The MADS transcription factor GhAP1 coordinates the flowering regulatory pathway in upland cotton (Gossypium hirsutum L.)

Xiaoqian Cheng  
Chinese Academy of Agricultural Sciences Cotton Research Institute

Hantao Wang  
Cotton Research Institute

Hengling Wei  
Cotton Research Institute

Lijiao Gu  
Cotton Research Institute

Pengbo Hao  
Cotton Research Institute

Huiru Sun  
Cotton Research Institute

Aimin Wu  
Cotton Research Institute

Shuaishuai Cheng  
Cotton Research Institute

Jinping Hua  
china agricultural university

Shuxun Yu  
Chinese Academy of Agricultural Sciences Cotton Research Institute  https://orcid.org/0000-0002-9715-3462

Research article

Keywords: GhAP1, Arabidopsis, VIGS, Yeast, Cotton, Early flowering

Posted Date: December 3rd, 2019

DOI: https://doi.org/10.21203/rs.2.18061/v1

License: ☺️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background: MADS-box gene family plays an important role in the molecular regulatory network of flower development. APETAL1 (AP1), a MADS-box gene, plays an important role in the development of flower organs. Although many studies about MADS-box family genes have been reported, the function of AP1 is still not clear in cotton. Results: In this study, GhAP1 (Gh_D03G0922), a candidate gene for cotton flower time and plant height obtained from our previous studies, was cloned from CCRI50 cotton variety and functionally characterized. Subcellular localization demonstrated that GhAP1 was located in nucleus. Infection test of Arabidopsis revealed that GhAP1 could cause precocious flowering and virus-induced gene silence (VIGS) assay demonstrated that GhAP1 could lead to delayed flowering of cotton plants. Yeast one-hybrid assays and transient dual-luciferase assays suggested that floral meristem identity control gene LEAFY (LFY) can bind the promoter of GhAP1 and negatively regulate it. Yeast two-hybrid assays suggested that GhAP1 can interact with pyridoxal phosphate (PLP)-dependent transferases superfamily protein PSAT2. Conclusions: Our research indicated that GhAP1 might work as a positive regulator in plant flowering. Moreover, GhAP1 may interact with GhPSAT2 and be negatively regulated by GhLFY in the regulatory pathways. This work laid the foundation for subsequent functional studies of GhAP1.

Background

Flowering is one of the important traits in plants. In 1990, Yanofsky et al. cloned an AGAMOUS gene for floral organ development from Arabidopsis for the first time, then research on flower development has developed rapidly [1]. Usually, the flowers of typical dicotyledonous plants consist of sepal, petal, stamen, pistil and carpel. These five floral organs are controlled by organ characteristic genes. The flower growth ABC model revealed the molecular mechanism underlying flower development which was accepted generally [2]. With the deepening of research, D functional genes were found in Petunia hybrida and E functional genes were found in Gerbera hybrida and Oryza sativa [3-5], then ABCDE model or AE model about flower was derived.

MADS-box genes are widely distributed in plants, and more than one hundred genes are found in Arabidopsis [6, 7]. Based on the structural differences, MADS-box genes can be divided into Type and Type [8]. Type genes have a high conserved SRF-like MADS domain and their functions are only reported in Arabidopsis [9]. Type genes also called MIKC type genes and they contain four domains: MEF2-like MADS, Intervening (I), Keratin (K), and C-terminal (C) domains. MADS domain is located in N-terminal consisting of 56 amino acids and is the most conservative domain [10]. I domain is less conservative, which is related with combination of DNA and protein [11]. According to different structures of I domain, these genes can be subdivided into MIKC and MIKC [12, 13]. K domain is a semi-conservative region, it is homologous to the helix of keratin and has about 70 amino acids [14] and is the signature sequence of MIKC type MADS-box genes. Expect for individual genes of maize, all the MADS-box genes contain K domain in plants [15]. The C domain is the least conservative region comes after the K domain, which is a transcriptional activation region composed of hydrophobic amino acids [16].
Different C domains can lead to different functions [17]. The MADS-box gene family encodes transcription factors (TFs) that share a highly conserved domain which can bind CArG box in promoter region of the MADS-box protein target genes [18].

MADS-box genes play important regulatory roles in plant growth and development. Most of them play important roles in the development of plant reproductive organs and some of them participate in the development of plant vegetative organs [19]. In different species, MADS-box genes have different degree of sub-functionalization but most of them are related to flowering. As a member of the MADS-box gene family, the function of AP1 in plants is very conservative. FUL (AGL8) promotes the initiation of flower development with SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) in the regulation of flowering time in Arabidopsis [20]. In the photoperiod dependent flowering pathway, FUL is involved in regulating the transformation of bud meristems into inflorescence meristems, and then the expression of FUL and SOC1 in the developed inflorescence procambium induces the transformation to flower meristems, thus controlling the flowering time [21]. Recent studies showed that PasAP1 is expressed during early development of tendrils with the A function. PasAP1 appears to play an important role in the tendril development and in maintaining the tissues with meristem characteristics [22]. NnAP1 shares 75-96% similarity with other plant AP1s and it plays an important role in floral meristem and floral organ formation [23]. AP1 was identified as a floral meristem related gene that regulates flowering in many plant species, including Moso bamboo, Cymbidium ssp and Dendrobium nobile species [24-26]. However, few studies have been done in cotton.

Cotton is the one of most important economic crop that is widely cultivated throughout the world [27]. The MADS-box genes have been extensively studied in cotton. Sarah Muniz Nardeli et al. identified a total of 207 MADS-box genes in tetraploid cotton, which opens an avenue to understanding the origin and evolution of each gene subfamily within diploid and polyploid species and paves the way for functional studies in cotton species [28]. The MIKC-type genes of MADS-box family in upland cotton were identified by Zhongying Ren et al. [9]. Additionally, 110 GhMIKC genes were identified and phylogenetically classified into 13 subfamilies, however, the Flowering locus C (FLC) subfamily was absent in the upland cotton [9]. It was found that GhmiR157/SPL promoted the development process of flower organs by regulating MADS-box genes [29]. Overexpression of both GhSOC1 and GhMADS42 in Arabidopsis accelerated flowering, and GhSOC1 can interact with AP1 protein in cotton [30]. Besides, GhMADS3 [31], GhMADS9 [32], GhMADS11 [33] have been studied in cotton.

Early maturity is one of the important traits of cotton. Early-maturing cotton is helpful for realizing both grain and cotton maturity, reducing the land conflict between grain and cotton. Flowering is one of the most important traits of early maturity. Many studies have shown that MADS TFs play important roles in all the stages of plant growth, especially flowering. Therefore, it is important to study the functional and molecular mechanism of MADS genes in cotton. In our previous studies of our lab, a MIKC type MADS gene, GhAP1 was identified as a candidate gene for cotton flower time [34] and plant height [35]. In this study, GhAP1 was cloned and characterized. The overexpression lines of GhAP1 in Arabidopsis and the VIGS plants of GhAP1 in cotton were used to explore its biological function. Yeast one-hybrid assays,
transient dual-luciferase assays and yeast two-hybrid assays were performed to examine the molecular mechanism of *GhAP1*. Our findings help to enrich flowering regulatory pathways and provide genetic resources for genetic improvement of cotton varieties.

**Results**

**Gene structure and protein sequence analysis of GhAP1**

Based on the results of genome-wide association study (GWAS) in our previous studies, *GhAP1* (Gh_D03G0922) was selected according to gene annotation, and its *Arabidopsis* homologs contain MADS-box domains [35]. *GhAP1* was located on D03 chromosome and its *Arabidopsis* homolog gene is Agamous-like MADS-box protein *AGL8* (AT5G60910). The protein sequence similarity between GhAP1 and AGL8 was 56.2%. The genome DNA length was 5452 bp and the CDS length was 690 bp, encoding 229 amino acids (Fig. 1a). In addition, *GhAP1* contained eight exons and seven introns (Fig. 1a). Protein sequence analysis showed its homologous genes of different species shared the high conserved MADS domain and the signature K-domain of MIKC type MADS-box genes. A putative nuclear localization signal (NLS) sequence (ALKRIRSRKNQLM) was predicted in GhAP1 protein sequence (Fig. 1b).

**Expression analysis of GhAP1 in different tissues and different varieties**

Quantitative real-time PCR (qRT-PCR) was performed to detect the expression levels of *GhAP1* in roots, stems, leaves, buds, fiber, petals, and sepals. *GhAP1* was found to be differentially expressed in different tissues. *GhAP1* was strongly expressed in sepals and leaves but was weakly expressed in petals, roots, and fiber (Fig. 2a). To evaluate the expression pattern of *GhAP1* during different growth stages, the expression level of *GhAP1* was examined in the leaves of two early maturing varieties (CCRI50 and CCRI74) and two late maturing varieties (BM and G11) at different growth stages. The results revealed that the expression level of *GhAP1* reached its maximum level in CCRI50 plants at the three-leaf stage. In addition, the expression levels of *GhAP1* increased gradually and then decrease during growth stages and were higher in early maturing varieties than in late maturing varieties (Fig. 2b).

**GhAP1 was localized to the nucleus**

GhAP1 was predicted by subcellular location software Plant-mPloc ([http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/](http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/)) and the results showed that the GhAP1 protein was localized to the nucleus. To confirm our prediction, the 35S-GhAP1::GFP vector was constructed and injected into tobacco leaves. Using a fluorescence microscope, we found that the tobacco leaves harboring the 35S-GhAP1::GFP construct emitted green fluorescence predominantly in nuclei under blue light (Fig. 3a, 3b).

**Overexpression of GhAP1 promoted flowering in Arabidopsis**

To further study the function of *GhAP1*, the overexpression vector 35S::GhAP1 was constructed and transformed into *Arabidopsis*. qRT-PCR results showed that *GhAP1* gene was hardly expressed in WT, but
was highly expressed in transgenic plants, which was significantly higher in transgenic plants than that of non-transgenic *Arabidopsis* plants (Fig. 4a). Early flowering was observed in GhAP1-OE transgenic *Arabidopsis* plants (Fig. 4b). The flowering time was measured using the days after sowing (DAS) to first flowering. Through investigation, we found that WT plants started flowering at 29 DAS and had 11 rosette leaves, whereas transgenic lines started flowering at 24 DAS and had 8 rosettes. The transgenic *Arabidopsis* bloomed about 5 days earlier than the WT and had 3 rosette leaves less than WT (Table 1).

**Table 1.** Flowering time and rosette leaf number of transgenic and WT *Arabidopsis* lines

<table>
<thead>
<tr>
<th>Plant lines</th>
<th>Start bolting(DAS)</th>
<th>Start flowering(DAS)</th>
<th>Number of rosette leaves</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>27.3±0.90</td>
<td>29.6±1.20</td>
<td>11.6±1.11</td>
</tr>
<tr>
<td><em>GhAP1</em>-4</td>
<td>20.6±0.49</td>
<td>25.6±0.80</td>
<td>8.4±0.49</td>
</tr>
<tr>
<td><em>GhAP1</em>-15</td>
<td>19.7±0.90</td>
<td>24.7±0.46</td>
<td>8.3±0.90</td>
</tr>
<tr>
<td><em>GhAP1</em>-17</td>
<td>19.4±1.36</td>
<td>24.6±0.49</td>
<td>9±0.63</td>
</tr>
</tbody>
</table>

* Data were shown as means ± SDs from ten plants.

**VIGS delayed cotton flowering**

SGN-VIGS Tool (http://vigs.solgenomics.net/) was used to analyze the *GhAP1* sequence and 250 bp of *GhAP1* was selected to construct pCLCrVA-GhAP1 vector. pCLCrVA-PDS vector was served as a positive control and pCLCrVA vector was served as an empty control. The cotton plants harboring pCLCrVA-PDS showed an albino phenotype, suggesting that the VIGS assay was successful. qRT-PCR was performed to evaluate the effect of gene silencing, the results showed that the expression level of *GhAP1* in positive plants was significantly lower than that in VA empty control (Fig. 5a). When flowering was observed in VA plants, the positive plants were not (Fig. 5b). Compared to VA plants, we found that positive plants flowered later.

**GhAP1 interacted with GhPSAT2**

A yeast library was constructed using the bud samples from two-leaf, three-leaf and five-leaf stages of early maturing variety CCR150. pGBK7-GhAP1 vector was build and transformed into Y2H Gold yeast cells. After three days both of pGBK7 control and pGBK7-GhAP1 can grow on SD-Trp medium plate, but can’t grow and become blue on SD-Trp/-His/-Ade and SD-Trp/-His/-Ade/X-α-gal medium plate (Fig. 6a), indicating that GhAP1 had no self-activation activity. According to gene annotation and yeast two-hybrid validation, a pyridoxal phosphate (PLP)-dependent transferases superfamily protein GhPSAT2 (Gh_D07G1721) was identified to be interacted with GhAP1 (Fig. 6b). The research on *GhPSAT2* was poor, but its homologous gene in soybean was proved to encode mitochondrial isozyme of aspartate aminotransferase (AAT4) when expressed in *E. coli* [36].

**GhLFY bound to the promoter of *GhAP1* in the yeast one-hybrid assay.**
Three GhLFY binding sites (CCAATGG) were founded in promoter of *GhAP1*, which located at -80, -1500 and -1700 bp. In order to validate which site can interact with GhLFY, approximately 51 bp around these sites were selected respectively and inserted into the pHis2 carrier vector with three series repeats (called pHis2-L plasmid). When 120 mM 3-amino-1,2,4-triazole (3-AT) were added, their self-activation could be inhibited. pGADT7-GhLFY vector was constructed and co-transformed into Y187 yeast cells with pHis2-L vector. The yeast cells contained pHis2-1500bp and pGADT7-GhLFY vector could growth on the SD-Trp/-His/-Leu/120mM 3-AT medium. These results indicate that GhLFY can directly bind the promoter of *GhAP1* at the -1500 bp site (Fig. 7).

**The expression of *GhAP1* was suppressed by GhLFY in the dual-luciferase assay**

A transient dual-luciferase assay was performed in tobacco leaves to elucidate the functional role of *GhAP1* in vivo. The experiment was conducted using a double reporter plasmid, pGreenII0800-LUC, containing the REN luciferase driven by the 35S promoter and the LUC luciferase driven by the *GhAP1* promoter. In addition, the assay includes an effecter plasmid, pGreenII62-SK, expressing the GhLFY. The constructs are shown in Fig. 8a. The ratio of LUC to REN was used to reflect the transcriptional activity. The results showed that compared with the control, GhLFY inhibited the *GhAP1* promoter by significantly reducing the LUC/REN ratio (Fig. 8b), suggesting that GhLFY inhibited the expression of *GhAP1*.

**Discussion**

**Functional characterization of *GhAP1* gene in cotton**

As an important commercial crop, cotton plays an important role in the development of the world’s agriculture. Flowering is one of the important traits of cotton. A number of studies have shown that MADS TFs regulate numerous physiological and biochemical processes in plants [37-39]. It involves every stage of the plant’s growth cycle and is considered to be a key gene family regulating flowering and vegetative organ development [40]. In this study, a MADS-box gene *GhAP1* was isolated and characterized from upland cotton CCRI50. Sequence analysis revealed that *GhAP1* contained a high conserved MADS domain and a moderate conserved K-box. These typical domains suggested that *GhAP1* was a MICK type gene. The MIKC family members are plant specific TFs, and play crucial roles in plant seed development and floral identification [9]. Organizational expression patterns were analyzed by qRT-PCR. The results showed that *GhAP1* was highly expressed in sepals, which was the same as described in the classic ABCDE model: the formation of sepals was determined by the formation of *AP1/AP1* homologous dimer and *SEP/SEP* dimer [41]. The *GhAP1* gene had highest expression at the three-leaf stage, and had higher expression in early maturing cultivars than in the late-maturing cultivars. Our *GhAP1* gene shared 72.27% similarity with the gene reported by Su et al. [34]. In addition, consistent with the conclusions of Su et al., *GhAP1* presented similar expression pattern to Su et al. gene [34]. Subcellular localization analysis revealed that the GhAP1 protein was located in nucleus, consistent with the website predicted.
As a class of a functional gene, AP1 plays an important role in the normal formation of flower organs. We carried out *Arabidopsis* transformation and cotton VIGS experiments. The overexpression showed that *GhAP1* could promote the flowering of *Arabidopsis*, and qRT-PCR indicated that the expression of *GhAP1* in transgenic plants was much higher than that in WT plants. *OsMADS18* was a homologous gene of *GhAP1* in rice, and overexpression of *OsMADS18* in rice can induce early flowering [42]. Overexpression of *PaAP1* in *Arabidopsis* can trigger early flowering in the 35S::*PaAP1* transgenic *Arabidopsis* lines [43]. *GmAAP1* could cause early flowering and alteration of floral organs when it was ectopically expressed in tobacco [44]. These studies are similar to the results of this study. In the VIGS assay, compared with the control group, the experimental group showed a late flowering phenotype, and the expression level of *GhAP1* in the experimental group was significantly lower than that of the VA plants. Silencing the SBP-box gene in Snapdragon by VIGS technology will down-regulate the expression of multiple MADS-box genes, such as *SQUA*, *AmFUL*, etc., resulting in late flowering or even no flowering of plants [45]. Silencing of *TaAG-A* and *TaAG-B* in a fertile wheat line resulted in green and yellow striped leaves, emaciated spikes and reduced seed setting rate [46]. These studies increased the credibility of our study.

**Molecular mechanism of GhAP1**

Proteins interact with each other to diversify their functions. Yeast hybrid system is an efficient and rapid method for analyzing protein interactions. For example, *OsMADS18* can interact with the seed-specific protein NF-YB by yeast two-hybrid system [47]. In addition, MADS-box proteins can regulate the temporal and spatial expression of target genes, and its own expression is also strictly regulated.

Previous reports showed that *AP1* can be directly and positively regulated by LFY [48]. Promoter sequence analysis revealed that three *GhLFY*-binding sites (CCAATGG) were present in the promoter region of *GhAP1*. Yeast one-hybrid assay showed that GhLFY can interact with the -1500 bp promoter site of *GhAP1*, indicating that GhLFY can regulate *GhAP1* by binding to the -1500 bp site. By dual-luciferase assay, it was found that the value of LUC/REN in the experimental group was significantly lower than that of the control group, indicating that GhLFY might negatively regulate *GhAP1* expression. This is in contrast to the reporting pathway in *Arabidopsis* [49]. This may be due to the functional redundancy of *GhLFY* and *GhAP1*, which is considered to be implemented by a set of shared genes. Moreover, studies have shown that *GhLFY* and *GhAP1* have antagonistic effects on target gene regulation [48]. Therefore, we suspected that this is an antagonist effect that leads to this result. As a complex biological phenomenon of plants, flowering was regulated by numerous genes. Therefore, it's understandable that there are different results, but this conjecture remains to be further proved.

As a TF, AP1 can regulate the expression of target genes, but most studies on signaling pathways comes to an abrupt end with AP1. However, little research has been done on its mechanism of action in cotton. In our study, the library screening experiment was carried out by using yeast two-hybrid. We found that a protein GhPSAT2 can interact with GhAP1. It is a mitochondrial isozyme of aspartate aminotransferase AAT4 in soybean [36]. Currently, there are few reports on GhPSAT2, indicating that it may be a new gene, and its mechanism of action on plant growth still needs to be further studied.
Conclusions

GhAP1 is a TF associated with flowering. It has a highly conserved MADS domain and belongs to the MADS-box gene family. GhAP1 is located in the nucleus. Overexpressed GhAP1 can promote flowering in Arabidopsis. The cotton VIGS showed a delayed flowering phenotype. Yeast two-hybrid assay showed that GhAP1 can interact with GhPSAT2. GhAP1 promoter contains GhLFY binding sites and GhLFY can interact with the promoter of GhAP1. In addition, GhAP1 is negatively regulated by GhLFY.

Methods

Plant materials and growth conditions

All plants used in our experiments were grown in a greenhouse at 25 °C, with a 16 h light/8 h dark cycle.

To examine gene expression at different stages of different maturity varieties, the leaves were collected from CCRI50, CCRI74, G11 and BM at the cotyledon and 1-, 2-, 3-, 4- and 5-leaf stages. Each sample included three repetitions. The growth period of CCRI50 is about 105 days. After Shuxun Yu and others investigation found that it has an obvious character of early maturity [50]. The growth period of CCRI74 is 100 days and it has obvious early maturity [51]. The growth period of G11 is about 123 days, which was identified by Guoxin rural technical service association [52]. BM plants have the characteristics of late emergence, tall and loose, it belongs to late maturing varieties [53]. As Gu described, these varieties are often used for testing and preservation in our laboratory [54]. All samples were quickly frozen in liquid nitrogen and stored at −80°C for subsequent experiments. In addition, two early-maturing varieties, CCRI50 and CCRI74, and two non-early-maturing varieties, G11 and BM, were used for expression pattern analysis. The cotton varieties were planted in the field of the Cotton Research Institute of the Chinese Academy of Agricultural Sciences (Anyang, Henan, China). Different tissues, including roots, stems, leaves, buds, fiber, petals and sepals were harvested from CCRI50 [54].

Gene cloning and sequence analysis

To amplify the CDS and promoter of GhAP1 (Gh_D03G0922), we designed primers using Oligo7. The primers used in this study are listed in Additional file 1: Table S1. The full-length CDS and promoter fragment of GhAP1 was cloned from cDNA and DNA of CCRI50 leaves, respectively. The fragments were inserted into PBI121 vector and transformed into Escherichia coli competent cells (E. coli DH5α) for sequencing. The genomic and coding sequences of GhAP1 were submitted to Gene Structure Display Server online software (GSDS2.0) (http://gsds.cbi.pku.edu.cn/) to predict gene structures. Multiple sequence alignment was conducted using DNAMAN software.

qRT-PCR

Total RNA was isolated using RNA prep Pure Plant Kit (Polysaccharides & Polyphenolics-rich) (Tiangen, China). One microgram of total RNA was prepared for cDNA synthesis in a 20 μl reaction system using a
Prime Script™ RT reagent kit with gDNA Eraser (RR047A) (TaKaRa, Japan). The cDNA was diluted 5 times for qRT-PCR. Transcript levels were detected using a 7500 Real-Time PCR system (Applied Biosystems) and UltraSYBR Mixture (Low ROX) (CW2601M) (CWBio, China). The 10 μl reaction volume contained the following components: 5 μl of 2×UltraSYBR Mixture, 0.2 μl of the PCR forward primer (10 μM), 0.2 μl of the PCR reverse primer (10 μM), 1 μl of cDNA and 3.6 μl of ddH₂O. The optimal PCR amplification procedure was used as follows: a pre-denaturation step at 95°C for 10 min; 40 cycles of 95°C for 10 s, 60°C for 30 s and 72°C for 32 s, a melting curve step at 95°C for 15 s, 60°C for 1 min, 95°C for 15 s and 60°C for 15 s. GhActin and AtUBQ10 were used as reference genes. The 2^ΔΔCT method was applied to calculate relative expression levels [55]. All reactions were performed with three technical replicates.

**Subcellular localization**

The CDS sequence of GhAP1 without termination code was amplified from CCRI50 cDNA using PCR. GhAP1 was cloned into PBI121-GFP to construct 35-GhAP1::GFP vector and transformed into Agrobacterium tumefaciens strain LBA4404. Approximately 200 μl LBA4404 strains were added into LB liquid medium containing kanamycin, rifampicin and streptomycin, shaking to OD600=1.8-2.0 at 28°C. The culture was centrifuged for 10 min at 4000 rpm. Adjust the OD600=1.5 using the transformation medium. The bacteria were left in the dark at room temperature for 3 h before injected into the tobacco leaves. The fluorescent signal was observed by fluorescence microscope after 2 days (dark culture for 24 h and normal culture for 24 h).

**Genetic transformation of Arabidopsis**

The CDS sequence of GhAP1 was inserted into PBI121 vector to construct 35S::GhAP1 and transformed into the Agrobacterium tumefaciens strain LBA4404 chemically competent cells. Arabidopsis was infected by dipping flower method [56]. The LBA4404 cells containing 35S::GhAP1 were cultured in the liquid medium and shaken to OD600=1.2-1.6 at 28 °C. The culture was centrifuged for 5 min at 5000 rpm to remove the supernatant. The bacteria was resuspended using a transformation medium (0.217 g 1/2MS + 5 g sucrose + 20 μl silwettl-77 per 100 ml H₂O) to adjust the OD600 = 0.8-1.0. The flower buds of WT plants were immersed in the suspension for 50 s and the transformed plants were placed in the dark condition for 24 h. After one week, the Arabidopsis plants were genetically transformed again. The seeds harvested form the transformed WT plants are T0 generation. The positive plants were screened using the 1/2MS medium containing kanamycin. Phenotypic observation and data statistics were carried out in T3 generation.

**VIGS assay**

For the VIGS assay, approximately 250 bp fragment of GhAP1 gene was amplified from the CCRI50 cDNA and integrated into the pCLCrVA vector to construct pCLCrVA-GhAP1. The pCLCrVA-GhAP1 plasmids were transformed into LBA4404. The LBA4404 strains carrying pCLCrVA-GhAP1, pCLCrVA (negative control) or pCLCrVA-PDS (positive control) were mixed with the strain harboring pCLCrVB (helper vector) (1:1 ratio, OD600 =1.5) and co-injected into two fully expanded cotyledons of TM-1 plants. In the VIGS assay, at
least 20 seedlings were used per group. For qRT-PCR detection, samples from at least 6 uniform injected plants were used. The cotton plants were cultivated in a greenhouse at 22°C with a 16 h light/8 h dark cycle. The experiment was repeated three times. The detailed VIGS procedure was performed as previously described [57, 58].

**Yeast two-hybrid**

To identify the interaction proteins of GhAP1, the buds and leaves of CCRI50 from the two-leaf, three-leaf and five-leaf stages were used to construct yeast two-hybrid library. *GhAP1* was cloned into pGBKTK7 to construct pGBKTK7-GhAP1 plasmids and transformed into Y2H yeast receptor cells. Three colonies were randomly selected for self-activation detection. The pGBKTK7 empty vector was used as a negative control. The three clones were diluted with water and placed on the defect medium of yeast growth, namely, SD/-Trp, SD/-Trp/-His/-Ade and SD/-Trp/-His/-Ade/X-a-gal.

The mating method was used to screen yeast library on SD/-Trp/-His/-Ade/-Leu plates to identify the proteins that can interact with GhAP1. The Y2H yeast transformants containing the correct pGBKTK7-GhAP1 was used to prepare the yeast competent cells, and the library plasmid pGBKTK7-cDNA was transferred into it. Screening was conducted on the defective culture plate of SD/-Trp/-His/-Ade/-Leu plates, and the normally growing spots were coated on the culture plate of SD/-Trp/-His/-Ade/-Leu with X-a-Gal for further screening. The spots that could grow normally and turn blue were considered to be positive spots. The positive clones were amplified via PCR and sequenced. Cotton genome blast analysis was performed to obtain the coding genes of potential proteins [59].

**Yeast one-hybrid assay**

A suitable 50 bp promoter fragment sequence of *GhAP1* was selected and three copies were designed to construct pHis2 recombinant vector. The recombinant plasmid was called pHis2-L. pHis2-L construct, positive control and negative control were applied to the corresponding defect culture plates (pGAD53m+pHis2 as negative control, pGAD53m+p53HIS as positive control). Different concentrations of 3-AT were added to the defect culture plate to screen the appropriate 3-AT concentration. The CDS sequence of *GhLFY* was inserted into the pGADT7 vector to construct pGADT7-GhLFY and co-transformed into Y187 yeast cells with pHis2-L constructs. The Y187 yeast cells containing pGADT7-GhLFY and pHis2-L were identified on SD/-Trp/-His/-Leu + 3-AT plates. If it can grow normally, it indicated that the *GhLFY* gene is the upstream gene of *GhAP1*.

**Dual-luciferase reporter assay**

Transient reporter expression was performed in tobacco leaves using a dual-luciferase reporter system [60]. The *GhAP1* promoter was amplified and inserted into the pGreenII0800-LUC vector as the reporter plasmids. The ORF of *GhLFY* was amplified and inserted into the pGreenII62-SK vector as the effector plasmids. The fusion construct plasmids were transformed into *Agrobacterium tumefaciens* strain GV3101 (pSoup-p19). The GV3101 (pSoup-p19) cells containing the recombinant plasmids were
incubated in LB liquid medium containing 50 mg L\(^{-1}\) kanamycin, gentamycin and rifampin until the OD600 value reached 0.5–0.6. Subsequently, the culture was adjusted to an OD600 value of 0.2 with the filtration buffer (10 mM MgCl\(_2\), 10 mM MES and 100 μM acetosyringone). The culture suspensions were left for 2 h at room temperature. The effector and reporter suspensions were mixed in a 1:1 ratio and co-infiltrated into tobacco leaves. After 2 days of infiltration, LUC and REN luciferase activity was detected using a dual-luciferase® reporter assay system (Promega, USA) on a Glomax 20/20 Luminometer (Promega, USA) according to the manufacturer's instructions. At least six independent replicates were performed.

**Abbreviations**

TF: Transcription factor; VIGS: Virus-induced gene silencing; CDS: Coding sequence; aa: amino acid; NLS: Nuclear localization signal; DAS: Days after sowing; WT: Wild type; qRT-PCR: Quantitative real-time PCR; Trp: Tryptophan; Leu: Leucine; Ade: Adenine; His: Histidine; 3-AT: 3-amino-1,2,4-triazole

**Declarations**

**Ethics approval and consent to participate**

The experimental research on plants (either cultivated or wild), including collection of plant material, complied with institutional, national, or international guidelines. Field studies were conducted in accordance with local legislation. The plant materials used in this study were previously preserved in our laboratory.

**Consent for publication**

Not applicable.

**Availability of data and materials**

The data sets supporting the results of this article are included within the article and its additional file.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This research was support by the Shangdong TAISHAN Industry Leading Talent Program to Shuxun Yu (LJNY201608). The funders had no role in the design of the study, the collection, analysis, and interpretation of data, and in writing the manuscript.

**Authors’ contributions**
S.X.Y. designed the research program. J.P.H., H.L.W., H.T.W. and L.J.G. revised the manuscript. X.Q.C. wrote the paper. X.Q.C., P.B.H. and A.M.W. performed the data analysis. X.Q.C., H.R.S. and S.S.C. performed the yeast two-hybrid test. X.Q.C. performed left experiments. All authors have read and approved the final manuscript.

Acknowledgments

We thank all authors for their contributions to the article. We also appreciate the reviewers and editors for their patience regarding this work.

References


38. Nobuhiro Kotoda a, Masato Wadaa, Shinnosuke Kusaba b, Yuriko Kano-, Murakamib TM, Junichi Soejima a: Overexpression of MdMADS5, an APETALA1-like gene of apple, causes early flowering in


50. Song SYSFZHMH: **Characteristics and key cultivation techniques of CCRI50 varieties** *China’s cotton* 2005, **10**:25-26.


**Figures**

![Gene structure and protein sequence analysis of GhAP1. a Gene structure of GhAP1 (Gh_D03G0922). b Sequence alignment of the deduced GhAP1 protein with its homologous proteins AtAGL8 (At5g60910), GmFULc (Glyma05g07380) and TaVRT-1 (TraesCS5B02g396600). The high conserved MADS domain](image-url)

**Figure 1**

Gene structure and protein sequence analysis of GhAP1. a Gene structure of GhAP1 (Gh_D03G0922). b Sequence alignment of the deduced GhAP1 protein with its homologous proteins AtAGL8 (At5g60910), GmFULc (Glyma05g07380) and TaVRT-1 (TraesCS5B02g396600). The high conserved MADS domain.
approximately 44 amino acids in length, semi-conservative K domain and the predicted NLS sequence were showed by the two-headed arrow. The abbreviations before the gene names of different species were as follows: At, Arabidopsis thaliana; Gh, Gossypium hirsutum; Ta, Triticum aestivum; Gm, Glycine max.

Figure 2

Expression profiles of GhAP1 in different tissues and at different stages in cotton. a Expression levels of GhAP1 in different tissues. b Expression level of GhAP1 in leaves at different developmental stages. GhActin served as the reference gene. The bars represent the means ± standard error (SEs) from three independent experiments.

Figure 3
Subcellular localization of GhAP1 protein. a Tobacco cells in bright light. b Tobacco cells with 35S-GhAP1::GFP expressed instantaneously under blue excitation fluorescence.

Figure 4

Overexpressed GhAP1 in Arabidopsis and qRT-PCR results. a qRT-PCR analysis of the transcript levels of GhAP1 in Arabidopsis. AtUBQ10 was used as a reference gene. The values represented the means ± SDs. The significance of the data was determined using Student’s t-test (** P<0.01). The experiment was conducted with three repetitions. b Early flowering phenotypes of WT plants and three transgenic lines. AP1-4, AP1-15 and AP1-17 are transgenic plants. WT is Col-0.

Figure 5

VIGS phenotype and qRT-PCR analysis. a The expression levels of GhAP1 in VIGS plant and VA empty control. GhActin was served as the reference gene. The values represented the means ± SDs from three independent experiments. Independent t-tests revealed significant (**P<0.01) differences between the control and silenced plants. VA-11 and VA-8 were empty control; AP1-17 and AP1-23 were VIGS plants. b
VIGS phenotype indicated positive plants flowering later than VA empty control. VA-11 was used as empty control; AP1-17 and AP1-23 were VIGS plants; PDS was pCLCrVA-PDS positive control.

Figure 6

GhAP1 self-activation detection and GhAP1 physically interacted with GhPSAT2 in the yeast two-hybrid system. a GhAP1 self-activation detection. pGBK7T was served as negative controls. b GhAP1 physically interacted with GhPSAT2. pGBK7T-GhAP1 and pGADT7-GhPSAT2 were co-transformed into Y2H Gold yeast cells and assayed on the SD-Trp/-Leu and SD-Trp/-His/-Ade/-Leu/X-α-gal medium. pGBK7T-p53 + pGADT7-largeT and pGBK7T-laminC + pGADT7-largeT were used as positive and negative controls, respectively.
**Figure 7**

GhLFY bound to the promoter of GhAP1 in the yeast one-hybrid assay. pHis2-ProGhAP1 and pGADT7-GhLFY were co-transformed to Y187 yeast cells. pGAD53m+p53His and pGAD53m+pHis2 were used as positive and negative controls, respectively.
GhLFY inhibited the expression of GhAP1 in the dual-luciferase reporter assay. a Sketch of the effector and reporter constructs. The GhLFY was cloned into the effector vector pGreenII62-SK. The promoter of GhAP1 was cloned into the reporter vector pGreenII0800-LUC. b GhLFY inhibited the expression of GhAP1. The data were indicated by the ratio of LUC to REN. The values represent the means ± SDs of six independent repeats. The significance of the data was determined using Student’s t-test (**) P<0.01).

**Figure 8**

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1TableS1.docx