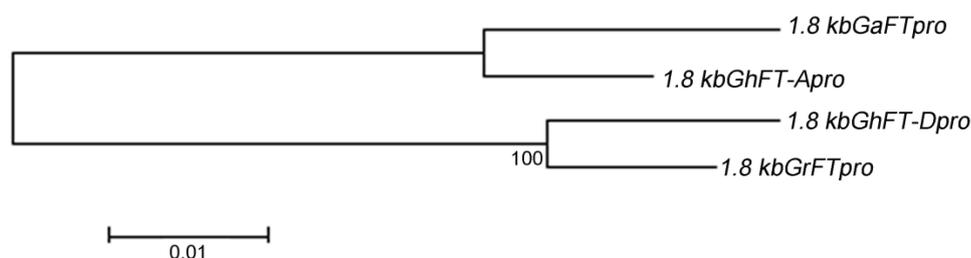


Figure 3. Phylogenetic analyses of the 1.8-kb *FT* promoter of *G. hirsutum*, *G. arboreum*, and *G. raimondii*. The phylogenetic tree was constructed using MEGA5.1. The numbers on the branches represent the reliability percent of bootstrap values based on 1000 replications, and the scale represents nucleotide substitution ratio.

We further amplified the approximately 1.8-kb *GhFT* promoter sequences from the A and D subgenomes of *G. hirsutum*. Sequencing confirmed that the length of the *GhFT-A* promoter was 1771 bp, and the length of the *GhFT-D* promoter was 1701 bp. A bioinformatics analysis by the PlantCare program revealed that these 1.8-kb fragments contain a number of putative plant *cis*-elements. The predicted core promoter elements, such as the TATA-box and CAAT-box, were identified (Tables S1 and S2). In addition, some hormone-related elements were also recognized, including the TCA-element, ABRE, the TATC-box and the P-box. Light-responsive elements, such as the as-2-box, the G-box, the ATCT-motif, ACE, Box1, Box4, Pc-CMA2c, the AT1-motif, the GT1-motif, Sp1 and the TCT-motif, were also observed. There were also abiotic stress-tolerance factors, such as HSE, TC-rich repeats, MBS, and the GC-motif. Among tissue-expression motifs, the CAT-box motif and circadian rhythm-related elements also existed in these two sequences (Tables S1 and S2).

2.4. The 1.0-kb Sequence Upstream of the FT Translation Initiation Site Concentrated CORE and CCAAT Domains

Two CO-responsive elements (CORE1 and CORE2) in the *AtFT* promoter share a consensus sequence, TGTG(N2-3)ATG [9]. CO proteins can interact with NF-Y transcription factors [10,11]. NF-Y proteins consist of three unique subunits, NF-YA, NF-YB, and NF-YC, which form a heterotrimeric complex that binds DNA and can recognize CCAAT *cis*-elements [29–31]. Because our experiment showed that the strong GUS activities were induced by the 1.8-kb cotton *GhFT* promoter fragment (Figure 2), we next analyzed the CORE and CCAAT domains in this 1.8-kb sequence to predict the CO and NF-Y transcription factor binding sites. The CORE and CCAAT domains were concentrated in the 1.0-kb region upstream of the translation start site, but there was no CCAAT domain between 1.0 and 1.8 kb in the promoter (Figure S1A,B), which was different from in *Arabidopsis* [12]. However, compared with the consensus sequence TGTG(N2-3)ATG reported previously, the conserved motif of CORE in the cotton *FT* promoter is the atypical TGTG(N2-3)ATG, and the number of arbitrary bases between TGTG and ATG is uncertain, being one or six (Figure S1A,B). The FT/CO regulon model is highly conserved in photosensitive plants [32–34]. Thus, whether the FT/CO regulatory model is conserved in upland cotton remains to be determined.

The RSs within the 1.8-kb region upstream of the *FT* start codon in *G. arboreum*, *G. raimondii*, and *G. hirsutum* were further analyzed. We found that a large number of repeated fragments exist between 1.1- and 1.4-kb upstream of the *FT* promoters, where the *cis*-acting elements are also enriched (Figure 4).

Thus, these repetitive sequences between 1.1- and 1.4-kb upstream of the translation start site may play important roles in the activity of the cotton *FT* promoter.

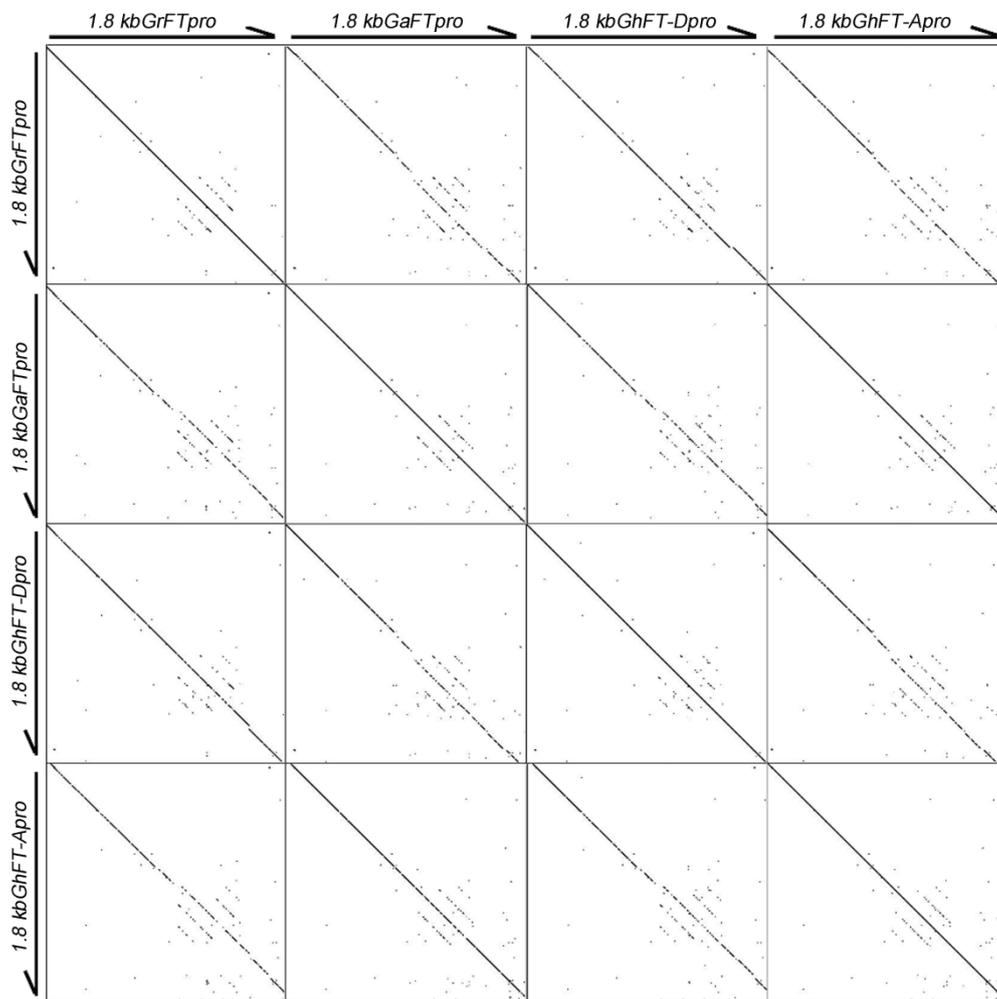


Figure 4. Repeat sequences in the 1.8-kb cotton *FT* promoter in *G. hirsutum*, *G. arboreum*, and *G. raimondii*. Harr-plot was used and the black arrow points downstream of the promoter.

2.5. Histochemical Activity of the 1.8-kb Sequence Upstream of the *FT* Promoter in Cotton

We hypothesized that the proximal cotton *FT* promoter (−1 to −1000 bp) may play important roles in *FT* regulation. To validate this hypothesis, two deletion constructs were generated (Figure 5A). The 1.0- and 1.8-kb *FT* promoters were independently fused to *GUS* to generate various *pFT:GUS* constructs, and they were stably transformed into Col-0 using *Agrobacterium*. *GUS*-staining assays in *Arabidopsis* showed that the strong *GUS* signals were induced by both 1.0- and 1.8-kb sequences upstream of cotton *FT*. Furthermore, the activity level of the *GUS* reporter induced by the 1.0-kb transgenic plants was stronger than in the 1.8-kb transgenic plants; however, there was no difference in the *GUS* signal induced by the A and D subgenomes' 1.0-kb *FT* promoter fragments (Figure 5B). A gene transcription analysis using quantitative real-time PCR (qRT-PCR) showed that the expression levels of *GUS* induced by both the 1.0- and 1.8-kb *FT* promoters were consistent with the observed *GUS* signaling levels (Figure 5C).

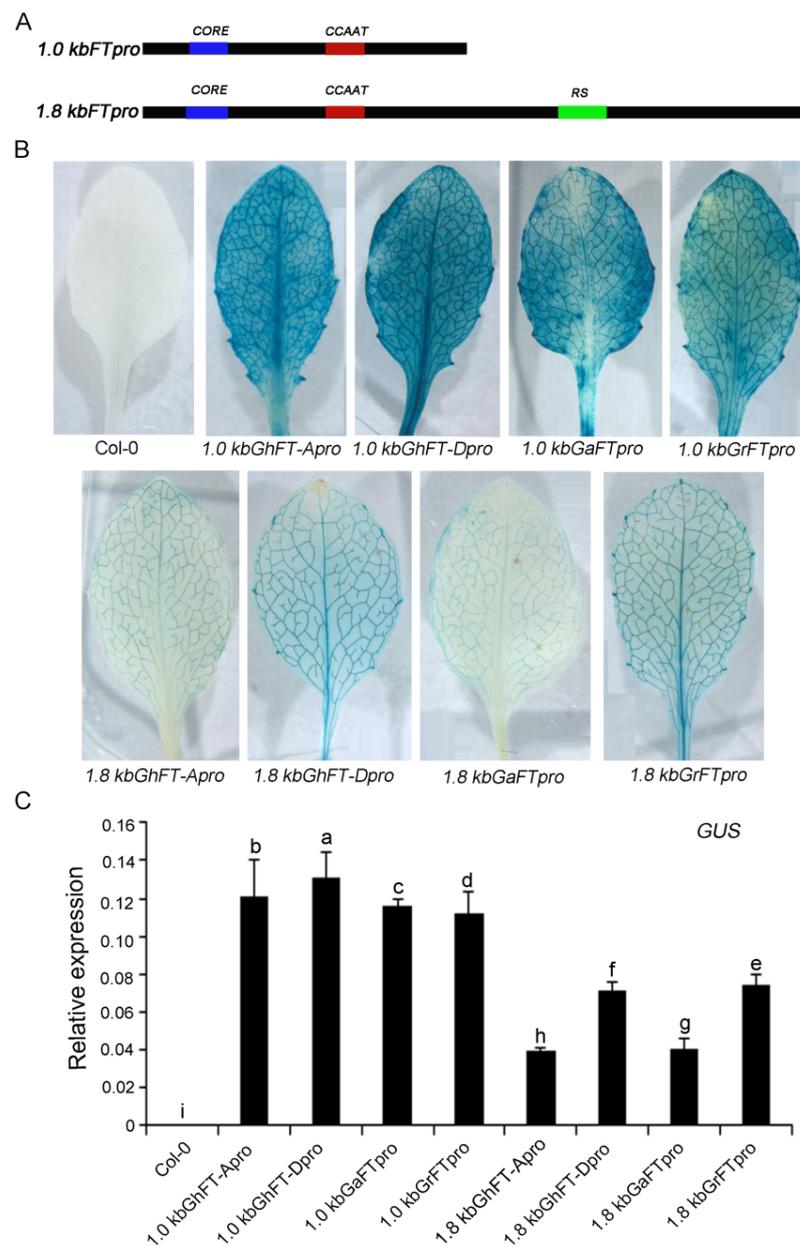


Figure 5. Histochemical activity analyses of the 1.8-kb sequence upstream of cotton *FT* promoters. **(A)** Truncation of the cotton *FT* promoters of *G. hirsutum*, *G. arboreum*, and *G. raimondii*. *RS* represents the domain of enriched repeat sequences. **(B)** The histochemical analysis of *GUS* activity in the leaves of *Arabidopsis* growing on flowering day; **(C)** The expression level of the *GUS* gene in the transgenic *Arabidopsis* and wild type plants (Col-0) by qRT-PCR. Total RNA was extracted from the true leaf tissues of *Arabidopsis* on flowering day. Data represent the mean \pm SE ($n = 3$) obtained from three biological replicates. Statistically significant differences are indicated by different lower-case letters ($p < 0.05$, Duncan's multiple range tests).

2.6. Functional Analysis of the 1.8-kb Sequence Upstream of the *FT* Promoter in Cotton

To further determine whether regulatory elements exist in the 1.8-kb sequence upstream of the cotton *FT* translation start site, we next generated transgenic *ft-10* mutant *Arabidopsis* harboring 1.0- or 1.8-kb promoter fragments fused to the cDNA of *FT*. Consistent with the reporter gene assays, complementation analyses revealed that these constructs could partially rescue the late-flowering phenotype of *ft-10* plants grown under LD conditions, and the 1.0-kb *FT* promoter fragments were more

efficient in rescuing the late-flowering phenotype than the 1.8-kb promoter fragments (Figure 6A,B) in transgenic *Arabidopsis*. In addition, gene transcription analyses in *Arabidopsis* using qRT-PCR showed that the *GhFT* gene was overexpressed in all the transgenic *ft-10* plants under LD conditions (Figure 6C). The data demonstrated that the 1.0-kb *FT* promoter fragment might play an important role in the activity levels of *FT* promoters in cotton, but the possibility of other regions having specific transcriptional regulatory functions was not excluded.

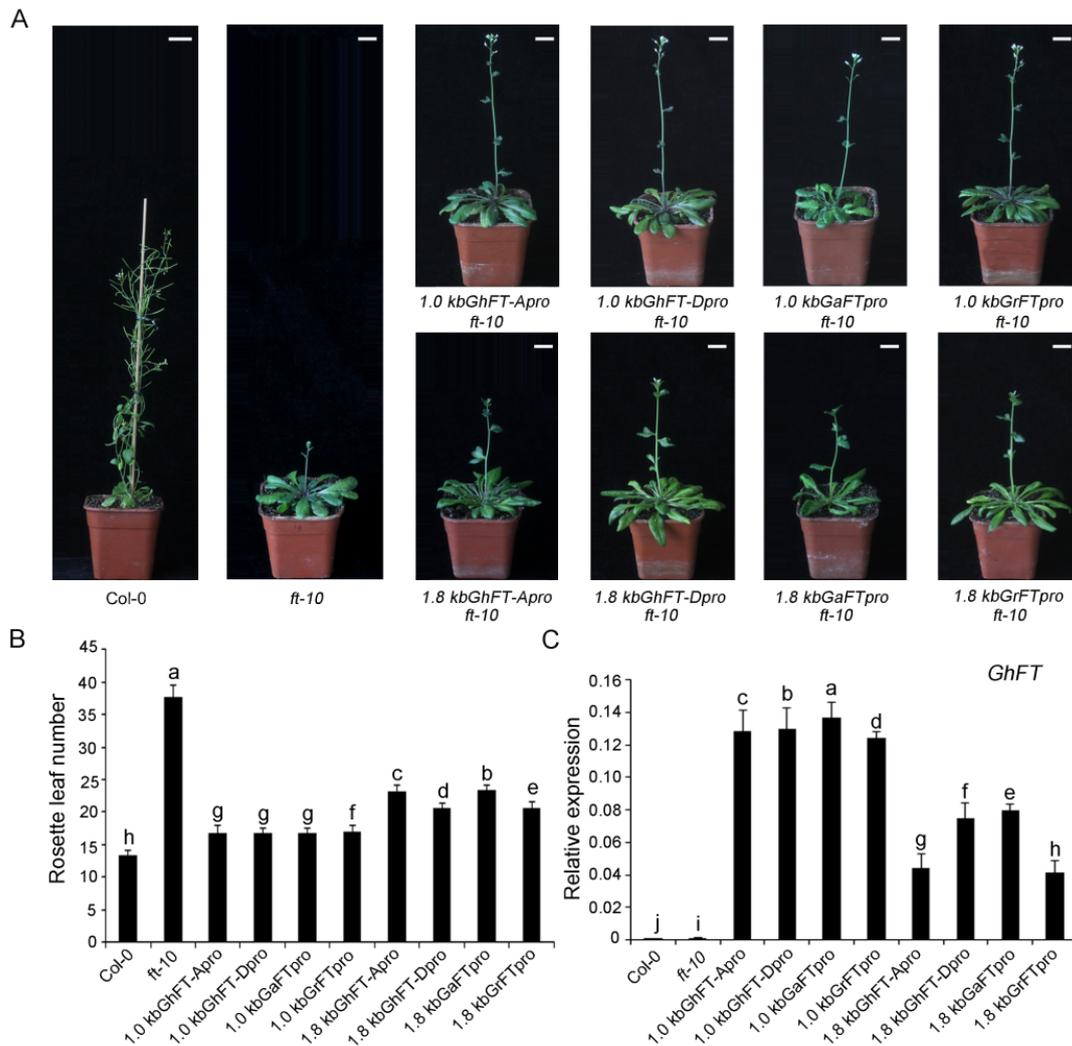


Figure 6. Functional analysis of the 1.8-kb sequence upstream of the *FT* promoter of *G. hirsutum*, *G. arboreum*, and *G. raimondii* using *Arabidopsis* as a model plant. (A) Appearance of 30 d Col-0, *ft-10*, *FTpro*:*GhFT* transgenic *Arabidopsis* lines grown under LD conditions. Scale bar, 2 cm. (B) Flowering time was measured as the rosette number of leaves per plant. (C) The expression level of *GhFT* was determined by qRT-PCR. Data represent the mean value obtained from three biological replicates. Statistically significant differences are indicated by different lower-case letters ($p < 0.05$, Duncan's multiple range tests).

3. Discussion

Cotton is an important economic crop, and upland cotton is the most widely planted species in the world [35]. A whole genome-wide analysis showed that there was only one homoeologous *FT* gene pair in the cotton genome, and the *FT*-homologue from *G. hirsutum* promotes early flowering in *A. thaliana* and *Nicotiana tabacum*, providing a good foundation for studying the molecular mechanisms of flowering regulation in cotton [24,25,28,36–38]. In this study, we studied the promoters of *GhFT*

from the A and D subgenomes of *G. hirsutum*, *GaFT* from the A₂ genome of *G. arboreum* and *GrFT* from the D₅ genome of *G. raimondii*.

3.1. Cotton FT Promoter Activity Levels Vary with Length and Origin

The promoter functions of the *FT* gene have been studied in several important crops, such as soybean [21], rice [22], and wheat [23], and the results suggested that the promoter activity varies among the different species. In this study, we identified the 1.8-kb promoters of *GhFT-A*, *GhFT-D*, *GaFT*, and *GrFT*, and a phylogenetic analysis classified the four promoters into two types, which corresponded to their genomic origins (Figure 3). A multiple sequence alignment revealed five distinct deletion fragments in the 5.9-kb cotton promoter (Figure 1). A large number of putative plant *cis*-elements were located in the proximal cotton end, such as the TATA-box and CAAT-box, for transcription initiation. In addition, some elements were found to be associated with hormones, light, stress tolerance, tissue expression, and circadian rhythm regulation (Tables S1 and S2).

The binding site CCAAT box of CO and NF-Y transcription factor are located approximately 5.3-kb upstream of *FT* in *Arabidopsis*, whereas the CORE1 and CORE2 sites are located between –220- and –161-bp upstream of the start codon [12]. Functional analyses of the *Arabidopsis FT* promoter indicated that the distal CCAAT element and the proximal CORE sites can regulate *FT* gene expression to control flowering transition. In this study, we analyzed the CORE elements and CCAAT domain of the cotton 1.8-kb *GhFT* promoter and found that they are all located 1.0-kb upstream of the promoter (Figure S1A,B), indicating that the 1.0-kb cotton *FT* promoter may contain the region responsible for activity, but the roles of other regions were not excluded. The CORE elements and CCAAT domain play key roles in the activation of the *FT* promoter in *Arabidopsis*, and the promoter has no activity if any of these motifs are deficient [12,16,18]. In addition, a number of repetitive fragments were found between the 1.1 and 1.4 kb in the *GhFT* promoters (Figure 4), which is an area that contained a large number of *cis*-regulatory elements. We hypothesized that this region may be a binding region of some unknown transcription factors and may also have an effect on the activity of the *FT* promoter. Therefore, we truncated the 1.8-kb *FT* promoter to 1.0 kb (Figure 5A).

GUS histochemical staining assays showed that the GUS activities induced by the 1.0-kb promoters of *GhFT-A*, *GhFT-D*, *GaFT*, and *GrFT* were stronger than their respective 1.8-kb promoters (Figure 5B,C), which indicated that the TGTG(N2-N3)ATG, and CCAAT domains might also determine the activity of the cotton *FT* promoter. Our study also demonstrated that there was no significant difference in promoter activities between the 1.0- or 1.8-kb promoter fragments of the A and D subgenomes using deletion constructs experiments. The promoter activity in the 1.8-kb fragment was weaker than that in the 1.0-kb fragment, implying that some unknown silencers may exist in the proximal cotton *FT* promoter region (–1.0 to –1.8 kb). A large number of repetitive sequences between 1.1- and 1.4-kb upstream of the *FT* translation start site were found (Figure 4), and we speculated that some unknown silencers are probably located in the repetitive enrichment region, which warrants further investigation.

Genetic complement assays showed that the 1.0-kb promoters of *FT* were more efficient in rescuing the late-flowering phenotype of *ft-10* plants than were the 1.8-kb promoters (Figure 6A), with similar rosette leaf numbers being obtained (Figure 6B). Furthermore, the bioinformatics analyses showed that the CORE and CCAAT domains existed in the 1.0-kb promoter of *GhFT*. Based on these experimental results, we speculate that the specific regulation of flowering transition, which involves CO and NF-Y, may exist in this region, but the influences of other regions were not excluded.

3.2. Photoperiod Sensitivities of the Promoter Regions from FT Homologues in Arabidopsis and Cottons Are Different

In *Arabidopsis*, under LD inductive conditions, transcriptional activation of *FT* is directly mediated by CO, which has two responsive elements, CORE1 and CORE2, in the *FT* promoter that share the consensus sequence TGTG(N2-3)ATG [1,9]. The CCAAT-box binding site of CO and NF-Y transcription factors exist approximately 5.3-kb upstream of *FT* in *Arabidopsis*. Rice is a short-day (SD) plant, and

the CCAAT box of the potential *cis*-element in the *Hd3a* promoter did not contain any nucleotide changes compared with in *Arabidopsis* [11,22]. The rice *Hd1* gene, which is an orthologue of the *Arabidopsis* CO gene, suppresses the transcriptional activation of *Hd3a* and inhibits flowering under LD conditions. In contrast, *Hd1* gene activates *Hd3a* expression, causing the promotion of flowering under SD conditions [15,39]. Wild cotton is a perennial plant, and most species respond to short day-photoperiods, resulting in a variety of plant types and different flowering time. However, domesticated cotton varieties have undergone extensive artificial selection, resulting in their gradual loss in photoperiodism [27]. Our research revealed that in upland cotton, CORE and CCAAT domains were concentrated in the 1.0-kb region upstream of the translation start site (Figure S1A,B), which was different from in *Arabidopsis* [12]. Compared with the consensus sequence TGTG(N2-3)ATG, the CORE in the cotton *FT* promoter is atypical TGTG(N2-3)ATG. We speculate that CO homolog may bind to the atypical TGTG(N2-3)ATG in upland cotton, but it remains to be confirmed in future. However, whether the FT/CO regulatory model in upland cotton is conserved, as in *Arabidopsis*, remains to be further studied.

4. Materials and Methods

4.1. Plant Materials and Growth Conditions

G. hirsutum cv. Xinluzao 42, *G. raimondii*, and *G. arboreum* were field-grown during the summer of 2016 in Shihezi (Xinjiang, China) under natural conditions. The seeds of the mutant *ft-10* (in the Col-0 background) were ordered from the Arabidopsis Biology Resources Center (ABRC, Columbus, OH, USA). The methods of seed sterilization and cultivation for *ft-10* and *A. thaliana* Col-0, as well as the seedling growth conditions, used in this study were published previously [23,24]. For gene expression analyses, the fresh leaves of Col-0, *ft-10*, and the transgenic lines growing for 20 d were harvested under LD conditions. All the samples were frozen immediately in liquid nitrogen and stored at -80°C until used.

4.2. Phylogenetic Analysis

The genomic sequences of *FTs* from *G. hirsutum*, *G. arboreum*, and *G. raimondii* were assembled using a batch Basic Local Alignment Search Tool (BLAST) search of shotgun sequences obtained from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). mVISTA was used to perform pairwise alignments (<http://genome.lbl.gov/vista>) [40,41]. The alignment of the nucleotide sequences and the phylogeny reconstruction analysis of the *FT* promoter in cotton were performed as previously described [42–44].

4.3. Cloning of the *FT* Promoter, Vector Construction, and Transformation of *Arabidopsis*

The cotton *FT* promoter, extending approximately 1.0- and 1.8-kb upstream of the ATG, was cloned by PCR amplification using gene-specific primers designed based on the putative sequences in the *G. hirsutum*, *G. arboreum*, and *G. raimondii* genome database (Table S3) and then ligated to the *pJET 1.2/blunt* vector following the manufacturer's instructions. The PlantCare database (<http://bioinformatics.psb.ugent.be/webtools/plant-care/html/>) was used to identify potential *cis*-regulatory elements. To analyze the RSs of the *FT* promoters, we used the Nucleic Acid Dot Plots website (<https://en.vectorbuilder.com/tool/nucleic-acid-dot-plot.html>) and entered the promoters as required. In the resulting figure, repetitive sequences were represented by point-intensive areas and were calculated according to the values on the coordinate axis. Then, the corresponding *FT* promoters, which were corrected by sequencing, and the *GhFT* gene of *35S:FT* [24] were digested by restriction enzymes to independently replace the *CaMV 35S* promoter and *GUS* gene of pCAMBIA 1301 to generate *pGhFT:FT* constructs. The resulting constructs were transformed into *Agrobacterium tumefaciens* GV3101, which was used to transform Col-0 and *ft-10*, as previously described [45]. Kanamycin-resistant transformants were selected, and homozygotes were replanted and subsequently monitored for flowering using

non-transgenic *ft-10* seedlings as controls. Flowering time was assessed using the number of rosette leaves per plant at the time of the first flower bloomed.

To evaluate the activity of the *GhFT* promoter, *pGhFT:GUS* was constructed and transformed into *A. thaliana* (Col-0) to generate independent transgenic lines. The activities of *GhFT* promoters were detected by histochemical GUS staining. Fresh plant tissues were harvested and immersed in GUS staining solution [0.1 M Na₂HPO₄·12H₂O, 0.1 M NaH₂PO₄·2H₂O, pH 7.0, 0.1% Triton X-100, 5 mM K₃[Fe(CN)₆], 5 mM K₄[Fe(CN)₆]·3H₂O, 10 mM EDTA-Na₂, pH 7.0 and 1 mM 5-bromo-4-chloro-3-indolyl-b-D-glucuronic acid]. After incubation at 37 °C overnight, the plant materials were rinsed with 70% ethanol. The resulting stained tissues were observed under a microscope.

4.4. Gene Expression Analysis

Total RNA for each sample was isolated using the RNAPrep pure Plant Kit (Tiangen, Beijing, China) and treated with RNase-free DNase according to the manufacturer's protocol. The quality, quantity, and integrity of the total RNA extracted were assessed as previously described [24]. cDNA was synthesized from 200 ng RNA using the Superscript First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. qRT-PCR was carried out using the SYBR Green Master Mixture (CWBIO, Beijing, China) on an Applied Biosystems 7500 Fast Real-Time PCR System (Life Technologies, Foster City, CA, USA). The PCR amplification system and program were as described previously [14]. Three biological replications were performed with independently isolated RNA in all the qRT-PCR assays. Relative gene expression levels were calculated using the 2^{-ΔΔCt} method [46].

4.5. Statistical Analysis

The results were based on three independent experiments with at least three replicates. The SPSS software package (ver.17.0; SPSS Inc., Chicago, IL, USA) was used for the statistical analysis, as previously described [47].

5. Conclusions

In this study, we identified a 5.9-kb promoter of *FT* from *G. hirsutum* and revealed that the activity level of the truncated *GhFT* promoter varied in different tissues. Further studies on the 1.0- and 1.8-kb fragments of the *FT* promoters in *G. hirsutum*, *G. arboreum*, and *G. raimondii* showed that the activity of the 1.0-kb *FT* promoter was higher than that of the 1.8-kb promoter. These results illustrate that the proximal *FT* promoter fragment might play an important role affecting the activity levels of *FT* promoters in cotton flowering.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/19/4769/s1>.

Author Contributions: X.H. and Y.S. conceived the experiments. N.S., D.C., and C.L. performed the experiments. N.S., D.C., and C.L. analyzed the data. N.S. and X.H. wrote the paper. All authors read and approved the final manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (31860393, 31360366), and the Technological Innovation Leading Talents of Xinjiang Production and Construction Corps (2006BC001).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Samach, A.; Onouchi, H.; Gold, S.E.; Ditta, G.S.; Schwarz-Sommer, Z.; Yanofsky, M.F.; Coupland, G. Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. *Science* **2000**, *288*, 1613–1616. [[CrossRef](#)] [[PubMed](#)]
2. Corbesier, L.; Coupland, G. The quest for florigen: A review of recent progress. *J. Exp. Bot.* **2006**, *57*, 3395–3403. [[CrossRef](#)] [[PubMed](#)]

3. Fornara, F.; de Montaigu, A.; Coupland, G. SnapShot: Control of flowering in *Arabidopsis*. *Cell* **2010**, *141*, 550. [[CrossRef](#)] [[PubMed](#)]
4. Schultz, E.A.; Haughn, G.W. *LEAFY*, a homeotic gene that regulates inflorescence development in *Arabidopsis*. *Plant Cell* **1991**, *3*, 771–781. [[CrossRef](#)] [[PubMed](#)]
5. Abe, M.; Kobayashi, Y.; Yamamoto, S.; Daimon, Y.; Yamaguchi, A.; Ikeda, Y.; Ichinoki, H.; Notaguchi, M.; Goto, K.; Araki, T. FD, a bZIP protein mediating signals from the floral pathway integrator *FT* at the shoot apex. *Science* **2005**, *309*, 1052–1056. [[CrossRef](#)] [[PubMed](#)]
6. Wigge, P.A.; Kim, M.C.; Jaeger, K.E.; Busch, W.; Schmid, M.; Lohmann, J.U.; Weigel, D. Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* **2005**, *309*, 1056–1059. [[CrossRef](#)] [[PubMed](#)]
7. Yoo, S.K.; Chung, K.S.; Kim, J.; Lee, J.H.; Hong, S.M.; Yoo, S.Y.; Lee, J.S.; Ahn, J.H. *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in *Arabidopsis*. *Plant Physiol.* **2005**, *139*, 770–778. [[CrossRef](#)]
8. Kaufmann, K.; Wellmer, F.; Muiño, J.M.; Ferrier, T.; Wuest, S.E.; Kumar, V.; Serrano-Mislata, A.; Madueño, F.; Krajewski, P.; Meyerowitz, E.M.; et al. Orchestration of floral initiation by *APETALA1*. *Science* **2010**, *328*, 85–89. [[CrossRef](#)]
9. Tiwari, S.B.; Shen, Y.; Chang, H.C.; Hou, Y.; Harris, A.; Ma, S.F.; McPartland, M.; Hymus, G.J.; Adam, L.; Marion, C.; et al. The flowering time regulator *CONSTANS* is recruited to the *FLOWERING LOCUS T* promoter via a unique *cis*-element. *New Phytol.* **2010**, *187*, 57–66. [[CrossRef](#)]
10. Ben-Naim, O.; Eshed, R.; Parnis, A.; Teper-Bamnolker, P.; Shalit, A.; Coupland, G.; Samach, A.; Lifschitz, E. The CCAAT binding factor can mediate interactions between *CONSTANS*-like proteins and DNA. *Plant J.* **2006**, *46*, 462–476. [[CrossRef](#)]
11. Wenkel, S.; Turck, F.; Singer, K.; Gissot, L.; Le Gourrierc, J.; Samach, A.; Coupland, G. *CONSTANS* and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. *Plant Cell* **2006**, *18*, 2971–2984. [[CrossRef](#)] [[PubMed](#)]
12. Cao, S.; Kumimoto, W.; Gnesutta, N.; Calogero, M.; Mantovani, R.; Holt, B.F. A Distal CCAAT/NUCLEAR FACTOR Y Complex Promotes Chromatin Looping at the *FLOWERING LOCUS T* Promoter and Regulates the Timing of Flowering in *Arabidopsis*. *Plant Cell* **2014**, *26*, 1009–1017. [[CrossRef](#)] [[PubMed](#)]
13. Li, F.; Fan, G.; Lu, C.; Xiao, G.; Zou, C.; Kohel, R.J.; Ma, Z.; Shang, H.; Ma, X.; Wu, J.; et al. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* **2015**, *33*, 524–530. [[CrossRef](#)] [[PubMed](#)]
14. Cai, D.; Liu, H.; Sang, N.; Huang, X. Identification and characterization of *CONSTANS*-like (*COL*) gene family in upland cotton (*Gossypium hirsutum* L.). *PLoS ONE* **2017**, *12*, e0179038. [[CrossRef](#)] [[PubMed](#)]
15. Hayama, R.; Yokoi, S.; Tamaki, S.; Yano, M.; Shimamoto, K. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* **2003**, *422*, 719–722. [[CrossRef](#)] [[PubMed](#)]
16. Adrian, J.; Farrona, S.; Reimer, J.J.; Albani, M.C.; Coupland, G.; Turck, F. *cis*-Regulatory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in *Arabidopsis*. *Plant Cell* **2010**, *22*, 1425–1440.
17. Liu, Y.; Li, X.; Li, K.; Liu, H.; Lin, C. Multiple bHLH proteins form heterodimers to mediate *CRY2*-dependent regulation of flowering-time in *Arabidopsis*. *PLoS Genet.* **2013**, *9*, e1003861. [[CrossRef](#)]
18. Liu, L.; Adrian, J.; Pankin, A.; Hu, J.; Dong, X.; von Korff, M.; Turck, F. Induced and natural variation of promoter length modulates the photoperiodic response of *FLOWERING LOCUS T*. *Nat. Commun.* **2014**, *5*, 4558. [[CrossRef](#)]
19. Li, D.; Liu, C.; Shen, L.; Wu, Y.; Chen, H.; Robertson, M.; Helliwell, C.A.; Ito, T.; Meyerowitz, E.; Yu, H. A repressor complex governs the integration of flowering signals in *Arabidopsis*. *Dev. Cell* **2008**, *15*, 110–120. [[CrossRef](#)]
20. Chen, L.; Bernhardt, A.; Lee, J.; Hellmann, H. Identification of *Arabidopsis* MYB56 as a Novel Substrate for CRL3BPM E3 Ligases. *Mol. Plant* **2015**, *8*, 242–250. [[CrossRef](#)]
21. Jiang, B.; Yue, Y.; Gao, Y.; Ma, L.; Sun, S.; Wu, C.; Hou, W.; Lam, H.M.; Han, T. *GmFT2a* polymorphism and maturity diversity in soybeans. *PLoS ONE* **2013**, *8*, e77474. [[CrossRef](#)] [[PubMed](#)]
22. Takahashi, Y.; Teshima, K.M.; Yokoi, S.; Innan, H.; Shimamoto, K. Variations in *Hd1* proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4555–4560. [[CrossRef](#)] [[PubMed](#)]

23. Yan, L.; Fu, D.; Li, C.; Blechl, A.; Tranquilli, G.; Bonafede, M.; Sanchez, A.; Valarik, M.; Yasuda, S.; Dubcovsky, J. The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19581–19586. [[CrossRef](#)] [[PubMed](#)]
24. Guo, D.; Li, C.; Dong, R.; Li, X.; Xiao, X.; Huang, X. Molecular cloning and functional analysis of the *FLOWERING LOCUS T (FT)* homolog *GhFT1* from *Gossypium hirsutum*. *J. Integr. Plant Biol.* **2015**, *57*, 522–533. [[CrossRef](#)] [[PubMed](#)]
25. Li, C.; Zhang, Y.; Zhang, K.; Guo, D.; Cui, B.; Wang, X.; Huang, X. Promoting flowering, lateral shoot outgrowth, leaf development and flower abscission in tobacco plants overexpressing cotton *FLOWERING LOCUS T (FT)*-like gene *GhFT1*. *Front. Plant Sci.* **2015**, *6*, 454. [[CrossRef](#)] [[PubMed](#)]
26. McGarry, R.C.; Ayre, B.G. Geminivirus-mediated delivery of florigen promotes determinate growth in aerial organs and uncouples flowering from photoperiod in cotton. *PLoS ONE* **2012**, *7*, e36746. [[CrossRef](#)] [[PubMed](#)]
27. McGarry, R.C.; Prewitt, S.; Ayre, B.G. Overexpression of *FT* in cotton affects architecture but not floral organogenesis. *Plant Signal. Behav.* **2013**, *8*, e23602. [[CrossRef](#)] [[PubMed](#)]
28. McGarry, R.C.; Prewitt, S.F.; Culpepper, S.; Eshed, Y.; Lifschitz, E.; Ayre, B.G. Monopodial and sympodial branching architecture in cotton is differentially regulated by the *Gossypium hirsutum* *SINGLE FLOWER TRUSS* and *SELF-PRUNING* orthologs. *New Phytol.* **2016**, *212*, 244–258. [[CrossRef](#)]
29. Mantovani, R. The molecular biology of the CCAAT-binding factor NF-Y. *Gene* **1999**, *239*, 15–27. [[CrossRef](#)]
30. McNabb, D.S.; Pinto, I. Assembly of the Hap2p/Hap3p/Hap4p/Hap5p-DNA complex in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **2005**, *4*, 1829–1839. [[CrossRef](#)]
31. Dolfini, D.; Gatta, R.; Mantovani, R. NF-Y and the transcriptional activation of CCAAT promoters. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 29–49. [[CrossRef](#)] [[PubMed](#)]
32. Hayama, R.; Coupland, G. The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. *Plant Physiol.* **2004**, *135*, 677–684. [[CrossRef](#)] [[PubMed](#)]
33. Böhlenius, H.; Huang, T.; Charbonnel-Campaa, L.; Brunner, A.M.; Jansson, S.; Strauss, S.H.; Nilsson, O. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* **2006**, *312*, 1040–1043. [[CrossRef](#)] [[PubMed](#)]
34. Ballerini, E.S.; Kramer, E.M. In the light of evolution: A reevaluation of conservation in the CO-FT regulon and its role in photoperiodic regulation of flowering time. *Front. Plant Sci.* **2011**, *2*, 81. [[CrossRef](#)] [[PubMed](#)]
35. Saha, S.; Jenkins, J.N.; Wu, J.; McCarty, J.C.; Gutie rrez, O.A.; Percy, R.G.; Cantrell, R.G.; Stelly, D.M. Effects of chromosome-specific introgression in upland cotton on fiber and agronomic traits. *Genetics* **2006**, *172*, 1927–1938. [[CrossRef](#)] [[PubMed](#)]
36. Zhang, X.; Wang, C.; Pang, C.; Wei, H.; Wang, H.; Song, M.; Fan, S.; Yu, S. Characterization and functional analysis of PEBP family genes in Upland Cotton (*Gossypium hirsutum* L.). *PLoS ONE* **2016**, *11*, e0161080. [[CrossRef](#)] [[PubMed](#)]
37. Prewitt, S.F.; Ayre, B.G.; McGarry, R.C. Cotton *CENTRORADIALIS/TERMINAL FLOWER 1/SELF-PRUNING* genes functionally diverged to differentially impact plant architecture. *J. Exp. Bot.* **2018**, *69*, 5403–5417. [[CrossRef](#)] [[PubMed](#)]
38. Wang, M.; Tan, Y.; Cai, C.; Zhang, B. Identification and expression analysis of phosphatidy ethanolamine-binding protein (PEBP) gene family in cotton. *Genomics* **2018**. [[CrossRef](#)] [[PubMed](#)]
39. Yano, M.; Katayose, Y.; Ashikari, M.; Yamanouchi, U.; Monna, L.; Fuse, T.; Baba, T.; Yamamoto, K.; Umehara, Y.; Nagamura, Y.; et al. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* **2000**, *12*, 2473–2484. [[CrossRef](#)]
40. Brudno, M.; Malde, S.; Poliakov, A.; Do, C.B.; Couronne, O.; Dubchak, I.; Batzoglou, S. Global alignment: Finding rearrangements during alignment. *Bioinformatics* **2003**, *19*, i54–i62. [[CrossRef](#)]
41. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)] [[PubMed](#)]
42. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [[CrossRef](#)] [[PubMed](#)]
43. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739. [[CrossRef](#)] [[PubMed](#)]

44. Si, Z.; Liu, H.; Zhu, J.; Chen, J.; Wang, Q.; Fang, L.; Gao, F.; Tian, Y.; Chen, Y.; Chang, L.; et al. Mutation of *SELF-PRUNING* homologs in cotton promotes short-branching plant architecture. *J. Exp. Bot.* **2018**, *69*, 2543–2553. [[CrossRef](#)] [[PubMed](#)]
45. Clough, S.J.; Bent, A.F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **1998**, *16*, 735–743. [[CrossRef](#)] [[PubMed](#)]
46. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
47. Yang, L.; Jin, Y.; Huang, W.; Sun, Q.; Liu, F.; Huang, X. Full-length transcriptome sequences of ephemeral plant *Arabidopsis pumila* provides insight into gene expression dynamics during continuous salt stress. *BMC Genomics* **2018**, *19*, 717. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).