Genome-wide analysis of the CalS gene family in cotton reveals their potential roles in fiber development and responses to stress

Jiajia Feng¹,², Yi Chen¹, Xianghui Xiao², Yunfang Qu¹, Pengtao Li², Quanwei Lu¹,², Jinling Huang¹

¹ College of Agriculture, Shanxi Agricultural University, Taigu, Shanxi, China
² School of Biotechnology and Food Engineering, Anyang Institute of Technology, Anyang, Henan, China

ABSTRACT

Callose deposition occurs during plant growth and development, as well as when plants are under biotic and abiotic stress. Callose synthase is a key enzyme for the synthesis of callose. In this study, 27, 28, 16, and 15 callose synthase family members were identified in Gossypium hirsutum, Gossypium barbadense, Gossypium raimondii, and Gossypium arboreum using the sequence of Arabidopsis callose synthase. The CalSs were divided into five groups by phylogenetic, gene structure, and conservative motif analysis. The conserved motifs and gene structures of CalSs in each group were highly similar. Based on the analysis of cis-acting elements, it is inferred that GhCalSs were regulated by abiotic stress. WGD/Segmental duplication promoted the amplification of the CalS gene in cotton, and purification selection had an important function in the CalS family. The transcriptome data and qRT-PCR under cold, heat, salt, and PEG treatments showed that GhCalSs were involved in abiotic stress. The expression patterns of GhCalSs were different in various tissues. We predicted that GhCalS4, which was highly expressed in fibers, had an important effect on fiber elongation. Hence, these results help us understand the role of GhCalSs in fiber development and stress response.

INTRODUCTION

Callose is a linear homopolymer composed of β-1, 3-linked glucose, widely found in higher plants as an important part of specialized cell walls or cell wall-associated structures (Chen & Kim, 2009). However, Callose synthesis mainly depends on callose synthase (CalS) (Hong, Delauney & Verma, 2001). In Arabidopsis thaliana, callose activity was significantly increased and deposited after overexpression of the AtCalS12 (Ellinger et al., 2013). In rice, callose deposition in the plasmodesmata of crr1 (AtCalS10 homologous gene) mutants was reduced (Song et al., 2016).

Callose can regulate the transport of plasmodesmata and phloem, affect plant development and response to biotic stress by controlling callose synthases. For example, the AtCalS5 could maintain the normal formation of callose walls during pollen development.
Meanwhile, CalS was regulated by various signaling pathways, and different biological regulatory processes involve hormones, transcription factors. For instance, external application of ABA can increase the activity of callose synthase of rice, promoting callose deposition, and thus increasing its resistance to BPH insects \(\textit{Liu et al., 2017}\). The expression of CalSs were also affected by developmental and stress conditions, such as pollen development \(\textit{Toller et al., 2008; Xie, Wang & Hong, 2010; Huang et al., 2009; Shi et al., 2015a; Shi et al., 2015b}\), low-temperature stimulation \(\textit{Fromm et al., 2013}\), mechanical wounding \(\textit{Cui & Lee, 2016; Jacobs et al., 2003; Xie et al., 2011}\), fungal diseases \(\textit{Dong et al., 2008; Oide et al., 2013; Ellinger et al., 2014; Blümke, Somerville & Voigt, 2013}\), bacterial diseases \(\textit{Granato et al., 2019; Enrique et al., 2011}\), and insect diseases \(\textit{Ahmad et al., 2011; Koh et al., 2012; Sun et al., 2018; Yao et al., 2020}\).

Due to the importance of callose, the callose synthase gene family has been identified in various plants. To date, it has been reported as 12 \textit{AtCalSs} in \textit{Arabidopsis thaliana} \(\textit{Hong, Delauney & Verma, 2001}\), 12 \textit{CsCalSs} in \textit{Citrus sinensis} \(\textit{Granato et al., 2019}\), 32 \textit{BnCalSs} in \textit{Brassica napus} \(\textit{Liu, Zou & Fernando, 2018}\), eight \textit{VvCalSs} in \textit{Vitis vinifera} \(\textit{Yu et al., 2016}\), seven \textit{HvCalSs} in \textit{Hordeum vulgare} \(\textit{Schöber et al., 2009}\), 15 \textit{BraCalSs} in Chinese cabbage \(\textit{Pu et al., 2019}\). Generally, according to the evolutionary analysis of CalSs in the above species, the CalS family can be divided into three of four main groups \(\textit{Chen & Kim, 2009; Liu, Zou & Fernando, 2018}\).

Cotton is an important economic crop in China which yield is affected by biotic and abiotic stresses, producing prevalent fibers for textile industry \(\textit{Wang et al., 2020; Fang et al., 2014}\). Some researches have reported that cotton fiber elongation was related to callose deposition which may be involved in the closing of plasmodesmata, then promoting the fiber length \(\textit{Ruan et al., 2004}\). It is possible that callose affects fiber elongation by controlling CalS. In light of the above, the CalS may play a significant role in cotton in responding various stresses and promoting fiber elongation. A \textit{callose synthase} gene, \textit{CFL1} was identified \(\textit{Cui et al., 2001}\), however, the \textit{callose synthase} gene family members, phylogenetic relationships and expression patterns in cotton are still unclear.

In the current study, we identified \textit{callose synthase} genes in two cultivated allotetraploids cotton, \textit{G. hirsutum} and \textit{G. barbadense}, and their two putative genome donors, \textit{G. raimondii} and \textit{G. arboreum}, then discussed their phylogenetic relationships, conserved domains, gene structures, synteny, and cis-acting elements. We also focused on the expression patterns of \textit{GhCalSs} in various tissues and their expression under abiotic stress. These findings provide a solid foundation for further study of the roles of \textit{CalSs} in cotton fiber development and stress responses.

**MATERIALS & METHODS**

**Plant materials and treatments**

Upland cotton TM-1, planted in Anyang Institute of Technology, was subjected to salt stress (350 mM NaCl) and drought stress (12% PEG6000) when the seedling reached two weeks. The leaves were collected 0 h, 1 h, 3 h, 6 h, 12 h, and 24 h after treatment. CCRI45 and MBI7747 were planted on farms managed by Cotton Research of Chinese Academy.
of Agricultural Sciences in Anyang. Cotton fibers were collected at 5, 10, 15, 20, 25 days post-anthesis (DPA). All samples were stored at $-80^\circ$C.

**Identification of CalS family members in Gossypium spp**

The genome sequences and annotated files of *G. hirsutum* (Hu et al., 2019), *G. barbadense* (Hu et al., 2019), *G. raimondii* (Paterson et al., 2012), and *G. arboretum* (Du et al., 2018) were downloaded from Cottongen (https://www.cottongen.org) (Yu et al., 2014). Both blast and HMMER were used to identify the CalS sequences. The 1,3 beta-glucan synthase (PF02364) and FKS1_DOM1 domain (PF14288) from the Pfam database (http://pfam.xfam.org) were searched by the HMMSearch program in TBtools to determine the presumed protein sequence (Chen et al., 2020). Besides, 12 Arabidopsis CalSs (Hong, Delauney & Verma, 2001) were used as queries sequences to identify family members using the Blastp program of TBtools (Chen et al., 2020). The protein sequences without above two domains were rejected and the domain which incomplete were also deleted. Finally, the final sequences were calculated by using ExPASy (https://www.expasy.org) to calculate the theoretical isoelectric points (pI) and molecular weights (MW) and using the CELLO (http://cello.life.nctu.edu.tw) for subcellular localization prediction (Yu et al., 2006).

**Phylogenetic tree construction, gene structure, and motif analysis**

The phylogenetic tree among four *Gossypium* species and *Arabidopsis thaliana* was constructed by MEGA7 (Kumar, Stecher & Tamura, 2016). It was constructed by the neighbor-joining (NJ) method, with 1,000 bootstrap replicates, then was drawn by using EvolView (He et al., 2016). TBtools was used to extract the location information of CalSs and visualize the gene structure. MEME (https://meme-suite.org/meme/tools/meme) was used to identify the conservative motif with the parameter set to the maximum number of motifs: 20.

**Chromosome location and synteny analysis for CalSs**

The locations of CalSs on chromosomes were shown by TBtools using four cotton species genomic annotation files (Chen et al., 2020). MCScanX was used to analyze the collinearity of the CalSs, that is, using CalS protein sequences to analyze the orthologous and paralogous gene pairs (Chen et al., 2020). Collinear gene pairs were visualized by using the circos (Chen et al., 2020). To investigate the selection pressure between homologous genes, we calculated the nonsynonymous substitutions rate (Ka) and synonymous substitutions rate (Ks) of homologous genes by KaKs_Calculator (Wang et al., 2010).

**Analysis of Cis-acting element in promoters and functional enrichment analysis**

The 2,000 bp sequence upstream of the translation initiation codon ATG of CalS gene was selected as promoter. The cis-acting elements contained in the promoter region of the CalSs were predicted using the PlantCare website (Lescot et al., 2002). For functional enrichment analysis, gene ontology (GO) analysis was performed using the OmicShare tool (https://www.omicshare.com/tools).
**GhCalSs expression patterns under different tissues and abiotic stresses**

In order to analyze the expression of GhCalSs in different tissues and under stress, we downloaded 11 tissues (bract, pental, torus, root, leaf, stem, pistil, sepal, anther, ovule, fiber) and abiotic stress treatment (cold, heat, drought, salt) data from Cotton Omics Database (http://cotton.zju.edu.cn) (accession number: PRJNA490626) (Hu et al., 2019). GhCalSs with FPKM > 1 were considered as expressed genes. The expression patterns of the GhCalSs were visualized by ComplexHeatmap (Gu, Eils & Schlesner, 2016) based on the value of log2(FPKM+1).

**RNA isolation and qRT-PCR analysis**

FastPure Plant Total RNA Isolation Kit (RC401, Vazyme) was used to extract RNA, and then we used 1µg to synthesize cDNA (HiScript III 1st Strand cDNA Synthesis Kit, R312 Vazyme). ChamQ Universal SYBR qRT-PCR Master Mix (Q711, Vazyme) was used for qRT-PCR in ABI 7500 Fast Real-time PCR System (Applied Biosystems, USA). Gene-specific primers for qRT-PCR were designed by using primer-blast in NCBI, with melting temperatures of 55–60 °C, product lengths of 101–221 bp, primer length of 18–25 bp (Table S1). For qRT-PCR, the reaction contains 10 µL 2x ChamQ Universal SYBR qPCR Master Mix, 0.4 µL of each primer, 3 µL template, and ddH2O to make up the total 20 µL volume. Then it was carried out in the following condition: one cycles of 95 °C for 30 s, 40 cycles of 95 °C for 10 s and 60 °C for 30 s. Each experiment was repeated three times, and two of the completed data were selected for drawing. Expression of all genes were calculated using a 2−ΔΔCt method (Livak & Schmittgen, 2001).

**RESULTS**

**Identification and characterization of CalSs in Gossypium spp**

Through the analysis of the CalS protein sequences in Arabidopsis thaliana, we found that all the protein sequences contain 1, 3-β-glucan synthase (PF02364) and FKS1_DOM1 domain (PF14288), total of 27 members of the CalS gene family in G. hirsutum, 28 in G. barbadense, 15 in G. raimondii, and 16 in G. arboretum were identified, all of which were named according to their chromosomal locations. The properties of CalSs in cotton were further analyzed (Table S2). The protein sequence length of CalSs ranged from 1,494 to 1,979 amino acids, with an average MW of 212.88 kDa, and shared high similarity to the Arabidopsis thaliana CalS proteins (Hong, Delauney & Verma, 2001). The isoelectric point (PI) values of the above genes were all greater than 7, indicating that CalSs in cotton were alkaline, which was the same as the biochemical properties of the CalSs in Chinese cabbage and Brassica (Pu et al., 2019; Liu, Zou & Fernando, 2018). The CalSs were most likely localized in the plasma membrane, as predicted in Arabidopsis thaliana and Chinese cabbage (Pu et al., 2019; Zavaliev et al., 2011).

**Classification and phylogenetic analysis of the cotton CalSs**

In order to investigate the evolutionary relationships of the CalSs in the four cotton species and its relationship with Arabidopsis thaliana, a phylogenetic tree was constructed using
the protein sequences of CalSs (Fig. 1). Based on the phylogenetic tree of this study, the CalSs were divided into five groups. The distribution of CalSs in each group was shown in Table S3. The members of Group A were homologous to AtCalS11/AtCalS12, the members of Group B were homologous to AtCalS9/AtCalS10, the members of Group C were homologous to AtCalS6/AtCalS7, the members of Group D were homologous to AtCalS8, and the members of Group E were homologous to AtCalS1-5. There were Arabidopsis genes homologous to cotton in each group, further indicating that the cotton CalSs and the Arabidopsis thaliana CalSs were close in evolutionary, which was consistent with the evolutionary relationship between Arabidopsis and cotton. It is observed that most of the CalSs derived from At-subgenome of two cultivated allotetraploids cotton stayed close together with the CalS gene of G. arboretum, and the CalSs of Dt-subgenome stayed close together with the CalS gene of G. raimondii, which was consistent with the hypothesis that the allotetraploid cotton species were produced by the recombination of two diploid cotton species (Liu et al., 2015). Phylogenetic tree analysis suggested that the CalS homologous gene in cotton may have similar functions.

**Gene structure and amino acid motif analysis of the CalSs**

The diversity of gene structure and differences in conserved motifs are the manifestations of the evolution of multigene families (Magwanga et al., 2018). The distribution of exon/intron regions of CalSs was analyzed to understand the diversity of gene structure (Fig. 2). The number of CalSs exons varied from 1–51, and most CalSs had more than 35 exons (57/86, 66.2%). Clearly, these CalSs were divided into an exon-poor group (<7 exons, group A) and other exon-rich groups (>37 exons, group B–E) (Fig. 2, Table S2). The exons of CalSs had high similarity in the same group, and the number of exons in group B, group D, and group E were the same (Fig. 2, Table S2).

The motif is a conserved region in the sequence (Magwanga et al., 2018). We identified 20 possible motifs using MEME (Fig. 2). Interestingly, all CalSs except GbCalS2/14/15 and GhCalS1 contained motif1-20 and were arranged in the sequence of motif15-9-8-13-12-11-7-20-14-16-3-2-1-6-19-5-17-4-18-10. The distribution of CalSs were slightly different among different groups, and only the number and arrangement position were different. The number and arrangement of motifs in the same group of CalSs were more similar than those in other groups.

**Chromosomal location, gene duplication, and syntenic analysis of the CalSs in Gossypium spp**

Based on the sequencing and annotated information of the four cotton genomes, the chromosome length and the distribution of genes on chromosome could be analyzed (Fig. 3). The distribution of CalSs in the two heterotetraploid cotton species chromosomes was highly similar. For example, CalSs had the same number and distribution on chromosomes A03, A04, A05, A08, A11. In G. hirsutum, 27 GhCalSs were distributed on 15 chromosomes, including 13 GhCalSs in At-subgenome and 14 GhCalSs in Dt-subgenome. In G. barbadense, 28 GbCalSs were distributed on 16 chromosomes, including 14 GbCalSs in At-subgenome and 14 GbCalSs in Dt-subgenome. In G. arboreum, 16 GaCalSs were
Figure 1  Phylogenetic analysis of CalS protein from *G. hirsutum*, *G. barbadense*, *G. raimondii*, *G. arboretum* and *Arabidopsis*. The bootstrap values are shown at the nodes. The CalSs from *G. hirsutum*, *G. barbadense*, *G. raimondii*, *G. arboretum*, and *Arabidopsis* are marked with red check, orange rect, purple star, green triangle, grey circle, respectively.

Gene duplication is the basis for the functional differentiation of homologous genes, the main reason for the generation of new functional genes (Conant & Wolfe, 2008). In order to explain the gene replication events of CalSs in cotton, we identified 15, 14 paralogous gene pairs in *G. hirsutum*, *G. barbadense* respectively, and one pair in *G. arboreum*. But there was no paralogous gene pair in *G. raimondii* (Table S4). GhCalS21/22, GbCalS1/2 as well as GbCalS22/23 were tandem duplication. In the four cotton species, the duplication
events of the CalSs were WGD/Segmental, Tandem Duplicates, Dispersed, and proximal duplication, and the main expansion mechanism was WGD/Segmental (Table S2).

In order to illustrate the collinearity of CalS genes, we analyzed the orthologous and paralogous gene pairs (Fig. 4, Table S4). There were 31 CalS orthologous gene pairs among G. arboretum and two allotetraploid cotton species, including 15 pairs between with At-subgenome of G. hirsutum and 16 pairs between with At-subgenome of G. barbadense. There were 17 CalS orthologous gene pairs among G. raimondii and two allotetraploid cotton species, including nine pairs between with Dt-subgenome of G. hirsutum and eight pairs between with Dt-subgenome of G. barbadense. Meanwhile, Ka/Ks of CalS homologous pairs were calculated to further understand the adaptation of the CDS region of CalSs (Fig. 4, Table S5). Most of the homologous gene pairs Ka/Ks < 1, and about 94.6% gene pairs had a Ka/Ks ratio less than 0.5, which meant that almost all gene pairs underwent purification selection. Only Ka/Ks > 1 of GaCalS2/GbCalS1 indicated that this was a positive selection for beneficial mutations.
Figure 3  Distribution of 86 CalSs on cotton chromosomes. The chromosome name is on the left of each chromosome, and the gene ID is on the right. (A) G. hirsutum; (B) G. barbadense; (C) G. arboretum; (D) G. raimondii.

Analysis of Cis-acting elements in promoter

Transcription factors can be combined with cis-elements in the promoter region to regulate gene transcription. Investigation of upstream regulatory sequence can help us to well understand the regulation mechanism and also supportive to estimate the potential function of the gene (Fig. 5, Table S6). Given the effect of plant hormones in abiotic stress, we focused on plant hormone responsive elements in promoter regions. ABA- (ABRE), auxin- (AuxRR-core, TGA-element), Gibberellin- (GARE-motif, P-box, TATC-box), MeJA- (CGTCA-motif, TGACG-motif), SA- (TCA-element) responsive elements were...
Collinearity analysis of CalSs in tetraploid and diploid cotton. (A) Orthologous and paralogous gene pairs among tetraploid and diploid cotton species. The lines represented by various colors indicates the syntenic regions around CalSs, and the color between the same species is the same (B) Ka, Ks, Ka/Ks distribution of CalS gene pairs. Ka, Ks, Ka/Ks analysis of GbCalS-GaCalS, GbCalS-GbCalS, GbCalS-GhCalS, GbCalS-GrCalS, GhCalS-GaCalS, GhCalS-GhCalS, GhCalS-GrCalS.

Full-size DOI: 10.7717/peerj.12557/fig-4

found in the promoters of 18, 6, 14, 18, 12 GhCalSs. All GhCalSs contained hormone response elements except the GhCalSs in Group D and GhCalS5 in Group C. More than half of the GhCalSs contained ABA/GA/MeJA-responsive elements. Auxin-responsive elements only in GhCalS6/14/20/21/22/23. Meanwhile, we also paid attention to elements related to stress. Low-temperature- (LTR), wound- (WUN-motif), drought- (MBS), stress- (TC-rich repeats), anaerobic induced response element (ARE), anoxic specific inducibility element (GC-motif) were found in the promoters of 11, 2, 15, 9, 23, 4 GhCalSs. Wound-responsive elements only in GhCalS5 and GhCalS18. Anoxic specific inducibility elements only in GhCalS11/15/17/25. In addition, these results suggested that CalSs might regulated by hormone and abiotic stresses.

Expression patterns of the GhCalSs under abiotic stresses

Previous studies have reported that the CalSs respond to abiotic stresses (Cui & Lee, 2016; Jacobs et al., 2003; Fromm et al., 2013). To understand the response of GhCalSs, we used public RNA-seq data of TM-1 treated with cold, hot, NaCl, and PEG to observe the expression patterns of GhCalSs (Fig. 6). Interestingly, all expressed GhCalSs were induced by different abiotic stresses, and the expression patterns were different. The expression of GhCalS3 was significantly up-regulated under cold, hot, NaCl, and PEG. The expression patterns of GhCalSs in the same group were slightly consistent, such as GhCalS3 and GhCalS6, GhCalS2 and GhCalS16. In order to verify the results obtained by the above transcriptome, cotton seedlings were treated with PEG and NaCl, and then the GhCalS2/3/6/9/16 were selected for qRT-PCR (Fig. 7). The expression of GhCalS3 and GhCalS6 in Group A were up-regulated within 24 h under PEG treatment and reached the peak at 24 h. The expression of GhCalS2, GhCalS9, GhCalS16 were up-regulated at first and then down-regulated and last up-regulated after PEG treatment. After NaCl treatment,
there was no consistent trend of gene expression. *GhCalS3* was significantly induced by NaCl and significantly up-regulated at 3 h. Both *GhCalS2* and *GhCalS16* of Group B were down-regulated within 24 h. The expression of *GhCalS6* and *GhCalS9* reached a peak at 12 h. These findings indicated that the expression patterns of several *GhCalSs* were changed after treatment, which proved that *GhCalSs* increased adaptability to abiotic stress (Fig. 6).

**Enrichment analysis of the GhCalSs**

In order to further understand the function of *GhCalSs*, we carried out functional enrichment annotation of gene ontology (GO) using p-value of \( \leq 0.05 \) as the cutoff. The
Figure 6  RNA sequence profiling of the CalS gene family. (A) Heatmap displaying expression of expressed GhCalSs under hot, cold, NaCl, and PEG treatment (B) Heatmap displaying expression of expressed GhCalSs in each tissue.

Figure 7  qRT-PCR results of GhCalSs under PEG and NaCl.
results improved our accurate understanding of gene function, including many significantly enriched terms (Fig. 8, Table S7). The GO-BP enrichment results showed 34 terms such as (1->3)-beta-D-glucan biosynthetic process (GO:0006075), beta-glucan metabolic process (GO:0051273), cellular carbohydrate biosynthetic process (GO:0034637), cellular macromolecule biosynthetic process (GO:0034645). The GO-CC enrichment results discovered 16 terms such as 1,3-beta-D-glucan synthase complex (GO:0000148), plasma membrane protein complex (GO:0098797), transferase complex (GO:1990234), catalytic complex (GO:1902494). The CC terms enriched by GO were consistent with the subcellular localization of \textit{GhCalSs}. GO-MF enrichment exposed 8 terms, including 1,3-beta-D-glucan synthase activity (GO:0003843), UDP-glucosyltransferase activity (GO:0035251), catalytic activity (GO:0003824), hexosyltransferase activity (GO:0016758). In short, the GO enrichment results confirmed the function of the \textit{GhCalSs} in many biological processes, which were associated with 1,3-beta-D-glucan synthetic activity, hydrolyzase activity, and membrane parts.

\textbf{GhCalSs expression patterns in various tissues and their role in fiber development}

We used transcriptome data from different tissues of \textit{GhCalSs} to gain insight the tissue-specific expression patterns of cotton. For instance, \textit{GhCalSs} were expressed in various tissues, and some of them were highly expressed. Some of \textit{GhCalSs} were expressed in one
Figure 9  Expression patterns of GhCalSs in cotton fiber. (A) The expression of GhCalSs of MBI7747 and CCRI45 at different fiber developmental stages. (B) qRT-PCR results of GhCalS4 at different fiber developmental stages.

or more tissues (GhCalS2, 3, 4, 6, 8, 9, 15, 16, 19, 20, 21). However, the expression of a few genes (GhCalS1, 5, 7, 10, 11, 12, 13, 14, 17, 18, 22, 23, 24, 25, 26, 27) did not show any expression in any tissues.

In order to determine the effect of GhCalSs in cotton fiber development, we focused on the expression of GhCalSs in different fiber developmental stages of two samples, MBI7747 and CCRI45, with different lengths and strengths (Lu et al., 2017) (Fig. 9). The expression of GhCalS4 was the highest in TM-1, MBI7747, CCRI45 fiber tissue, so it was speculated that GhCalS4 had an important function in cotton fiber development. In order to further determine its function in fiber development, qRT-PCR was used to analyze the GhCalS4 expression differences in two samples (Fig. 9). The results showed that the expression level of GhCalS4 in the two samples gradually increased from 5 DPA to 25 DPA, which was consistent with the transcriptome data of TM-1 used above. The expression of GhCalS4 in CCRI45 was higher than in MBI7747 at 5DPA, 10DPA, 15DPA but was significantly lower than that of MBI7747 at 25DPA (Lu et al., 2017). Thus, GhCalS4 may be involved in cotton fiber elongation.
DISCUSSION

Callose plays a vital role in plant growth, development, and resistance to various adverse factors (Píršelová & Matušíková, 2012). The gene family of *callose synthase* has been identified in a variety of plants. In this study, we identified the *CalSs* in *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum*, aiming to understand the role of *CalS* family in the cotton development.

A total of 86 *CalSs* were identified in four cotton species. They were divided into five groups based on evolutionary relationships. We divided *CalSs* (*AtCalS9-12* homologous) into two groups due to the large difference of the CDS number, and the other groups were the same as those in *Arabidopsis thaliana*. The *CalSs* number was 2:2:1:1 in group B/C/D, which was consistent with the evolutionary relationship among cotton species (Table S3). Compared with *Arabidopsis thaliana*, different percentages existed between subgroups. The percentages of Group A and Group E were significantly different from those of the corresponding *CalSs* in *Arabidopsis thaliana*, suggesting that these genes in cotton may have functional differences with homologous genes in *Arabidopsis thaliana* to a certain extent. These results will help to validate the function of cotton *CalS* homologous gene with *Arabidopsis thaliana*.

Tetraploid cotton species are formed by natural crossbreeding between *G. raimondii* and *G. arboretum* (Wendel et al., 2009). Thus, the four cotton species are closely related in evolution. In Fig. 4, the orthologous gene pairs of *CalSs* were all clustered in the same branch or group. Phylogenetic and orthologous genes of *CalS* further indicated that the results of this study were consistent with the evolutionary view.

A large number of hormone-responsive elements were identified on *GhCalSs* promoters which may be involved in the regulation of *GhCalSs*. Salicylic acid (SA) was an endogenous signal molecule in plants (Loake & Grant, 2007). In *Arabidopsis thaliana*, the expressions of *AtCalS1/5/9/10/12* were up-regulated by exogenous SA. Abscisic acid (ABA) played an important part in coping with a variety of adverse factors, closely related to callose synthesis (Liu et al., 2017). During the dormancy of *Populus tomentosa* buds, short-day induced ABA biosynthesis, promoted the expression of *PtCalS1*, callose deposited at the plasmodesmata to form blockage, which prevented the growth signal molecules from entering the cell and kept the dormancy state of buds (Tylewicz et al., 2018). Jasmonic acid (JA) was also involved in callose regulation, and Methyl Jasmonate (MeJA) application promoted callose deposition in grape leaves. Inhibition of the expression of *Cationic peroxidase 3* (*OCP3*), a negative regulator of the JA pathway, increased callose deposition (Repka, Fischerová & Šilhárová, 2004). In conclusion, ABA, JA, and SA were involved in the regulation of callose deposition. *GhCalSs* promoters with ABA, SA and JA response elements were highly likely to be regulated by them in cotton. However, how *CalS* gene is regulated by these hormones in the face of biotic-abiotic stress or growth and development is not known, which needs to be further studied.

Callose deposition is one of a series of coping strategies in plants to abiotic stress. Low temperature stimulation of maize leaves increased callose content and reduced transport of photosynthate in phloem (Wu et al., 2018). In *Arabidopsis thaliana*, *AtCalS7, AtCalS8*
and AtCalS12 were associated with callose synthesis under the condition of wound (Jacobs et al., 2003; Cui & Lee, 2016; Xie et al., 2011). In this study, public transcriptome data were used to analyze the responses of cotton leaves to cold, heat, salt and drought, and qRT-PCR was used to verify the results, which showed that CalSs were involved in abiotic stresses.

Callose deposits regulate material transport and control plant development. In this study, GhCalS4 was highly expressed in fibers and differentially expressed in MBI7747 and CCRI45 fibers at each fiber developmental stage (5/10/15/20/25 DPA). It has been reported that callose deposition may be involved in the closure of plasmodesmata, and the closure of plasmodesmata had an important function in the elongation of cotton fibers (Ruan et al., 2004). In Sea Island cotton, plasmodesmata remain open longer than in Upland cotton, allowing sucrose to be fed into fibroblasts, which eventually increase osmotic potential by hydrolysis to fructose and glucose. The more soluble sugar, K+ accumulated, the higher the cell leavening pressure, which promoted the elongation of cotton fiber (Hu et al., 2019). MBI7747 is a chromosome segment substitution line (CSSL) with different genetic background constructed by crosses between the upland cotton CCRI45 as the recurrent parent and the Sea Island cotton Hai 1 with outstanding fiber quality through the combination of high-generation backcrossing and molecular marker-assisted selection. The fiber length and strength of MBI7747 are better than CCRI45. In CCRI45, the expression level of this gene at 5DPA, 10DPA and 15DPA were all higher than those of MBI7747 during fiber elongation, and it was speculated that the degree of callose deposition in MBI7747 was lower than that of CCRI45, which made more sucrose input into fiber cells to increase osmotic potential and promote fiber elongation. Thus, GhCalS4 may be an introgression gene or there was difference in epigenetic regulation.

**CONCLUSIONS**

In this study, we identified 86 CalSs from *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum* using conserved domains. Phylogeny, gene structure, motif, chromosome location and homologous genes were analyzed. It indicated that CalSs have been highly conserved during evolution by the analysis of CalSs structure, conserved motifs, and syntenic blocks. WGD/Segmental replication was the main driving force for the amplification of CalS family in cotton, and purification selection played an important role in the evolution of CalSs. In addition, the cis-acting elements of GhCalSs related to hormone regulation and development and their expression patterns in stresses and tissues were also analyzed. CalS gene can be induced by abiotic stress. Furthermore, the expression difference of GhCalS4 in fiber of different length and strength materials was analyzed. It was speculated that GhCalS4 played a major role in fiber elongation. These findings could lay a foundation for further study on the role of CalS gene in stress response and fiber development.
ADDITIONAL INFORMATION AND DECLARATIONS

Funding
This work was supported by the National Natural Science Foundation of China (U1804103), the State Key Laboratory of Cotton Biology Open Fund (CB2020A10). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
The National Natural Science Foundation of China: U1804103.
The State Key Laboratory of Cotton Biology Open Fund: CB2020A10.

Competing Interests
The authors declare there are no competing interests.

Author Contributions
- Jiajia Feng performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Yi Chen and Xianghui Xiao performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yunfang Qu and Pengtao Li analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Quanwei Lu and Jinling Huang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability
The following information was supplied regarding data availability:
The raw data are available in the Supplemental File.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.12557#supplemental-information.

REFERENCES
Ahmad S, Veyrat N, Gordon-Weeks R, Zhang Y, Martin J, Smart L, Glauser G, Erb M, Flors V, Frey M, Ton J. 2011. Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize. Plant Physiology 157(1):317–327 DOI 10.1104/pp.111.180224.
Blümke A, Somerville SC, Voigt CA. 2013. Transient expression of the Arabidopsis thaliana callose synthase PMR4 increases penetration resistance to powdery
mildew in barley. *Advances in Bioscience and Biotechnology* **04(08)**:810–813 DOI 10.4236/abb.2013.48106.

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* **13**(8):1194–1202 DOI 10.1016/j.molp.2020.06.009.

Chen XY, Kim JY. 2009. Callose synthesis in higher plants. *Plant Signaling and Behavior* **4**(6):489–492 DOI 10.4161/psb.4.6.8359.

Conant GC, Wolfe KH. 2008. Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics* **9**(12):938–950 DOI 10.1038/nrg2482.

Cui X, Shen H, Song C, Laosinchai W, Amano Y, Brown Jr RM. 2001. A putative plant homolog of the yeast beta-1, 3-glucan synthase subunit FKS1 from cotton (*Gossypium hirsutum* L.) fibers. *Planta* **213**(2):223–230 DOI 10.1007/s00425000496.

Dong X, Hong Z, Chatterjee J, Kim S, Verma DP. 2008. Expression of callose synthase genes and its connection with Npr1 signaling pathway during pathogen infection. *Planta* **229**(1):87–98 DOI 10.1007/s00425-008-0812-3.

Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, Jia Y, Pan Z, Gong W, Liu Z, Zhu H, Ma L, Liu F, Yang D, Wang F, Fan W, Gong Q, Peng Z, Wang L, Wang X, Xu S, Shang H, Lu C, Zheng H, Huang S, Lin T, Zhu Y, Li F. 2018. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nature Genetics* **50**(6):796–802 DOI 10.1038/s41588-018-0116-x.

Ellinger D, Glockner A, Koch J, Naumann M, Sturtz V, Schutt K, Manisseri C, Somerville SC, Voigt CA. 2014. Interaction of the Arabidopsis GTPase RabA4c with its effector PMR4 results in complete penetration resistance to powdery mildew. *The Plant Cell* **26**(7):3185–3200 DOI 10.1105/tpc.114.127779.

Ellinger D, Naumann M, Falter C, Zwikowics C, Jamrow T, Manisseri C, Somerville SC, Voigt CA. 2013. Elevated early callose deposition results in complete penetration resistance to powdery mildew in *Arabidopsis*. *Plant Physiology* **161**(3):1433–1444 DOI 10.1104/pp.112.211011.

Enrique R, Siciliano F, Favaro MA, Gerhardt N, Roeschlin R, Rigano L, Sendin L, Castagnaro A, Vojnov A, Marano MR. 2011. Novel demonstration of RNAi in citrus reveals importance of citrus callose synthase in defence against *Xanthomonas citri* subsp. *citri*. *Plant Biotechnology Journal* **9**(3):394–407 DOI 10.1111/j.1467-7652.2010.00555.x.

Fang L, Tian R, Li X, Chen J, Wang S, Wang P, Zhang T. 2014. Cotton fiber elongation network revealed by expression profiling of longer fiber lines introgressed with different *Gossypium barbadense* chromosome segments. *BMC Genomics* **15**:838 DOI 10.1186/1471-2164-15-838.

Fromm J, Hajirezaei MR, Becker VK, Lautner S. 2013. Electrical signaling along the phloem and its physiological responses in the maize leaf. *Frontiers in Plant Science* **4**:239 DOI 10.3389/fpls.2013.00239.
Granato LM, Galdeano DM, D’Alessandre NDR, Breton MC, Machado MA. 2019. Callose synthase family genes plays an important role in the Citrus defense response to Candidatus Liberibacter asiaticus. *European Journal of Plant Pathology* 155(1):25–38 DOI 10.1007/s10658-019-01747-6.

Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32(18):2847–2849 DOI 10.1093/bioinformatics/btw313.

He Z, Zhang H, Gao S, Lercher MJ, Chen WH, Hu S. 2016. Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Research* 44(W1):W236–W241 DOI 10.1093/nar/gkw370.

Hong Z, Delauney AJ, Verma DP. 2001. A cell plate-specific callose synthase and its interaction with phragmoplastin. *The Plant Cell* 13(4):755–768 DOI 10.1105/tpc.13.4.755.

Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T, Lian J, Baruch K, Fang D, Liu X, Ruan YL, Rahman MU, Han J, Wang K, Wang Q, Wu H, Mei G, Zang Y, Han Z, Xu C, Shen W, Yang D, Si Z, Dai F, Zou L, Huang F, Bai Y, Zhang Y, Brodt A, Ben-Hamo H, Zhu X, Zhou B, Guan X, Zhu S, Chen X, Zhang T. 2019. Gossypium barbadense and Gossypium hirsutum genomes provide insights into the origin and evolution of allotetraploid cotton. *Nature Genetics* 51(4):739–748 DOI 10.1038/s41588-019-0371-5.

Huang L, Chen XY, Rim Y, Han X, Cho WK, Kim SW, Kim JY. 2009. Arabidopsis glucan synthase-like 10 functions in male gametogenesis. *Journal of Plant Physiology* 166(4):344–352 DOI 10.1016/j.jplph.2008.06.010.

Jacobs AK, Lipka V, Burton RA, Panstruga R, Strizhov N, Schulze-Lefert P, Fincher GB. 2003. An arabidopsis callose synthase, GSL5, Is required for wound and papillary callose formation. *The Plant Cell* 15(11):2503–2513 DOI 10.1105/tpc.016097.

Koh EJ, Zhou L, Williams DS, Park J, Ding N, Duan YP, Kang BH. 2012. Callose deposition in the phloem plasmodesmata and inhibition of phloem transport in citrus leaves infected with “Candidatus Liberibacter asiaticus”. *Protoplasma* 249(3):687–697 DOI 10.1007/s00709-011-0312-3.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33(7):1870–1874 DOI 10.1093/molbev/msw054.

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* 30(1):325–327 DOI 10.1093/nar/30.1.325.

Liu J, Du H, Ding X, Zhou Y, Xie P, Wu J. 2017. Mechanisms of callose deposition in rice regulated by exogenous abscisic acid and its involvement in rice resistance to Nilaparvata lugens Stal (Hemiptera: Delphacidae). *Pest Management Science* 73(12):2559–2568 DOI 10.1002/ps.4655.

Liu X, Zhao B, Zheng HJ, Hu Y, Lu G, Yang CQ, Chen JD, Chen JJ, Chen DY, Zhang L, Zhou Y, Wang LJ, Guo WZ, Bai YL, Ruan JX, Shangguan XX, Mao YB, Shan
CM, Jiang JP, Zhu YQ, Jin L, Kang H, Chen ST, He XL, Wang R, Wang YZ, Chen J, Wang LJ, Yu ST, Wang BY, Wei J, Song SC, Lu XY, Gao ZC, Gu WY, Deng X, Ma D, Wang S, Liang WH, Fang L, Cai CP, Zhu XF, Zhou BL, Jeffrey Chen Z, Xu SH, Zhang YG, Wang SY, Zhang TZ, Zhao GP, Chen XY. 2015. *Gossypium barbadense* genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites. *Scientific Reports* 5:14139 DOI 10.1038/srep14139.

Liu F, Zou Z, Fernando WGD. 2018. Characterization of callose deposition and analysis of the callose synthase gene family of *brassica napus* in response to *leptosphaeria maculans*. *International Journal of Molecular Sciences* 19(12):3769 DOI 10.3390/ijms19123769.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25(4):402–408 DOI 10.1006/meth.2001.1262.

Loake G, Grant M. 2007. Salicylic acid in plant defence—the players and protagonists. *Current Opinion in Plant Biology* 10(5):466–472 DOI 10.1016/j.pbi.2007.08.008.

Lu Q, Shi Y, Xiao X, Li P, Gong J, Gong W, Liu A, Shang H, Li J, Ge Q, Song W, Li S, Zhang Z, Rashid MHO, Peng R, Yuan Y, Huang J. 2017. Transcriptome analysis suggests that chromosome introgression fragments from sea island cotton (*Gossypium barbadense*) increase fiber strength in upland cotton (*Gossypium hirsutum*). *G3-Genes Genomes Genetics* 7(10):3469–3479 DOI 10.1534/g3.117.300108.

Magwanga RO, Lu P, Kirungu JN, Cai X, Zhou Z, Wang X, Diouf L, Xu Y, Hou Y, Hu Y, Dong Q, Wang K, Liu F. 2018. Whole genome analysis of cyclin dependent kinase (CDK) gene family in cotton and functional evaluation of the role of CDKF4 gene in drought and salt stress tolerance in plants. *International Journal of Molecular Sciences* 19(9):2625 DOI 10.3390/ijms19092625.

Oide S, Bejai S, Staal J, Guan N, Kaliff M, Dixielius C. 2013. A novel role of PR2 in abscisic acid (ABA) mediated, pathogen-induced callose deposition in *Arabidopsis thaliana*. *New Phytologist* 200(4):1187–1199 DOI 10.1111/nph.12436.

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo MJ, Byers R, Chen W, Doron-Faigenboim A, Duke MV, Gong L, Grimwood J, Grover C, Grupp K, Hu G, Lee TH, Li J, Lin L, Liu T, Marler BS, Page JT, Roberts AW, Romanel E, Sanders WS, Szadkowski E, Tan X, Tang H, Xu C, Wang J, Wang Z, Zhang D, Zhang L, Ashrafi H, Bedon F, Bowers JE, Brubaker CL, Chee PW, Das S, Gingle AR, Haigler CH, Harker D, Hoffmann LV, Hovav R, Jones DC, Lemke C, Mansoor S, Rahman Mur, Rainville LN, Rambani A, Reddy UK, Rong JK, Saranga Y, Scheffler BE, Scheffler JA, Stelly DM, Triplett BA, Van Deynze A, Vaslin MF, Waghmare VN, Walford SA, Wright RJ, Zaki EA, Zhang T, Dennis ES, Mayer KF, Peterson DG, Rokhsar DS, Wang X, Schmutz J. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492(7429):423–427 DOI 10.1038/nature11798.
Pu Y, Hou L, Guo Y, Ullah I, Yang Y, Yue Y. 2019. Genome-wide analysis of the callose enzyme families of fertile and sterile flower buds of the Chinese cabbage (Brassica rapa L. ssp. pekinensis). FEBS Open Bio 9(8):1432–1449 DOI 10.1002/2211-5463.12685.

Repka V, Fischerová I, Šilhárová K. 2004. Methyl jasmonate is a potent elicitor of multiple defense responses in grapevine leaves and cell-suspension cultures. Biologia Plantarum 48:273–283 DOI 10.1023/B:BIOP.0000033456.27521.e5.

Ruan YL, Xu SM, White R, Furbank RT. 2004. Genotypic and developmental evidence for the role of plasmodesmatal regulation in cotton fiber elongation mediated by callose turnover. Plant Physiology 136(4):4104–4113 DOI 10.1104/pp.104.051540.

Schober MS, Burton RA, Shirley NJ, Jacobs AK, Fincher GB. 2009. Analysis of the (1,3)-beta-D-glucan synthase gene family of barley. Phytochemistry 70(6):713–720 DOI 10.1016/j.phytochem.2009.04.002.

Shi X, Sun X, Zhang Z, Feng D, Zhang Q, Han L, Wu J, Lu T. 2015a. GLUCAN SYNTHASE-LIKE 5 (GSL5) plays an essential role in male fertility by regulating callose metabolism during microsporogenesis in rice. Plant Cell Physiology 56(3):497–509 DOI 10.1093/pcp/pcu193.

Shi ZH, Zhang C, Xu XF, Zhu J, Zhou Q, Ma LJ, Niu J, Yang ZN. 2015b. Overexpression of AtTTP affects ARF17 expression and leads to male sterility in Arabidopsis. PLOS ONE 10(3):e0117317 DOI 10.1371/journal.pone.0117317.

Sun M, Voorrips RE, Steenhuis-Broers G, Van’t Westende W, Vosman B. 2018. Reduced phloem uptake of Myzus persicae on an aphid resistant pepper accession. BMC Plant Biology 18(1):138 DOI 10.1186/s12870-018-1340-3.

Song L, Wang R, Zhang L, Wang Y, Yao S. 2016. CRR1 encoding callose synthase functions in ovary expansion by affecting vascular cell patterning in rice. Plant Journal 88(4):620–632 DOI 10.1111/tpj.13287.

Toller A, Brownfield L, Neu C, Twell D, Schulze-Lefert P. 2008. Dual function of Arabidopsis glucan synthase-like genes GSL8 and GSL10 in male gametophyte development and plant growth. Plant Journal 54(5):911–923 DOI 10.1111/j.1365-313X.2008.03462.x.

Tylewicz S, Petterle A, Marttila S, Miskolczi P, Azeez A, Singh RK, Immanen J, Mähler N, Hvidsten TR, Eklund DM, Bowman JL, Helariutta Y, Bhalerao RP. 2018. Photoperiodic control of seasonal growth is mediated by ABA acting on cell–cell communication. Science 360(6385):212–215 DOI 10.1126/science.aan8576.

Wang X, Lu X, Malik WA, Chen X, Wang J, Wang D, Wang S, Chen C, Guo L, Ye W. 2020. Differentially expressed bZIP transcription factors confer multi-tolerances in Gossypium hirsutum L. International Journal of Biological Macromolecules 146:569–578 DOI 10.1016/j.ijbiomac.2020.01.013.

Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. 2010. KaKs_Calculator 2.0: a toolkit incorporating Gamma-Series methods and sliding window strategies. Genomics, Proteomics & Bioinformatics 8(1):77–80 DOI 10.1016/s1672-0229(10)60008-3.
Wendel JF, Brubaker C, Alvarez I, Cronn R, Stewart JM. 2009. Evolution and natural history of the cotton genus. *Genetics and Genomics of Cotton* 3–22 DOI 10.1007/978-0-387-70810-2_1.

Wu SW, Kumar R, Iswanto ABB, Kim JY. 2018. Callose balancing at plasmodesmata. *Journal of Experimental Botany* 69(22):5325–5339 DOI 10.1093/jxb/ery317.

Xie B, Wang X, Hong Z. 2010. Precocious pollen germination in Arabidopsis plants with altered callose deposition during microsporogenesis. *Planta* 231(4):809–823 DOI 10.1007/s00425-009-1091-3.

Xie B, Wang X, Zhu M, Zhang Z, Hong Z. 2011. CalS7 encodes a callose synthase responsible for callose deposition in the phloem. *Plant Journal* 65(1):1–14 DOI 10.1111/j.1365-313X.2010.04399.x.

Yao L, Zhong Y, Wang B, Yan J, Wu T. 2020. BABA application improves soybean resistance to aphid through activation of phenylpropanoid metabolism and callose deposition. *Pest Management Science* 76(1):384–394 DOI 10.1002/ps.5526.

Yu CS, Chen YC, Lu CH, Hwang JK. 2006. Prediction of protein subcellular localization. *Proteins* 64(3):643–651 DOI 10.1002/prot.21018.

Yu Y, Jiao L, Fu S, Yin L, Zhang Y, Lu J. 2016. Callose synthase family genes involved in the grapevine defense response to downy mildew disease. *Phytopathology* 106(1):56–64 DOI 10.1094/PHYTO-07-15-0166-R.

Yu J, Jung S, Cheng CH, Ficklin SP, Lee T, Zheng P, Jones D, Percy RG, Main D. 2014. CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Research* 42:D1229–D1236 DOI 10.1093/nar/gkt1064.

Zavaliev R, Ueki S, Epel BL, Citovsky V. 2011. Biology of callose (beta-1, 3-glucan) turnover at plasmodesmata. *Protoplasma* 248(1):117–130 DOI 10.1007/s00709-010-0247-0.