Plant Science xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Functional characterization of *GhSOC1* and *GhMADS42* homologs from upland cotton (*Gossypium hirsutum* L.)

Xiaohong Zhang[a,b], Jianghui wei[a,b], Shuli Fan[b], Meizhen Song[b], Chaoyou Pang[b], Hengling Wei[b], Chengshe Wang[a,**], Shuxun Yu[a,b,*]

[a] College of Agronomy, Northwest A&F University, Yangling, 712100 Shaanxi, People's Republic of China
[b] State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang, 455000 Henan, People's Republic of China

ARTICLE INFO

Article history:
Available online xxx

Keywords:
GhSOC1
GhMADS42
Arabidopsis
Gossypium hirsutum
ChIP
Yeast two-hybrid

ABSTRACT

In *Arabidopsis* flowering pathway, MADS-box genes encode transcription factors, with their structures and functions highly conserved in many species. In our study, two MADS-box genes *GhSOC1* and *GhMADS42* (*Gossypium hirsutum* L.) were cloned from upland cotton CCRI36 and transformed into *Arabidopsis*. *GhSOC1* was additionally transformed into upland cotton. Comparative analysis demonstrated sequence conservation between *GhSOC1* and *GhMADS42* and genes of other plant species. Tissue-specific expression analysis of *GhSOC1* and *GhMADS42* revealed spatiotemporal expression patterns involving high transcript levels in leaves, shoot apical buds, and flowers. In addition, overexpression of both *GhSOC1* and *GhMADS42* in *Arabidopsis* accelerated flowering, with *GhMADS42* transgenic plants showing abnormal floral organ phenotypes. Overexpression of *GhSOC1* in upland cotton also produced variations in floral organs. Furthermore, chromatin immunoprecipitation assay demonstrated that GhSOC1 could regulate *GhMADS41* and *GhMADS42*, but not *FLOWERING LOCUS T*, by directly binding to the genes promoter. Finally, yeast two-hybrid and bimolecular fluorescence complementation approaches were undertaken to better understand the interaction of GhSOC1 and other MADS-box factors. These experiments showed that GhSOC1 can interact with APETALA1/FRUITFULL-like proteins in cotton.

1. Introduction

Cotton (*Gossypium spp.*) is one of the most important natural textile fiber crops. In addition, cotton seed is a worldwide source of oil and protein meal. *Gossypium hirsutum* L., a tetraploid species known as upland cotton, is thought to arisen via an allopolyploidization event that occurred approximately 1–2 million years ago involving a D-genome species *Gossypium raimondii* as the pollen-providing parent and an A-genome species *Gossypium arboreum* as the maternal parent [1,2]. New insights into *Gossypium* biology have been provided by the whole genome sequencing of these latter two species [3,4]. The availability of *G. raimondii* and *G. arboreum* reference sequences and the high-throughput sequencing of *G. hirsutum* have revealed a great deal of information about tetraploid cotton genome structure and polyploid evolution, and are essential for the identification, isolation and manipulation of important cotton genes conferring agronomic traits for molecular breeding and genetic improvement [5].

Upland cotton is a short day cultivated cotton species whose sensitivity to the day length was lost during domestication [6]. Because of conflicts in land use involving grain vs. cotton cultivation, early maturation has become a key trait in short-season cotton breeding programs in China [7]. Short-season cotton varieties are suitable for wheat, barley, rape and cotton double cropping farm practice in Yellow River and Yangzi River cotton areas of China and the short frost-free regions of northwestern China and the American high plains [8]. The time and concentration of flowering affects the earliness of short-season cotton varieties. In the flowering plant life, the transition from vegetative to reproductive growth is

Abbreviations: SOC1, SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1; AP1, APETALA1; FUL, FRUITFULL; SVP, SHORT VEGETATIVE PHASE; FT, FLOWERING LOCUS T; CO, CONSTANS; *FLC*, *FLOWERING LOCUS C*; GA, gibberellic acid; RIN, RIPENING INHIBITOR; Col-0, Colombia 0; RT-PCR, reverse transcription-PCR; ORF, open reading frame; ChIP, chromatin immunoprecipitation; DAS, days after sowing; X-α-galactosidase, X-α-Gal; BiFC, bimolecular fluorescence complementation; LD, long day.

* Corresponding author at: College of Agronomy, Northwest A&F University, Yangling, 712100 Shaanxi, People's Republic of China. Tel.: +86 0372 2562201; fax: +86 372 2562256.
** Corresponding author. Tel.: +86 029 87032198; fax: +86 029 87082845.
E-mail addresses: xiaohongzhang910112@gmail.com (X. Zhang), jiang850515@163.com (J. wei), fasl@cricaas.com.cn (S. Fan), songmzccri@163.com (M. Song), chypang@163.com (C. Pang), henglingwei@163.com (H. Wei), wangcs2008@126.com (C. Wang), yu@cricaas.com.cn (S. Yu).

http://dx.doi.org/10.1016/j.plantsci.2015.05.001
0168-9452/

regulated by both endogenous and exogenous signals [9–11]. Extensive genetic analyses to elucidate the molecular mechanisms of flowering in *Arabidopsis* have revealed several flowering pathways controlling flowering time, such as photoperiod, vernalization, gibberellin and autonomous pathways [12–17].

Among floral transition integrators, several members of the MADS-box family, such as *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *APETALA1* (*AP1*)/*FRUITFULL* (*FUL*) and *SHORT VEGETATIVE PHASE* (*SVP*), have major roles in *Arabidopsis* and other species. *SOC1* belongs to the F class gene group, an extension of the ABC and ABCDE models of floral development. In *Arabidopsis*, *SOC1* is regulated by *FLOWERING LOCUS C* (*FLC*), *FLOWERING LOCUS T* (*FT*) and *CONSTANS* (*CO*) genes, which act together with *FLOWERING LOCUS T* (*FT*) to determine precise flowering time [18]. However, ectopic expression of a *SOC1* homolog in *Gerbera hybrida* has been found to lead to phenotypic changes, such as a partial loss of floral organ identity, but does not affect flowering time [19]. Other *SOC1*-like genes are flowering promoters that have played an evolutionarily conserved role in floral regulation of strawberry and orchidaceae family [20,21]. Two recent studies of *SOC1* DNA-binding profiles using chromatin immunoprecipitation (ChIP)-chip and ChIP sequencing have revealed that SOC1 activates its own transcription and regulates the *SPL3*, *SPL4* and *SPL5* genes by directly binding to the conserved CArG motifs of their promoters [22,23].

In *Arabidopsis*, the MADS-box gene *AP1* shared overlapping roles in promoting flower meristem identity while preventing the formation of flowers in the axils of sepals and suppressing of cytokinin biosynthesis and activation of cytokinin degradation [24]. *AP1* also controls sepal, carpel, stamen and petal formation in herbaceous peony. However, the ortholog of *AP1* in trees has an evolutionary novel function in photoperiodic regulation of seasonal growth [25,26]. Transcripts of another MADS-box gene, *AGL8*/*FUL*, accumulates at high levels in the inflorescence apical meristem, in the inflorescence stem and cauline leaves. This gene also regulates the transcription of genes required for cellular differentiation during fruit and leaf development in *Arabidopsis* [27,28]. *FUL* genes control phase transition, cauline leaf growth, compound leaf morphogenesis and fruit development in eudicot [29]. In tomato, FUL proteins can regulate fruit ripening through ethylene biosynthesis and interact with the master regulator MADS-box protein *RIPENING INHIBITOR* (*RIN*) [30,31].

In *Arabidopsis* and rice, the *SVP* gene encoding a MADS-domain transcription factor is broadly expressed in leaves and shoot apices during vegetative development and functions as a flowering time repressor [32,33]. *SVP* delay flowering by repressing gibberellic acid (GA) biosynthesis and integrate gene expression in the shoot apical buds [34]. Over-expression of *MtSVP1* has been found to alter floral development but not flowering time in *Medicago truncatula*, which suggests that this gene does not repress the floral transition [35]. During the floral transition, the sequential formation of FUL-SVP and FUL-SOC1 heterodimers may mediate the vegetative and meristem identity transitions, counteracting the repressive effect of *FLC* and *SVP* on flowering [36].

Upland cotton is a photoperiod-insensitive perennial in nature. In previous studies, two *SQUAMOSA-like* genes and the flowering promoter factor *GhFPF1* were cloned from *G. hirsutum* and analyzed in transgenic *Arabidopsis* [37]. The *GhLFY* and *GhSPL* gene family has been identified and analyzed, and GhSOC1 has been confirmed to bind to the promoters of *GhSPL3* and *GhLFY* to regulate flowering [38,39]. Although several studies have described the MADS-box genes of *G. hirsutum*, little work has been carried out on the flowering pathway regulatory mechanism operating in this species. The purpose of this study was consequently to determine the function and regulation of *GhSOC1* and *GhMADS42*, which should be useful information for the development of early-maturing cotton varieties.

2. Materials and methods

2.1. Plant materials and growth conditions

Samples were collected from plants of *G. hirsutum* 'CCRI36', a short-season Chinese variety of upland cotton grown, that were cultivated on the farm of the Cotton Research Institute of the Chinese Academy of Agricultural Sciences, Anyang, Henan. We sampled roots, stems, leaves, shoot apical buds, full-bloom flowers at anthesis, ovules and fiber 10 days post anthesis. Wild type *Arabidopsis thaliana* Colombia 0 (Col-0) was grown in the greenhouse at 22 °C under long day (LD) condition (16/8 h light/dark photoperiod).

2.2. Gene cloning, vector construction and transformation

We screened a cotton cDNA library constructed in our laboratory. We designed specific primers to amplify the open reading frames (ORFs) of *GhSOC1* and *GhMADS42* genes using cDNA templates prepared from different tissues of CCRI36. PrimeSTAR GXL DNA polymerase (TaKaRa Tokyo, Japan) was used to amplify the two genes using the following cycling profile: 98 °C for 1 min, followed by 30 cycles of 98 °C for 10 s, 55 °C for 15 s and 68 °C for 1 min. The amplified products were cloned into a pBI121 vector (Clontech, Palo Alto, CA, USA) and sequenced from both ends. *Arabidopsis* plants were transformed using the *Agrobacterium*-mediated gene transfer method described previously [40]. Cotton 'CCRI24' hypocotyl explants were transformed with GhSOC1 fusion genes via *Agrobacterium*-mediated DNA transfer through a series of co-cultivations involving callus, differentiated callus, embryogenic callus, plantlets and grafted plants [41]. Transgenic *Arabidopsis* plants were selected with kanamycin, and flowering times in T3 generation were monitored by counting rosettes and caulines leaves on main inflorescences. Samples for quantitative real-time PCR (qRT-PCR) were harvested 14 days after sowing (DAS). Leaf samples of transgenic cotton plants from T0 and T1 generation were harvested at the third leaf expansion stage. Statistical testing was performed using one-way analysis of variance Duncan's method.

2.3. Sequence alignment and phylogenetic analysis

The amino acid sequences of the proteins used in this study were downloaded from GenBank. Accession numbers of all species are listed in Supplementary Table S1. Multiple sequence alignment was performed using ClustalW (http://www.ebi.ac.uk). A phylogenetic tree was constructed using the neighbor joining method in Molecular Evolutionary Genetics Analysis (MEGA) software MEGA5.05 [42]. The reliability of nodes in the tree was evaluated by bootstrapping with 1000 replicates.

2.4. Quantitative real-time PCR

Total RNA was isolated from samples using a Plant RNA purification kit (Tiangen, Beijing, China). Reverse transcription-PCR was carried out using a SuperScript III First-Stand Synthesis System for RT-PCR (Invitrogen, Carlsbad, USA). Transcript levels were then determined by qRT-PCR using a 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) and SYBR Premix Ex Taq (2×) (TaKaRa). Gene-specific primer pairs used for the PCR amplifications are listed in Supplementary Table S2. Determination of reaction specificities and data processing were performed as described previously [39,43]. Three biological replicates were analyzed.

Please cite this article in press as: X. Zhang, et al., Functional characterization of *GhSOC1* and *GhMADS42* homologs from upland cotton (*Gossypium hirsutum* L.), Plant Sci. (2015), http://dx.doi.org/10.1016/j.plantsci.2015.05.001