Genome-wide comparative analysis of *LEAFY* promoter sequence in angiosperms

Zhichao Sun¹ · Zheng Li¹ · Jianqin Huang¹ · Bingsong Zheng¹ ·
Liangsheng Zhang¹ · Zhengjia Wang¹

Received: 7 July 2016 / Revised: 7 November 2016 / Accepted: 18 November 2016 / Published online: 19 December 2016
© Prof. H.S. Srivastava Foundation for Science and Society 2016

Abstract Regulation of the flowering mechanism is influenced by many environmental factors. Dissecting the regulatory processes upstream of the *LFY* (*LEAFY*) gene will help us to understand the molecular mechanisms of floral induction. In total, 53 *LFY* sequences were identified in 37 species. Among the 53 selected *LFY* promoters and after eliminating the short sequences, 47 *LFY* promoters were analyzed. Comparative genome studies for *LFY* promoters among plants showed that TATA-box existed in all herbaceous plants. The 1345-bp promoter sequence upstream to hickory *LFY* gene was cloned and analyzed, together with functional studies. The result of sequence alignment showed that the region of the hickory *LFY* promoter has only two conserved auxin response elements (AuxRE), whereas other plants had four. The positions of AuxRE in hickory and walnut were the same, but they were different from the positions from other plants. Furthermore the sequence analysis showed that the promoter have TATA-box and CAAT-box motifs. Deletion analysis of these motifs did not block β -glucuronidase (GUS) activity during the transient expression assay, suggesting that it may be a TATA-less promoter. Low temperature and light significantly induced the full-length promoter to increase

about two folds of the GUS enzymatic activity, suggesting these environmental factors induced flowering in hickory.

Keywords *LFY* genes · Promoter · *Carya cathayensis* · Sequence analysis · Transient expression · Fluorometric analysis

Introduction

The transition from vegetative growth to sexual reproduction in angiosperms is regulated by responses mainly to internal and external cues including fluxes in plant growth regulators, and changes in photoperiod and temperature (Mouradov et al. 2002). Four major transition routes for initiation of flowering have been described and include the following: photoperiodic, gibberellin, autonomous pathway, and vernalization (Liu et al. 2009; Mouradov et al. 2002; Simpson and Dean 2002). Increased expression of the floral meristem identity gene *SOC1* (*SUPPRESSOR OF OVEREXPRESSION OF CO1*) was common in all four pathways in flowering (Lee and Lee 2010; Ma et al. 2011). The floral genes *API* (*APETALA1*), *AP3*, *PI* (*PISTILLATA*), and *AG* (*AGAMOUS*) are all necessary for initiation and development of flowers, and activated after the *LFY* (*LEAFY*) gene is prompted by *SOC1* (Du and Pijut 2010; Hou et al. 2011; Li et al. 2012; Liu et al. 2010). Therefore, *LFY* is the key gene for transition to flowering.

LFY encodes a plant-specific transcription factor that regulates the change from vegetative to flowering stage, and directs the development of initiated floral meristems. *LFY* plays a central role during the floral meristem development (Moyroud et al. 2010; Shiokawa et al. 2008) by inducing the expression of floral homeotic genes (Hamès et al. 2008). The *LFY* genes are present in all terrestrial plants and its sequence

Electronic supplementary material The online version of this article (doi:10.1007/s12298-016-0393-8) contains supplementary material, which is available to authorized users.

✉ Liangsheng Zhang
fafuzhang@163.com

✉ Zhengjia Wang
wzhj21@163.com

¹ School of Forestry and Biotechnology, Zhejiang Agriculture and Forestry University, Dong Hu Campus, 88 Northern Circle Road, Linan 311300, China

is highly conserved throughout the plant kingdom, however the translated protein is not similar to other proteins (Maizel et al. 2005). Floral transition is governed by the gradual increase in *LFY* expression. *LFY* is present as a single copy in most angiosperms, which differs from other transcription factor families that have multiple copies (Riechmann and Ratcliffe 2000; Shiu et al. 2005). Because of its pivotal and critical role in initiating floral meristems, *LFY* is placed at the center of different evolutionary scenarios that attempt to explain the evolution of successful angiosperm groups (Albert et al. 2002; Dornelas and Rodriguez 2006; Theissen and Melzer 2007). Many homologs of *LFY* have been isolated from diverse species, including woody plants such as *Cedrela fissilis* (Dornelas and Rodriguez 2006), *Hevea brasiliensis* (Dornelas and Rodriguez 2005a), and *Pinus caribaea* (Dornelas and Rodriguez 2005b).

Woody plants exhibit vegetative growth during early development, although they receive flowering signals every year. However, when the *LFY* gene was transferred from *Arabidopsis* to *Populus*, a woody plant, the juvenile phase (non-reproductive) was substantially reduced (Weigel and Nilsson 1995). Therefore, transfer of *LFY* gene could shorten the breeding cycle of some plants, many woody species, which have a long juvenile phase (Shiokawa et al. 2008).

Hickory (*Carya cathayensis* Sarg.) and walnut belong to the same family Juglandaceae, and both trees are known for their nuts, which are of high nutrition. However, hickory trees raised from seed typically requires more than 10 years before it flowers, which causes prolonged or protracted breeding programs. Additionally, female and male flowers have different growth periods. The *CcLFY* gene, a *LFY* homolog, has been identified in hickory (Wang et al. 2012). Dissecting the regulatory processes upstream of the *LFY* gene will help us to understand the molecular mechanisms of floral induction sensor relay. In this study, we isolated and analyzed a 1302-bp section of the 5'-flanking region of the *CcLFY* gene. We constructed various promoter deletion constructs and conducted transient GUS expression analyses in tobacco to clarify the roles of the various promoter regions. We demonstrated that *CcLFY* has a TATA-less promoter and that low temperature and light induce *CcLFY* expression.

Materials and methods

Computer analysis

DNA sequences were analyzed by DNAMAN software, Clustalx, and MEGA.5, the *CcLFY* and other plants promoter elements by PLACE, and the transcription start sites (TSS) and TATA-box were predicated on the website: http://www.fruitfly.org/seq_tools/promoter.html and <http://linux1.softberry.com/berry.phtml>.

WaLNUT promoter sequence was obtained from the website: http://dendrome.ucdavis.edu/ftp/Genome_Data/genome/Reju/. BioMart of phytosome was used to search other plants promoter sequences (1500 bp). Significant differences in GUS activity among treatment groups was tested by analysis of variance (ANOVA) using Microsoft Excel software.

Amplification of *CcLFY* promoter

The *CcLFY* promoter was isolated using the amplification procedure shown in Table S1. After three steps of DNA walking, a sequence approximately 1300 bp was cloned. The first amplification product and this fragment were assembled and a 1345-bp 5' flanking region was obtained. This region contained ATG of the *CcLFY* gene (Fig. 1). Genomic DNA was extracted from *C. cathayensis* cotyledons and used as the template for anchored PCR (A-PCR) amplifications. The 5'-flanking region of the *CcLFY* gene was isolated by chromosome walking A-PCR. Sequences of DNA adaptors and primers used for promoter amplification are shown in Table S2.

Construction of chimeric promoters

Based on the results of computer analysis and the flowering characteristics of hickory, we selected four light cis-acting elements and one low-temperature cis-acting element for analysis. To identify the functions of these cis-acting elements, we constructed several expression vectors in which various deleted fragments of the *CcLFY* promoter were fused with the GUS reporter gene. A series of PCR amplifications with several primers were carried-out to construct the various length deletions of *CcLFY* promoter/GUS fusion products (Table S2). Different amplified promoters sequences were cut with HindIII and BamHI, and the digested fragments were retrieved. The retrieved fragments were ligated with the CaMV35S eliminated promoter sequence in the PBI121 vector where the elimination of CaMV35S sequence in the vector was carried out using HindIII and BamHI. We constructed 11 expression vectors containing various lengths of *CcLFY* promoters, designated as follows; 1-C, L1, L2, L3, L4, T1, SPL4, SPT1, SSPL4, SSPT1, and W (Fig. 2). The CaMV35S promoter served as the positive control and tobacco leaves treated with the buffer is taken as the negative control. Each step was checked by PCR and digested by HindIII and BamHI to confirm the identity and size of each construct.

Transient expression analyses in tobacco

The expression vector constructs of different promoters and PBI-35S were introduced into *Agrobacterium tumefaciens*

1 ttgtatacct tgggccatcg atgtagatag tagtattatg ctaacttcaa caaccacatc
 61 gttttaactt cttattaact acaaatagtg agccgaagcc agaggccacc aagactgac
 121 ccattgaaat aatcagatt gatgaatac tcaaatagga actgtgtgtg cagcttcaag
 181 tacaagaat aagtggcaca agaagtggg tttgggtgac attgtagtag aatagaggg
 241 ctgcaacgac gaaagggggt gtaagtattt ttgggggtga cacaatccga gcataatcgg
 301 ttctgttttc acacctttga ataaaagaca tctgttttaa atttgaggtt atgactctgt
 361 gtgaattggt tttgtttttg ggaataact atagatataa agagattata taaagtaac
 421 tcacaattg attgtctttt attgaatcgg tttagatcat tttacaataa aaataatttt
 481 ataatttgac aagtcaaatc gaactacatc aatttataa attacttttg tatattaata
 541 tatttatacc tctaaccatt ccttttctct agagaaatag gagacttgac ttgtcaatgt
 601 cccattggat gatataaact ttttaattac accaaataat ccgatccata ggttaagtc
 661 taattccaat atttatttta taaaataact cataatttaa ttgctatttg tacggatat
 721 tttcagctag cctttgact cttattagta aattactata ttaatatatg catccatgga
 781 gggggaaaaa agattaatga gatattgagc atcaaatgt cctactctcg acttttcaa
 841 cattgttgta atggcagcct taagggtatt tatgtgtcta taccgctat ttgtccatt
 901 aaaagacgtg taagaagaa agataaaat tctgacatat ttgaaaaac ttgattgggt
 961 attactgcc actgggagta gctagcccta agagtgggag cggaaagagt acttattctg
 1021 tctgttttca ttggactaa ctagtttcaa tgaacagca tctagtttct tcttaattt
 1081 tccctgggag aaaaaccctc ccacgctcga tatatatcgt cttagcaaa gagaagaggg
 1141 ggcAgtttcg gaaacccgag aaaatactgg tgccttata ggaagcaaca gccaacagcg
 1201 tgcaccttc acgaacatt tatcaagcgg caaaagagga cgtgagtga ggaattggag
 1261 tcaaatata gaagtgtat acagaaggtt ttgtatccat cgATG
 5' UTR

Fig. 1 Nucleotide sequence of *CcLFY* promoter. Transcription start site shown in red uppercase font. Black uppercase font shows initiator codon; primers designed for deletion analysis are underlined in red. Blue underlining shows 5'UTR; cis-acting element associated with low temperature (LTR) shown in blue font; red font shows cis-acting element associated with light (ACE, G-box, SP1, GT1-motif); 1 sequence similar to CAAT-box sequence, 2 sequence similar to TATA-box sequence

strain EHA105 using the freeze–thaw method (Goto et al. 1999) and then checked by PCR to confirm introduction. Leaf discs (approx. 2 cm in diameter) were excised from leaves of *Nicotiana tabacum* using a leaf punch on the bench top and incubated for 15 min at 25 °C with transformed *A. tumefaciens* carrying the various *CcLFY* promoter constructs (Horsch et al. 1985). The discs were incubated in X-Gluc at 37 °C overnight. The chlorophyll was removed from green tissues by immersing the leaf discs in 75% ethanol and acetone:carbinol (1:3) two to

Fig. 2 Deletion constructs of *CcLFY* promoter. A in translation start site ATG represents +1; first base before ATG represents −1; 1-C full-length *CcLFY* promoter, L1 1-C 5'-end deletion of 344 bp, L2 1-C 5'-end deletion of 699 bp, L3 1-C 5'-end deletion of 869 bp, L4 1-C 5'-end deletion of 987 bp, T1 1-C 5'-end deletion of 1117 bp, SPL4 1-C deletion of L4 cis-acting element, SSPL4 promoter L4 deletion of L4 cis-acting element, SSPT1 promoter T1 deletion of T1 cis-acting element, W 1-C 5'-end deletion of 1151 bp

three times. The discs were examined and photographed under a microscope. The GUS-specific signal was visible as blue spots on the leaf discs (Horsch et al. 1985).

Activities of promoters under different environment conditions

Leaf discs infected with *A. tumefaciens* containing various *CcLFY* promoter constructs were cultured on Murashige and Skoog (MS) (An et al. 1989) medium for 8, 16, and 24 h under the conditions shown in Table S3. Activity of GUS was determined by fluorometric GUS assays as described below after grinding the leaf disc tissues (approx. 0.1 g) in liquid nitrogen (Xu et al. 1993).

Protein extraction and fluorometric GUS assays

Assays for GUS activity in infected leaf discs were based on the protocol as described by Jefferson et al. (1987) using 4-methyl-umbelliferyl glucuronide (4-MUG) as the substrate for the fluorescent assay (Jefferson et al. 1987). Proteins were quantified using the Bradford assay using Bovine Serum Albumin (BSA) as the standard (Bradford 1976). Fluorescence was measured by a spectrofluorometer (Hitachi F-4500). The excitation wavelength was 365 nm and the emission wave length was 455 nm.

Results

LFY gene analysis in different plant genomes

To identify the *LFY* genes in different plant species, a genome-wide search was performed using *Arabidopsis AtLFY* nucleotide sequences as query and BLAST against 57 genomes listed in the Phytozome database. *LFY* gene sequences were taken from 37 genomes including mosses, monocots and eudicots. Although *LFY* has been identified as a key regulator of flower development and it was also present in non-flowering plants, such as mosses and ferns. In total, 53 *LFY* sequences with e value $<4.4e-54$, coverage percentage $\geq 70\%$ and identity percentage $\geq 38.1\%$ were identified. One *LFY* sequence was observed for almost all plant genomes examined, but for *M. esculenta*, *F. vesca*, *M. domestica*, *G. max*, *L. usitatissimum*, *P. virgatum*, *Z. mays*, *M. guttatus*, *B. Rapa* FPSC, *S. moellendorffii* and *P. patens* multiple possible genes were identified. To clarify whether or not the sequences obtained from BLAST are *LFY* genes, phylogenetic analysis for the 53 full-length *LFY*-like gene sequences was carried out using sequences from *P. patens* as functional outgroups (Fig. S1). *LFY*-like genes reported from the same plants were grouped into a single clade and among which hickory and walnut *LFY* genes shared same clade. We analyzed the conserved motifs of *LFY* gene from those plant species using the MEME program (Fig. S1). The motif distribution analyses of the *LFY* proteins revealed that 25 of 53 (47.2%) from different species *LFY* proteins contain all ten domains, named as motif1 to motif10. 14 of 53 (26.4%) *LFY* proteins lack one domain; three (5.7%) *LFY* proteins lack two domains and others lack exceed two domains (Fig. S1). *LFY* genes contained both highly conserved N-terminal and C-terminal.

Promoter analysis of the *LFY* genes

Earlier phylogenetic shadowing in seven Brassicaceae plants uncovered two highly conserved, functionally important, regions in the *LFY* promoter (*pLFY*), which overlapped with the previously defined functionally important distal and proximal regions (Blázquez and Weigel 2000; Yamaguchi et al. 2013). Among the 53 selected *pLFYs* and after eliminating the short sequences, 47 *pLFYs* were analyzed. Phylogenetic trees were constructed separately for the first 1500 bp length sequences (Fig. 3a) and for the P region (Fig. 3b). Phylogenetic tree built based on the P region displayed better evolutionary relationship than that built based on D region and it was conserved among all plants examined. A TATA-box was identified in all the herbaceous plants including monocots and species in the *Brassicaceae*, but was absent in the woody plants except for members of the *Salicaceae*.

Fig. 3 Phylogenetic tree analysis for the different parts of the promoter regions; 1500 bp region (a) and P-region (b). Gene sequences and their corresponding accession numbers were taken from the phytozome genome portal. TATA box presence in the different plants are marked with red filled circle in (a) and (b). In both regions, hickory promoter regions share same clade of evolutionary relationship with walnut. Tree was constructed by MEGA 5.1 programme using the Maximum likelihood method with the bootstrap value of 1000

Analysis of the conserved motifs in the P region of the promoter

Closer inspection of the P region revealed four conserved auxin-response element (AuxRE) core motifs (Fig. 4a). These conserved (AuxRE) motifs are responsible for auxin-mediated *LFY* activation. Two of the AuxRE elements (AuxRE1 and AuxRE3) were conserved among hickory and other plants, whereas AuxRE2 motif was not found in

Fig. 4 General depiction of the LFY promoter architecture and AuxRE core motif sequences. **a** D and P are the functionally defined regions of the Arabidopsis LFY promoter. Below is the alignment of the P region sequences of the LFY promoters from different plant species. Conserved regions of the P region (AuxRE) were drawn with

black box. Walnut and hickory share the same location of AuxRE and other plants have their at different location. AuxRE 2 was not found in the hickory plant. **b** WebLogos of different AuxRE core motif sequences between hickory and other plant species are drawn

hickory. The sequence of the fourth AuxRE4 (GACA) in other plants was changed to (GACT) in hickory (Fig. 4b). The position of the AuxRE motif was also different between hickory and other plants. The sequence and position of the motif regions for hickory and walnut are at the same location whereas the positions in other plants were different.

Conservation analysis and distribution of TATA-box segments

In order to study the conserved sequence positions in the TATA and TATA less box promoters, we selected 22 TATA promoters and 23 TATA-less promoters sequences from various plants and aligned them with Clustal X. In the

promoter sequence, the regions for the TATA-box were highly conserved than others (Fig. 5a, c). The location of the TATA-box sequences in the promoter strongly influences the downstream transcript. In general, the location of the TATA-box was confined to very narrow regions between −35 to −23 bp, and −32 bp upstream of the TSS is the optimum position for plant promoters (Fig. 5a). Furthermore, the TATA-box location in mammalian core promoters are constrained to more narrow windows (−32, −29 bp) (Molina and Grotewold 2005; Ponjavic et al. 2006). The existence of wide range of TATA-box locations (−35 to −23 bp) in different plant promoters allow them to perform highly tissue specific transcription of the genes. For the TATA-less promoters, there is no specific conserved site as the TATA box (Fig. 5b), however, it contains

A

Alyrata | 917939
 Mesculenta | cassava4.1 032718m.g
 Athaliana | AT5G61850
 Sitalica | S011756m.g
 Esalsugineum | Thhalv10005606m.g
 Osativa | LOC_Os04g51000
 Graimondii | Gorai.001G053900
 Ptrichocarpa | Potri.015G106900
 Fvesca | gene30683-v1.0-hybrid
 Sbicolor | Sobic.006G201600
 Cgrandiflora | Cagra.4553s0022
 Pvirgatum | Pavir.Gb00795
 BrapaFPsc | Brara.B03775
 Bstricta | Bostr.26833s0050
 Zmays | GRMZM2G180190
 Zmays | GRMZM2G098813
 Mtruncatula | Medtr3g098560
 Bdistachyon | Bradi5g20340

B

Rcommunis | 30147.1000672
 Csativus | Cucsa.156880
 Vvinifera | GSVIVG01008655001
 Mesculenta | cassava4.1 031323m.g
 Csinensis | orange1.lg044766m.g
 Celestialinal | Ciclev1003942m.g
 Aeocerulea | Aquea 043 00044
 Mdomestica | MDP0000186703
 Mdomestica | MDP0000229889
 Mdomestica | MDP0000492089
 Lusitissimum | Lus10016732.g
 Lusitissimum | Lus10022427.g
 Egrandis | Euegr.K02192
 Stuberousum | PGSC0003DMG400014174
 Fvesca | gene33406-v1.0-hybrid
 Fvesca | gene30750-v1.0-hybrid
 Sycopersicum | Solyc03g118160.1
 Tcacao | ThecelEG014768
 Ppatens | Phpat.003G044500
 Mguttatus | Migut.100669
 Gmax | Glyma.06G163600
 Carya cathayensis
 WALNUT_00015719-RA

C

Fig. 5 Sequence analysis of the TATA-box (a) and TATA-less box (b) motif regions in different plant promoter regions. The transcription start site (TSS) is positioned at 0 and the predicted regions of the TATA and TATA-less regions are marked in between –30 and

–20 bp regions of the promoters. The weblogo picture (c) predicted the highly conserved sequences at the corresponding *box* region of the promoter

TC-rich elements. TC-rich elements might constitute a class of novel regulatory elements participating in the complex modulation of gene expression in plants (Bernard et al. 2010).

Sequence analysis of the *CcLFY* promoter

We analyzed the sequence of *CcLFY* promoter using PLACE databases, and predicted its key cis-acting elements and their locations. We did not find typical ‘TATA’ and ‘CAAT’ box motifs, but found ‘TATA-TATAT’ and ‘GTTTCAATGT’ at 186- and 252-bp upstream, respectively, of the initiation codon (ATG) of

the *CcLFY* open reading frame. We speculated that these two fragments function as a TATA-box and a CAAT-box, respectively. We also found a cis-acting element associated with the circadian cycle at –1242 bp. Two homologous sequences of anaerobic-related cis-acting elements (ARE) were located at –256 and –922 bp. One cis-acting regulatory element involved in methyl jasmonate (MeJA)-responsiveness was located at –575 bp, and one homologous sequence of LRT, which participates in low-temperature responsiveness, was located at –132 bp. One homologous sequence of an auxin-related cis-acting element AuxRR-core was located at –1283 bp (Fig. 1).

Table 1 Prediction of cis-elements of in *LFY* promoters in TATA-less and TATA-box using database analysis

Element	Sequence	Number of elements						Function
		TATA-less			TATA-box			
		Hickory	Walnut	Grape	Arabidopsis	Alyrata	Populus	
AAGAA-motif	GAAAGAA	1	2	0	3	4	0	Cis-acting element involved in light responsiveness
ACE	CTAACGTATT	1	0	0	0	1	1	
ARE	TGGTTT	2	2	3	0	0	0	Cis-acting regulatory element essential for anaerobic induction
ATC-motif	AGTAATCT	1	0	0	0	0	0	Part of conserved DNA module involved in light responsiveness
AuxRR-core	GGTCCAT	1	0	0	0	0	0	Cis-acting regulatory element involved in auxin responsiveness
Box 4	ATTAAT	3	2	3	3	1	13	Part of conserved DNA module involved in light responsiveness
CAT-box	GCCACT	1	0	2	0	0	2	Cis-acting regulatory element related to meristem expression
CGTCA-motif	CGTCA	1	1	2	2	0	2	Cis-acting regulatory element involved in MeJA-responsiveness
ELI-box3	AAACCAATT	1	0	0	0	0	0	Elicitor-responsive element
G-box	CACGTC	2	2	5	2	3	4	Cis-acting regulatory element involved in light responsiveness
GA-motif	AAGGAAGA	1	1	0	2	1	1	Part of a light responsive element
GT1-motif	GGTTAA	1	2	0	4	3	2	Light responsive element
LTR	CCGAAA	1	1	1	1	1	0	Cis-acting element involved in low-temperature responsiveness
Skn-1_motif	GTCAT	1	3	1	4	5	2	Cis-acting regulatory element required for endosperm expression
Sp1	CC(G/A)CCC	1	2	5	2	0	5	Light responsive element
TC-rich repeats	ATTTTCTCCA	1	1	1	2	1	3	Cis-acting element involved in defense and stress responsiveness
TCT-motif	TCTTAC	1	1	1	2	0	1	Part of a light responsive element
TGA-element	AACGAC	1	1	0	4	0	1	Auxin-responsive element
TGACG-motif	TGACG	1	1	2	2	0	2	Cis-acting regulatory element involved in the MeJA-responsiveness
Circadian	CAANNNNATC	1	0	1	1	1	2	Cis-acting regulatory element involved in circadian control

Although there has been extensive research on the *LFY* gene and its promoter in *Arabidopsis*, the *CcLFY* promoter was studied in this hickory plant for the first time. In this study, we compared the *CcLFY* promoter to *LFY* promoters in other plants using database analysis. Although low homology exists among the six promoters studied, the cis-acting elements predicted were nearly identical in all these promoters (Table 1).

Transient expression analysis in tobacco

We compared six *CcLFY* promoter constructs (1-C, L1, L2, L3, L4, and T1) for their ability to drive the GUS reporter gene (Fig. 2). The constructs were separately introduced into tobacco by *Agrobacterium*-mediated

gene transfer. All five of the deleted promoter constructs were able to drive GUS expression (Fig. 6a–e). GUS staining on the disc transformed with the 1-C construct was faint, but that on discs transformed with L1–L4 was stronger, especially that with L4 (Fig. 6). The promoter T1, which started at –185 bp, does not have the TATA and CAAT-box-like regions, but it was still able to drive GUS expression (Fig. 6f). To identify the key cis-acting element, we designed the promoter constructs SPL4 and SPT1, which had L4 and T1 elements deleted from the full-length promoter, respectively (Fig. 6g, h). As the regions surrounding the cis-acting element may also influence its activity, we deleted 42-bp from each construct, including the cis-element and its surrounding sequence. Transient expression analyses showed that

Fig. 6 Transient GUS expression analysis for the various promoter constructs. Various promoter constructs designed in this work were as follows **a** 1-C, **b** L1, **c** L2, **d** L3, **e** L4, **f** T1, **g** SPT1, **h** SPL4, **i** W, **j** SSPL4, **k** SSPT1, **l** PBI35S. Tobacco leaf discs were incubated with different constructs transformed *Agrobacterium* strains for 15 min and then histochemical staining was performed to visualize the GUS expression in the transformed leaf discs

both constructs were able to drive GUS expression, but the intensity of the GUS staining was similar to that in 1-C. The promoter SPL4 lacked the putative TATA-box region, but was able to drive expression of the GUS gene.

In order to determine which one of the cis-acting element is essential, we delete the different regions of the promoter (L4 and T1), and thus creating the SSPL4 and SSPT1 constructs, respectively (Fig. 6j, k). These constructs lacked the inhibitory trans-regulatory factor binding sites, and both showed stronger GUS staining than the other promoters when expressed in tobacco. The ability of SSPL4 to drive the GUS gene was stronger than that of SSPT1. A promoter construct was designed that started from the end of the T1 element to determine the location of the region that substituted for activity of the TATA-box. This promoter, designated as W, consists of the region from –151 bp and contained no cis-acting elements (Fig. 6i). The W construct was able to drive GUS gene

Fig. 7 GUS activity measurement for the full length promoter construct (1-C) under varying temperature (25 °C) and light (dark) conditions

expression in the transient expression analyses and it is the simplest structure to drive gene expression in this experiment. The strongest expression of GUS was achieved using the positive control promoter, PBI35S (Fig. 6i).

Fluorometric analysis of GUS activities in tobacco under different environment conditions

When the tobacco leaf discs were transformed with the full-length promoter constructs and incubated in darkness for 24 h at 25 °C, the GUS enzymatic activity was 8.93 units. When the discs were kept under dark with low temperature for 8, 16, and 24 h, the activity was found to be 12.84, 16.28, and 17.77 units, respectively. Under low temperature and light provided conditions, the measured GUS activity was found to be 12.22, 15.91, and 18.24 units, respectively (Fig. 7).

To further study the roles of the various promoter segments on gene expression, we conducted fluorometric analyses of GUS activities in tobacco leaves under normal and different temperature and light treatment conditions. The GUS activity was 8.92 units with the 1-C construct, 10.11 units with L1, 10.18 with L2, 10.28 with L3, 10.84 with L4, and 9.5 with T1. For all the promoter constructs, GUS enzymatic activity was significantly induced by both low temperature and light, but the differences among the various promoters diminished over time (Fig. 8).

The different promoter constructs such as SPL4, SSPL4, and W promoted more GUS expression than 1-C in the control group, but less than that of PBI121 (Fig. 8). In light, the activity of GUS increased over time. The most rapid increase was with SPL4; GUS activity increased to 1.47, 1.77, and 2.2 times its initial activity at 8, 16, and 24 h, respectively. The second most rapid increase was with 1-C. The highest GUS activity (20.56) was observed after 24 h with the SSPL4 promoter construct. In the low

Fig. 8 Fluorimetric analysis of GUS activity in tobacco leaves incubated with different promoter constructs. Assays were performed under normal, varying temperature (25 °C), and light (dark) conditions at different time intervals of 0, 8, 16 and 24 h. Error bars at 8, 16 and 24 h indicate the standard deviation for three replicates. Assays in all the treatments include the full length promoter (*I-C*), and 5' end deleted *I-C* promoter (*T1*). For light treatment, the constructs deleted with *L4* cis-acting elements (*SPL4*, *SSPL4*) and for temperature treatment constructs deleted with *T1* cis-acting elements (*SPT1* and *SSPT1*) were included along with *W* construct

temperature treatment, the GUS activity also increased over time in the transformed promoters. The most rapid increase was with *I-C*, which increased GUS activity to 1.43, 1.82, and 2.04 times its initial activity at 8, 16, and 24 h, respectively. The next most rapid increase was with

the *SSPT1* construct. The highest GUS activity (20.04) was after 24 h of cold treatment with the *SSPT1* promoter construct (Fig. 8).

Discussion

Promoters mediate the transcription of many genes during developmental processes of the plants under both optimum and under environmental stress conditions and thus they are the key mediators during developmental and stress causing conditions (Bernard et al. 2010; Molina and Grotewold 2005). In this work, the characterization of the *CcLFY* promoter was carried out in order to understand the gene expression during the flowering stage and also to study the promoter influence under stress causing conditions. The *CcLFY* gene was strongly expressed in flower buds and leaves, but only slightly expressed in stems, and was undetectable in root tissue. Sequence analyses suggest that the promoter harbors various regions for the developmental process of the plants and also the regions for the stress causing conditions. Our sequence analysis, predict that *CcLFY* promoter exhibit circadian cycle control region (−1242 bp) (Fig. 1). Various environmental stress responsive regions such as oxygen, low temperature, some plant hormones such as abscisic acid (ABA), MeJA and light-responsive elements are also present in the promoter sequence (Table 1). One of the *LFY* gene functions are described to play a major role in plants development (Tanahashi et al. 2005). Expression of *CcLFY* began during the transition from vegetative to floral meristems, and continued to increase and peaked 16 days after bloom (Wang et al. 2012). This is the time period (mid of April) when the majority of floral buds in Hickory are formed. Thus, characterizing the *CcLFY* promoter is an extremely valuable approach for biotechnological applications aimed at shortening the juvenile phase.

Previous studies have shown that *LFY* exist as copy in almost all the plants, even in the polyploid species (Riechmann and Ratcliffe 2000; Shiu et al. 2005). Phylogenetic analysis of the sequence of *LFY* genes showed that *LFY* gene is found in primitive to higher order plants and the analyses revealed that monocotyledons and gymnosperms share close evolutionary relationships than that of the dicotyledons and gymnosperms (Fig. S1). This indicates that this gene may have important roles for the growth of all plants (Moyroud et al. 2010; Shiokawa et al. 2008). The presence of *LFY* in moss may have important role for cell division and meristem development than the later acquired floral function (Tanahashi et al. 2005). Among the different plants, *CcLFY* gene from hickory and *LFY* gene from walnut shared the same clade of evolutionary relationship (Fig. S1). *CcLFY* gene exists as a

single copy in the hickory genome (Wang et al. 2012), and this was the same for walnut. The copy number of *LFY* genes present in individual plants varies and even among the different copies of *LFY* genes, only one possess TATA-box, except in maize. Therefore, studying the TATA-box of *LFY* homologous gene promoters can predict evolutionary relationships between plants. Compared with the full length promoter, P region is highly conserved among all plants mainly based on the AuxRE motif (Fig. 4). The cis-acting elements predicted in the *CcLFY* promoter and the *LFY* promoters from other plants were nearly identical, which showed the high level of *LFY* promoter's conservation across species. Cis-acting regulatory elements are important molecular switches involved in the transcriptional regulation of dynamic networks of genes. The activities of cis-acting regulatory element mainly control the biological processes, that include responses to abiotic stresses, plant growth regulators, and development (Guo et al. 2010).

TATA-box presence in the promoter is associated with the precise initiation of transcription and the CAAT-box plays a major role in controlling the initiation of transcription (Zhu et al. 2002). Transient expression analyses using the truncated promoter constructs which lack the putative TATA-box and CAAT-box regions, could able to drive the GUS expression (Fig. 6) which shows that *CcLFY* promoter is a TATA-less promoter. It shows that it has alternative mechanism to the TATA-box for binding to RNA polymerase II (Fridlender et al. 1996). TATA-less promoters were reported in human (Uchiumi et al. 2011) and in *Saccharomyces* spp. (Erb and Van Nimwegen 2011), and such promoters contain alternative elements or motifs. Based on the deletion analysis for the various regions of the promoter we assumed that the alternative mechanism may be located at 151-bp region of the 5' flanking sequence (Fig. 1).

Transcriptome expression profiling technologies have identified various combinations of cis-acting elements in promoters of stress-inducible genes (Yamaguchi-Shinozaki and Shinozaki 2005). GUS-assay experiments with several deletion constructs of the *CcLFY* promoters (L1–L4) increased the GUS expression. The data suggest that some trans-regulatory factor binding sites might exist between –1302 and –958 bp and from –592 to –315 bp, because suppression was eliminated after these regions were deleted. Promoter constructs SPL4, SPT1, SSPL4, and SSPT1 were designed to identify the essential cis-acting elements. We anticipated that one or some of these promoters would result in sharp decreases in GUS activity, but was not observed. It was interesting to note that SSPT1, which lacked all of the predicted elements, was able to drive expression of the GUS gene (Fig. 6). Therefore, the key cis-acting elements were not identified. Two-fold

increased GUS enzymatic activity during low temperature and light treatment conditions, suggest that under natural conditions, low temperature and light could induce the *CcLFY* promoter to drive floral gene expression. In the GUS assays, the GUS activity in the leaf discs diminished at later time of analysis and this may be due to the accumulation of the GUS protein in plant tissue and limitation of the promoters' activities.

Conclusion

In this work, in depth sequence analysis of the *CcLFY* promoter and expression studies were carried out in order to understand the optimum environmental conditions for the hickory flowering. Our sequence analysis shows that *LFY* promoters reported from Hickory and walnut share same clade of evolutionary relationship and also show other similarities based on the conservation of AuxRE in the P region and also the location of the TATA-box region. GUS expression analysis using different deletion constructs of *CcLFY* promoter suggests that the promoter reported in this work is a kind of TATA-less promoter and its activity is highly influenced by low temperature and dark conditions.

Acknowledgements This work was supported by the Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholar (LR14C160002), the National High Technology Research and Development Program 863 (2013AA102605), the National Natural Science Foundation of China (30872047, 31170637, 31070604, and 31570666), the Zhejiang Provincial New Varieties Breeding Major Agricultural Science and Technology Projects (2012C12904) and the initial project of the National Basic Research Program of China (2001CB111510).

References

- Albert VA, Oppenheimer DG, Lindqvist C (2002) Pleiotropy, redundancy and the evolution of flowers. *Trends Plant Sci* 7:297–301
- An G, Mitra A, Choi HK, Costa MA, An K, Thornburg RW, Ryan CA (1989) Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. *Plant Cell* 1:115–122
- Bernard V, Brunaud V, Lecharny A (2010) TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genom* 11:1
- Blázquez MA, Weigel D (2000) Integration of floral inductive signals in Arabidopsis. *Nature* 404:889–892
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
- Dornelas MC, Rodriguez APM (2005a) A *FLORICAULA/LEAFY* gene homolog is preferentially expressed in developing female cones of the tropical pine *Pinus caribaea* var. *caribaea*. *Genet Mol Biol* 28:299–307

- Dornelas MC, Rodriguez APM (2005b) The rubber tree (*Hevea brasiliensis* Muell. Arg.) homologue of the *LEAFY/FLORICAULA* gene is preferentially expressed in both male and female floral meristems. *J Exp Bot* 56:1965–1974
- Dornelas MC, Rodriguez APM (2006) The tropical cedar tree (*Cedrela fissilis* Vell., Meliaceae) homolog of the Arabidopsis *LEAFY* gene is expressed in reproductive tissues and can complement arabidopsis leafy mutants. *Planta* 223:306–314
- Du N, Pijut PM (2010) Isolation and characterization of an AGAMOUS homolog from *Fraxinus pennsylvanica*. *Plant Mol Biol Report* 28:344–351
- Erb I, Van Nimwegen E (2011) Transcription factor binding site positioning in yeast: proximal promoter motifs characterize TATA-less promoters. *PLoS ONE* 6:e24279
- Fridlender M, Harrison K, Jones JD, Levy AA (1996) Repression of the Ac-transposase gene promoter by Ac transposase. *Plant J* 9:911–917
- Goto F, Yoshihara T, Shigemoto N, Toki S, Takaiwa F (1999) Iron fortification of rice seed by the soybean ferritin gene. *Nat Biotechnol* 17:282–286
- Guo L, Yu Y, Xia X, Yin W (2010) Identification and functional characterisation of the promoter of the calcium sensor gene *CBL1* from the xerophyte *Ammopiptanthus mongolicus*. *BMC Plant Biol* 10:1
- Hamès C et al (2008) Structural basis for *LEAFY* floral switch function and similarity with helix-turn-helix proteins. *EMBO J* 27:2628–2637
- Horsch R, Fry J, Hoffmann N, Eichholtz D, Rogers SA, Fraley R (1985) A simple and general method for transferring genes into plants. *Science* 227:1229–1231
- Hou J-H, Gao Z-H, Zhang Z, Chen S-M, Ando T, Zhang J-Y, Wang X-W (2011) Isolation and characterization of an AGAMOUS homologue *PmAG* from the Japanese apricot (*Prunus mume* Sieb. et Zucc.). *Plant Mol Biol Report* 29:473–480
- Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* 6:3901
- Lee J, Lee I (2010) Regulation and function of *SOC1*, a flowering pathway integrator. *J Exp Bot* 61:2247–2254
- Li H-Y, Liu F-F, Liu G-F, Wang S, Guo X-H, Jing J (2012) Molecular cloning and expression analysis of 13 MADS-box genes in *Betula platyphylla*. *Plant Mol Biol Report* 30:149–157
- Liu C, Thong Z, Yu H (2009) Coming into bloom: the specification of floral meristems. *Development* 136:3379–3391
- Liu X, Anderson JM, Pijut PM (2010) Cloning and characterization of *Prunus serotina* AGAMOUS, a putative flower homeotic gene. *Plant Mol Biol Report* 28:193–203
- Ma G, Ning G, Zhang W, Zhan J, Lv H, Bao M (2011) Overexpression of *Petunia SOC1*-like gene *FBP21* in tobacco promotes flowering without decreasing flower or fruit quantity. *Plant Mol Biol Report* 29:573–581
- Maizel A, Busch MA, Tanahashi T, Perkovic J, Kato M, Hasebe M, Weigel D (2005) The floral regulator *LEAFY* evolves by substitutions in the DNA binding domain. *Science* 308:260–263
- Molina C, Grotewold E (2005) Genome wide analysis of arabidopsis core promoters. *BMC Genom* 6:1
- Mouradov A, Cremer F, Coupland G (2002) Control of flowering time interacting pathways as a basis for diversity. *Plant Cell* 14:S111–S130
- Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) *LEAFY* blossoms. *Trends Plant Sci* 15:346–352
- Ponjavic J, Lenhard B, Kai C, Kawai J, Carninci P, Hayashizaki Y, Sandelin A (2006) Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. *Genome Biol* 7:R78
- Riechmann JL, Ratcliffe OJ (2000) A genomic perspective on plant transcription factors. *Curr Opin Plant Biol* 3:423–434
- Shiokawa T et al (2008) Isolation and functional analysis of the *CjNdy* gene, a homolog in *Cryptomeria japonica* of *FLORICAULA/LEAFY* genes. *Tree Physiol* 28:21–28
- Shiu S-H, Shih M-C, Li W-H (2005) Transcription factor families have much higher expansion rates in plants than in animals. *Plant Physiol* 139:18–26
- Simpson GG, Dean C (2002) Arabidopsis, the Rosetta stone of flowering time? *Science* 296:285–289
- Tanahashi T, Sumikawa N, Kato M, Hasebe M (2005) Diversification of gene function: homologs of the floral regulator *FLO/LFY* control the first zygotic cell division in the moss *Physcomitrella patens*. *Development* 132:1727–1736
- Theissen G, Melzer R (2007) Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Ann Bot* 100:603–619
- Uchiumi F, Miyazaki S, S-i Tanuma (2011) The possible functions of duplicated ets (GGAA) motifs located near transcription start sites of various human genes. *Cell Mol Life Sci* 68:2039–2051
- Wang ZJ, Huang JQ, Huang YJ, Chen FF, Zheng BS (2012) Cloning and characterization of a homologue of the *FLORICAULA/LEAFY* gene in Hickory (*Carya cathayensis* Sarg). *Plant Mol Biol Report* 30:794–805
- Weigel D, Nilsson O (1995) A developmental switch sufficient for flower initiation in diverse plants. *Nature* 377:495–500
- Xu D, McElroy D, Thornburg RW, Wu R (1993) Systemic induction of a potato pin2 promoter by wounding, methyl jasmonate, and abscisic acid in transgenic rice plants. *Plant Mol Biol* 22:573–588
- Yamaguchi N et al (2013) A molecular framework for auxin-mediated initiation of flower primordia. *Dev Cell* 24:271–282
- Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends Plant Sci* 10:88–94
- Zhu Y, Li Y, Zheng X (2002) Modern molecular biology. Beijing Higher Education Press, Beijing