

Multi-responses of O-methyltransferase genes to salt stress and fiber development of *Gossypium* species

Abdul Hafeez

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Qún Gě

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Qí Zhāng

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Jùnwén Lǐ

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Jǔwǔ Gōng

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Ruìxián Liú

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Yùzhēn Shí

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Hǎihóng Shāng

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Àiyīng Liú

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Muhammad S. Iqbal

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Xiǎoyīng Dèng

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Abdul Razzaq

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Muharam Ali

Sindh Agriculture University Tandojam

Yǒulù Yuán

Institute of Cotton Research, Chinese Academy of Agricultural Sciences

Gong Wankui (✉ gongwankui@caas.cn)

Institute of Cotton Research, Chinese Academy of Agricultural Sciences <https://orcid.org/0000-0002-8556-0669>

Research article

Keywords: O-methyltransferase, *Gossypium*, fiber development, biotic and abiotic stress

Posted Date: December 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-36601/v3>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at BMC Plant Biology on January 11th, 2021. See the published version at <https://doi.org/10.1186/s12870-020-02786-6>.

Abstract

Background: O-methyltransferases (OMTs) are an important group of enzymes that catalyze the transfer of a methyl group from S-adenosyl-L-methionine to their acceptor substrates. OMTs are divided into several groups according to their structural features. In *Gossypium* species, they are involved in phenolics and flavonoid pathways. Phenolics defend the cellulose fiber from dreadful external conditions of biotic and abiotic stresses, promoting strength and growth of plant cell wall.

Results: An *OMT* gene family, containing a total of 192 members, has been identified and characterized in three main *Gossypium* species, *G. hirsutum*, *G. arboreum* and *G. raimondii*. Cis-regulatory elements analysis suggested important roles of *OMT* genes in growth, development, and defense against stresses. Transcriptome data of different fiber developmental stages in Chromosome Substitution Segment Lines (CSSLs), Recombination Inbred Lines (RILs) with excellent fiber quality, and standard genetic cotton cultivar TM-1 demonstrate that up-regulation of *OMT* genes at different fiber developmental stages, and abiotic stress treatments have some significant correlations with fiber quality formation, and with salt stress response. Quantitative RT-PCR results revealed that *GhOMT10_Dt* and *GhOMT70_At* genes had a specific expression in response to salt stress while *GhOMT49_At*, *GhOMT49_Dt*, and *GhOMT48_At* in fiber elongation and secondary cell wall stages.

Conclusions: Our results indicate that O-methyltransferase genes have multi-responses to salt stress and fiber development in *Gossypium* species and that they may contribute to salt tolerance or fiber quality formation in *Gossypium*.

Background

Cotton (*Gossypium* Species) has the importance for natural fiber all over the globe. The primary goals of upland cotton (*G. hirsutum*) perspectives have been always to achieve better quality with higher yield [1]. Mostly *G. hirsutum* bears staple fibers 25-40 mm in length and 15 μ m in thickness at their full maturity. Fiber cells must undergo four distinct but partially overlapped developmental stages, including initiation, elongation, secondary cell wall deposition, and maturation. The secondary cell wall of fiber, which is mainly composed of cellulose, is important especially for fiber quality perspective. However, some studies have shown that secondary cell wall of fibers of flax (*Linum usitatissimum* L.), ramie (*Boehmeria nivea* L.), and Spanish broom (*Spartium junceum* L.) also contain phenolics along with cellulose. Their fibers are known for their physical properties such as length and strength and have been used for textile purposes. A thicker secondary cell wall was estimated to contain no less than 70% cellulose content while the cotton fiber contains almost 90% cellulose [2, 3]. Lignin is another important component in cell wall [4]. It provides strength to plant cell wall and response to biotic and abiotic stresses in vascular plants [5]. The presence of lignin, which is reported at lower level in secondary cell wall of cotton fibers [6], negatively regulates fiber elongation and secondary cell wall synthesis in cotton. Studies demonstrated that the cotton plants that accumulate less lignin and lignin-like phenolics in mature fibers tend to have longer and stronger fibers [7]. From an active perspective, lignin and phenolics defend the cellulose fiber against dreadful conditions and increase the ability of response to biotic and abiotic stresses, and thus influence the growth and strength of plant cell walls [8]. Previous studies in herbaceous plants demonstrated the involvement of O-methyltransferases (*OMTs*) in lignin biosynthesis [9]. The involvement of *OMTs* mediate normal plant growth in the presence of lignin [10]. The initial *OMT* cDNA was described in 1991 [11], then a series of *OMT* cDNAs have been cloned from diverse plants species, including *Zea mays*, *Arabidopsis thaliana*, *Iris hollandica*, and *Nicotiana tabacum* [12].

According to substrate classification, plant methyltransferases have three major categories, I. O-methyltransferases (*OMTs*), II. N-methyltransferases (*NMTs*), and III. C-methyltransferases (*CMTs*). Category I *OMTs* are further classified into five sub-categories. Sub-category I-a comprises caffeoyl coenzyme A 3-O-methyltransferase (*CCoAOMT*) and caffeic acid 3-O-methyltransferases (*COMTs*), which are involved in methylation in phenylpropanoids. Sub-categories I-b, I-c, and I-d act in methylation of hydroxyl in flavonoid, alkaloids, and myoinositol, respectively. The fifth sub-category I-e takes part in methylation of carboxyl of diverse acids. The results of a study discovered the crystal structure of *OMTs* from *Medicago sativa* [13]. In the light of the explanations, the *OMT* gene that was cloned and characterized from a medicinal plant *Ligusticum chuanxiong* and contained higher ferulic acid was named as *LcCOMT*. The differential expression of *LcCOMT* gene under chilling stress was more than 6-fold higher than that under controlled conditions, suggesting that ferulic acid may increase plant tolerance to cold stress. BLAST analysis showed that *LcCOMT* was 23.9–40.2% similar to *OMTs* of alkaloid, flavonoid, isoflavonoid, and phenylpropanoids [14].

In the whole life cycle of cotton plant, it undergoes various environmental conditions from the cold spring in April when it is sowed to hot mid-summer when it grows rapidly in vegetation and reproduction and to late freezing autumn when it gets mature and is harvested. During the whole growth procedure, the cotton plant maintains an exquisite molecular controls and regulations. But little is known what roles the *OMT* family genes have played in cotton plant especially in early or late growth stage when season transition occurs, or in various stress conditions. Therefore, in this study, we identified the *OMT* family genes in the genome-wide scale and made detailed bioinformatics analysis of gene structure, chromosomal distribution, selection pressure during their evolution, sub-cellular localization, cis-regulatory elements etc, together with their expression profiling in different developmental stages and in responses to various stresses. Their expression profiling in developing fiber cells was verified using RNA sequencing data from RILs, CSSLs, and TM-1 at different fiber development stages. This study could open the way to comprehend the functions of *OMTs* in fiber quality advancement and in cotton plant responses to abiotic stresses, and thus could assume a noteworthy part for further investigation in the molecular mechanism of fiber improvement and stress tolerance.

Results

Genomewide identification and characterization of *OMT* genes

A genome wide analysis was conducted to characterize *OMT* family genes in three *Gossypium* species. A total of 192 *OMT* members were identified, including 82 in *G. hirsutum*, 55 in *G. arboreum*, and 55 in *G. raimondii* (Table S3. Sheet A). For phylogenetic analysis [15], 33 *OMT* members in *A. thaliana*, and 26 members in *T. cacao* species were also retrieved (Table S3. Sheet B). Retrieving information of *OMT* genes in *G. hirsutum* revealed that *GhOMT75_Scaf*, which was detected in scaffold, coded the smallest protein of 62 amino acids (aa) with a molecular weight of 6.642 kDa. While *GhOMT33_Dt*, which was identified on chromosome D_t02, coded the largest protein of 969 aa with a molecular weight of 108.296 kDa among all *OMT* members in three *Gossypium* species.

In domain analysis of *OMT* family genes in *Gossypium* species, the results revealed that 64, 45 and 47 *OMTs* in *G. hirsutum*, *G. arboreum* and *G. raimondii* contained Pfam domain Pf00891, and that only 20, 10 and 9 *OMTs* in *G. hirsutum*, *G. arboreum* and *G. raimondii* contained Pfam domain Pf01596. In *A. thaliana* and *T. cacao*, 25, 24 *OMTs* contained Pf00891 domain, and 8, 2 *OMTs* contained Pf01596 domain respectively.

Chromosomal distribution, collinearity, duplication, and loss of *OMT* genes

The analysis of chromosomal positioning was performed by using TBtool software [16]. A total of 161 *OMT* genes were positioned on their respective chromosomes, while seven of *G. raimondii*, one of *G. arboreum*, and 23 of *G. hirsutum* were positioned in scaffolds (Figure S1). In *G. raimondii* (D genome), chr11 was mapped with 13 genes followed by chr08 with nine genes. The minimum number of genes in a chromosome was one in chr2, chr6, and chr10 respectively. There was no *OMT* family members identified in chr01 and chr07 (Figure S1.a). In *G. arboreum* (A-genome) (Figure S1.b), 54 *OMT* genes were mapped in all chromosomes except chr1. Chr10 harbored 13 *OMT* genes which were the highest per chromosome, followed by chr12 and chr04 with 10 and 9 genes respectively. The minimum number of genes located in a chromosome was one in chr02 and chr11 respectively. In *G. hirsutum* (A_tD_t genome) (Figure S1.c), unexpectedly, there were no *OMT* genes in A_t02, A_t05, A_t07, D_t03, D_t09, and D_t11 chromosomes. The distribution of genes in D_t sub-genome (33 genes) was higher than in A_t sub-genome (26 genes). The maximum number of genes in a chromosome was seven in D_t04 and A_t12, followed by four in D_t10 and A_t10 chromosomes, respectively. D_t01, D_t05, A_t01, A_t06, and A_t11 only had one *OMT* gene, and D_t06, D_t07, A_t03, A_t08, A_t09, A_t13 two *OMT* genes and D_t02, D_t08, and D_t13 three *OMT* genes respectively (Figure S4.c). A collinearity analysis of the *OMT* family genes in *Gossypium* species chromosomes was shown in Figure 1. The results demonstrated a pair wise collinearity of *OMT* genes between the chromosomes on which *OMT* family genes were mapped. Noticeably, a number of available genes in A_t and D_t scaffolds were collinear with their homologues in A and D genomes suggesting the collinearity of the DNA fragments between the scaffolds and chromosome where these *OMT* genes locate (Figure 1). Taken the *OMT* gene numbers identified in each A/D genome or A_t/D_t sub-genome, collinearity analysis also revealed that there were totally 21 and 19 *OMT* genes exclusively detected in A and D genomes respectively. Their homologous counterparts in A_tD_t genomes of *G. hirsutum* are lost. There are also a few *OMT* genes that are exclusively detected in A_tD_t genome of *G. hirsutum* without homologous counterparts in A and D genomes (Figure S1).

According to previous studies there are five types of duplications including singleton, dispersed, proximal, tandem, and segmental or whole-genome duplication [17]. In the present study, the analysis of gene pairs duplication events predicted a total of 31, 28, and 54 gene pairs of D_tD_t, A_tA_t and D_tA_t genomes from their common ancestor, 33 gene pairs of A_tD_t subgenomes in segmental duplication, and 5 gene pairs of A_tD_t subgenomes in tandem duplication events (Table S4 Sheet A).

Analysis of selection pressure

In genetics, the Ka/Ks ratio used to estimate the balance between neutral mutations, purifying selections, and positive mutations based on a set of homologous genes [18]. The ratio of the number of non-synonymous substitutions per non-synonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks) represents selection pressure of the gene [19]. Ka/Ks<1 demonstrates purifying selection pressure, while Ka/Ks=1 and Ka/Ks>1 show neutral and positive selection pressures respectively. Analysis of Ka/Ks ratio of homologous *OMT*s in three *Gossypium* species revealed that they are under purifying selection pressure. The Ka/Ks ratio of homologous *OMT*s in *G. raimondii* and *G. arboreum* ranged from 0.09 to 0.8, in *G. raimondii* and *G. hirsutum* ranged 0 to 0.7, and in A_t and D_t of *G. hirsutum* ranged 0.4 to 0.7 (Table S4 Sheet B).

Sequences alignment, phylogenetic Analyses, conserved motifs and gene structure

The sequence alignment of 251 *OMT* genes, including 192 genes from three *Gossypium* species, 33 from *A. thaliana*, and 26 from *T. cacao* species was performed to understand the phylogenetic relationship of these genes. The evolutionary relationship of *OMT* genes in three *Gossypium* species was monophyletic (Figure 2.a), and the member of *A. thaliana* and *T. cacao* were distributed in paraphyletic manner (Figure 2.b). According to the topology of constructed tree, the *OMT* gene family is divided into five clades (I, II, III, IV, and V) in *Gossypium*, *A. thaliana*, and *T. cacao* species. The results showed that each clade of *OMT* genes were symmetrically distributed within *Gossypium* species (Figure 2.a), while in *A. thaliana* and *T. cacao*, *OMT* genes were identified in cluster forms (Figure 2.b). The results demonstrated that these *Gossypium OMT* members might be evolutionary close within respective species and their identified clades.

To examine the conserved motifs of each clade, the analysis of representative motif logo and conserved motifs prediction were conducted (Figure S2). The results revealed that motif 1, enriched with leucine, valine, and glycine, motif 2, enriched with leucine and valine, and motif 3, motif 4, motif 5, and motif 6 were common in clades I, II, III, IV and clade V. While motif 7 was found missing in some members of clade V, which was then replaced with motif 8 at same positions (Figure S2). The enriched amino acid residues of conserved motif 1 (L/VDVGGG/TG) was previously identified in S-adenosyl-L-methionine (SAM)-dependant *OMT*s that shared 95% similarity with *G. hirsutum OMT* [20].

Investigation of gene structure has uncovered the different number of exons and introns of *OMT* genes. Exon and intron number of *OMT* genes varied from the least one exon and no intron to the most 7 to 9 exons and 6 to 8 introns (Table S5). Two members in *G. hirsutum* including *GhOMT82_At* and *GhOMT82_Dt* contain nine exons as the highest (Table S5). Same as, two members including *GaOMT82_A* in *G. arboreum*, and *GrOMT82_D* in *G. raimondii* contain 9 exons (Table S5). Gene structure analysis revealed that the *OMT* genes with higher number of exons had a shorter exons and introns, and vice versa. These results demonstrated that *OMT* members possess different structural patterns in accordance with their features.

Identification of cis-regulatory elements in *OMT* family

The promoter regions of the *OMT* family contain precisely a large number of cis-regulatory elements. The analysis of cis-regulatory elements revealed the enrichment of MYB cis-regulatory elements, which was detected more than 350 times in *OMT* genes (Figure S3). The MYC was another important element that was found 183 times in enlisted *OMT* genes. Box 4 (part of a conserved DNA module involved in light responsiveness) was found 152 times in 43/82 genes in *G. hirsutum*. ABRE elements was detected 119 times in 29/82 *OMT* genes in *G. hirsutum*. The ERE element was detected 113 times in 37/82 and G-Box 97 times in 37/82 *G. hirsutum OMT* genes. An auxin RR-core and cis-acting regulatory element involved in the MeJA-responsiveness (TGACG-motif) were also observed in *Gossypium OMT* genes where this element was identified 48 times in 25/82 genes. Some other important cis-regulatory elements including wun-motif 44 times in 26/82, W-box 39 times in 31/82, GATA-motif 32 times in 27/82, O2-site 30 times in 22/82 *OMT* genes respectively, in *G. hirsutum* (Figure S3). These cis-regulatory elements might function collectively in accordance with their specific roles and with specific conditions as well as growth and development stages (Figure S3).

Sub-cellular localization prediction of *OMT* genes

Understanding and determining the sub-cellular localization of proteins is an important strategy to identify the function of protein at cellular level [21]. This approach includes proteomic-based experiments and microscopic high throughputs [22, 23]. Several sequence-based approaches have been developed to predict the sub-cellular localization by providing amino acid sequences including PSORT [24], Yloc [25], BaCelLO [26], LOctree [27]. According to CELLO prediction, most of *OMT* genes were located in the cytoplasm (Table 1), while seven genes were predicted in periplasm, including, *GhOMT45_At*, *GhOMT45_Dt*, *GhOMT46_Dt*, *GhOMT48_At*, *GhOMT48_Dt*, *GhOMT49_At*, and *GhOMT49_Dt*. Five *OMT*s were predicted to be localized in both periplasm and cytoplasm, including *GhOMT47_At*, *GhOMT47_Dt*, *GhOMT53_At*, *GhOMT54_Dt*, and *GhOMT68_At*. Two genes *GhOMT82_At* and *GhOMT82_Dt* were predicted in the outer membrane. Only *GhOMT55_At* was predicted in inner membrane and cytoplasm (Table. 1). The results of Wolf Psort were highly in agreement with those of CELLO analysis regarding the presence of most of the *OMT* genes in cytoplasm, however, with exceptions of *GhOMT48_At*, *GhOMT82_At*, *GhOMT82_Dt*, which were predicted in chloroplast and one gene *GhOMT76_Dt* in mitochondria (Table. 1). The function of the *OMT* genes might be related to their predicted localizations, though the experimental approach is still needed for further confirmation.

GO enrichment and KEGG pathway analyses

To understand the functional annotations of *OMT* family genes of *G. hirsutum*, 82 genes in *G. hirsutum* were undergone through gene ontology (GO) enrichment, kyoto encyclopedia of genes and genomes (KEGG Pathway), and InterPro analyses. GO term analysis verified their O-methyltransferase activity of all 82 *OMT* genes, while 62 of the 82 genes were also enriched in methyltransferase activity and 53 of the 82 genes in protein dimerization activity (Figure 3.a). KEGG Pathway analysis revealed that these *OMT*s were involved in different metabolic pathways. Twenty-nine *OMT*s were involved in monolignol biosynthesis, phenylpropanoid, secondary metabolism, and metabolic pathways respectively. Eleven genes were involved in phenylalanine and flavonoid biosynthesis pathways respectively (Figure 3.b). InterPro analysis (<http://www.ebi.ac.uk/interpro/>) categorized these 82 *OMT* genes as functional genes of S-adenosyl-L-methionine-dependent methyltransferase (Figure 3.c). Sixty-two genes were also predicted in categories of methyltransferase_2 and O-methyltransferase COMT-type respectively (Figure 3.c), while fifty-seven in winged helix-turn-helix DNA-binding domain, fifty-three in plant methyltransferase dimerization (Figure 3.c).

Expression profiling of *OMT* genes and their homologues in fiber development and salt stress

In order to verify the biological functions of *OMT* family genes, several transcriptome data sets including TM-1 [28], *G. arboreum*, *G. raimondii*, CSSLs [29], and RILs [30], were applied to analyze their expression profiles in different developmental stages, organs, or tissues, and responses to various abiotic stress treatments. The transcriptome clusters showed that the *OMT* genes can be assorted into three basic groups (Figure 4A): Those that have a broad responses to different developmental stages from germination to fiber maturation, typical examples of which included *GhOMT48_At*, *GhOMT48_Dt*, *GhOMT49_At* and *GhOMT49_Dt*; those that have specific responses to root development, including *GhOMT1_At*, *GhOMT2_At* and *GhOMT40_At*; and those that have responses to early germination in seed, cotyledon, root and stem, including *GhOMT47_At*, *GhOMT9_Dt*, and *GhOMT58_Dt*. When fiber specific transcriptome data sets of *G. arboreum*, *G. raimondii* were applied to observe the expression profiling diploid *OMT* family genes, the result also supported specific expression profiling of some *OMT* genes in diploid species of *G. arboreum* (Figure 4.B) and *G. raimondii* (Figure 4.C).

The gene expression profiling was further verified with transcriptome datasets of RILs (Figure 5.A) two CSSLs (Figure 5.B and 5.C). The results showed that the genes that had specific expressions during fiber development (Figure 4.A) also had specific expressions in fiber development of RILs and CSSLs materials. These genes had a highly consistent expression profiling among the different cotton cultivars and lines during fiber development. Some selected *GhOMT* examples genes, *GhOMT49_At* (Figure 5.D), *GhOMT70_At* (Figure 5.E), *GhOMT48_At* (Figure 5.F), *GhOMT10_Dt* (Figure 5.G), and *GhOMT49_Dt* (Figure 5.H), were verified through qRT-PCR using sGK9708 and 0-153, the two parental lines of the RIL population with different fiber quality traits. The results showed that *GhOMT48_At*, *GhOMT49_At*, and *GhOMT49_Dt* were significantly up-regulated during fiber development in sGK9708 than in 0-153 (Figure 5.D, 5.F and 5.H) and that *GhOMT70_At* and *GhOMT10_Dt* did not show differences between the two cultivars (Figure 5.E and 5.G). Noticeably, *GhOMT49_At* and *GhOMT49_Dt* reached the highest expression levels at 20 DPA and their high expression lasted in a short time as compared with that of *GhOMT48_At*. *GhOMT48_At* had a rapid expression increase from 10 DPA to 15DPA and then its expression steadily increased until 25 DPA when it reached its highest expression level.

Based on the expression profiling of the *OMT* gene family in responses to cold, hot, osmotic, and salt stress treatments (Figure 6.A), two genes specific in salt stress responses, *GhOMT70_At* and *GhOMT10_Dt*, and three genes specific in fiber development, *GhOMT48_At*, *GhOMT49_At*, and *GhOMT49_Dt*, were verified by qRT-PCR with RNA samples extracted from salt treatment. The results indicated that both *GhOMT70_At* and *GhOMT10_Dt* showed an elevated expression in salt treatments in salt-tolerant cultivar as compare to the control treatments (Figure 6.B and 6.C). These two genes had different expression profiles from 2h to 6h after salt treatment. *GhOMT70_At* had the highest expression at 2h and then its expression went down at 6h; whereas *GhOMT10_Dt* had an increasing expression pattern from 2h to 6h. Both genes had much higher expression in roots than in stem or leaf.

Discussion

A genome-wide survey of *OMTs*

A genome wide search of *G. hirsutum* [28], *G. arboreum* [31], and *G. raimondii* [32] resulted in the identification of 192 genes (82 in *G. hirsutum*, 55 in *G. arboreum*, and 55 in *G. raimondii*). Recent study testified that modern allotetraploid *Gossypium* species were developed from a natural hybridization between the ancestors of two diploid species of *G. raimondii* (D-genome) [33] and *G. arboreum* (A-genome) [31] 1.7 to 1.9 million years ago [28]. The results of current study revealed a loss of quite a large number of *OMT* genes in *G. hirsutum* A_tD_t genome as compared to the total number of *OMT* genes in A and D genomes. Possibly 19 *OMT* genes in A_t sub-genome and 17 in D_t sub-genome in *G. hirsutum* (Figure S1) were lost during the evolution procedure after it arose from above mentioned hybridization [28]. Gene losses can be the result of premature stop codon, disruption of genes as compared to their orthologous [34], and rapid genome re-organization during polyploidization and diploidization process [35-37]. Previous studies have evidenced that polyploidization processes may result in losing of homologous members or altered expression profiles of the homologous genes or both [38-41]. Similar phenomenon was noticed in the expression profiling of homologous *OMT* genes between A_t and D_t in *G. hirsutum*, which clued that these genes might have experienced abovementioned events during evolution processes. Collectively, a higher number of genes were also identified in whole genome duplication event. The whole genome duplication may have resulted from an organism that inherited two genomes from each parent. Whole genome duplication events results duplicate genes that may lost through fractionation [42]. Besides the whole genome duplication, segmental duplication events were also identified with a large number in *OMT* gene pairs. Segmental duplication is widespread in flowering plants, which might lead to the evolution of novel genes and their functions [43].

Phylogenetic analysis showed high similarity and monophyletic distribution of *OMTs* within *Gossypium* species that might support the conservative evolution mode of *OMT* genes within five phylogenetic clades. Previous study also reported five clades of *OMT* genes in *Catalpa bungei* [44]. Analyses of selection pressures revealed that most of *OMT* genes in *Gossypium* species were under a purifying selection pressure. The purifying selection pressure might suggest the importance of *OMT* genes in *Gossypium* species. But noticeable exceptions were also observed in some interspecific homologous pairs, in which their Ka/Ks values were >1, indicating these homologous pairs were under a positive selection pressure. These homologous pair exceptions included *GrOMT52_D-GhOMT52_Scaf* and *GrOMT29_D-GhOMT29_At* in *G. raimondii* and *G. hirsutum*, *GaOMT30_A-GhOMT30_At* in *G. arboreum* and *G. hirsutum*, *GrOMT63_D-GaOMT64_A* and *GrOMT29_D-GaOMT29_A* in *G. raimondii* and *G. arboreum*. These results suggested that the *OMT* genes might had experience positive selection pressures during the evolution from diploids to tetraploids. Previous studies have evidenced that the positive selection pressure might be associated with the onsets of new functions in genes [45, 46]. Considering the fact that quite a proportion of *OMT* genes were lost during the formation and evolution of allotetraploid cotton (see afore discussion and Figure S1). In the current study, six *OMT* family members in *G. hirsutum*, one in *G. arboreum* and nine in *G. raimondii* were characterized as (R,S)-reticuline 7-O-methyltransferase. *OMTs* convert reticuline to laudanine in tetrahydrobenzylisoquinoline biosynthesis in the opium poppy *Papaver somniferum*, however, this enzymatic activity is unknown in most higher plants [47]. Therefore, how these genes function is still open to discussion. Taken all findings together, the results might suggest that the *OMTs* that experienced positive selective pressure be lost or take on some novel functions in *G. hirsutum* during the processes of its evolution and ancestor formation.

Previous findings have reported that the *G. raimondii* (D-genome) and *G. arboreum* (A-genome) are the closest relatives to the D_t and A_t sub-genomes of allotetraploids, respectively [28]. Each gene in A or D genome will always have a homolog in the correspondent A_t or D_t sub-genomes of *G. hirsutum* [48]. However, in both A and D genomes we detected quite a large number of *OMT* genes that do not have homologs in their relative A_t and D_t sub-genomes (Figure S1). Previous studies evidenced that such

homolog loss could result from two possible reasons: one is that the homologs were lost during the procedure of polyploidization from diploids to tetraploid; the other is that after the tetraploid formation, the *OMT* members in each genome started their separate evolution procedure. This separate evolution procedure makes the newly evolved members have no homologs in its relative genomes [28]. Previous studies revealed that in A, D, A_tD_t genomes do not maintain same speed of evolution. A faster evolution rate was observed in allotetraploid cottons than in diploid cottons [28]. Taken the fact that *OMT* genes undergo purifying selection procedures (Table S4. Sheet B), the first reason is possibly endorsed as the main cause for the current evolution status of *OMT* gene family and the second reason may also played a role.

Function prediction of *OMT* candidates

OMTs are involved in diverse cis-regulatory elements

Plants encounter various biotic and abiotic stresses during their entire life cycles that negatively affect growth, development, and productivity [49]. Under exposure of these stresses, plants require some potential mechanism, which can be activated in critical circumstances, to support whole plant life cycle [50]. Excessive salinity is also a major factor that affects the cotton production all around the world [51]. Identification of cis-regulatory elements revealed that the *OMT* genes are enriched with important cis-regulatory elements that are essential against negative environmental stresses. Some important regulatory elements, including W-box, MYB, MYC, DRE, ABRE, G-Box, MBS [52], were identified in *OMT* genes. W-box is important to regulate the expression of genes and to bind *WRKY TFs*. *WRKY TFs* are important to mediate plants to defense against chilling, wounding, drought, salinity and heat stresses [53-61]. MYB and MYC have been identified as involved in dehydration-response [62]. DRE [63], which up-regulate gene expression under cold stress and increase the tolerance of plants was also identified in these specific genes. ABRE is an important regulatory element that enhances salt stress tolerance in plants. It plays a key role in dehydration and in response to salinity stress in *Arabidopsis thaliana*, soybean and rice, and in response to chilling or cold in *Paeonia suffruticosa* [64]. G-box is identified in several gene promoters in previous studies and it contributes to development, hormone response, and tolerance against fungal infections in plants. Besides, a gibberellins response element (GARE) was also identified to be important to promote flowering in plants. The auxin hormones play a major role in growth and development of diverse plant species [65]. These results were in accordance with our findings. Especially the repetitively identified cis-regulatory elements might have biological functions in plants under specific conditions and development stages.

OMTs are possibly involved in secondary metabolic pathways

The KEGG pathways enrichment analysis revealed the involvement of *OMT* genes in secondary metabolism and metabolic pathways including monolignol, phenylpropanoid, flavonoid, and phenylalanine metabolisms. Secondary metabolic pathways are demonstrated to have exceptional impacts on biotic and abiotic stresses. Secondary metabolites are phytochemicals, which are synthesized through secondary metabolism. In plants, phenylpropanoids are categorized in several groups such as phenolic acids, flavonoids, and lignins, which are involved in diverse physiological processes and tolerance under unfavorable conditions [66-70]. The activity of secondary metabolites increases during the response of abiotic stresses. These phenolics provide plants with higher tolerance against heavy metals [71, 72], salinity [73], drought [74], and temperature stresses [69]. These pathways also play an important role in plant cell elongations [75, 76]. Same as, plant *OMT* genes have been identified in secondary metabolism [77]. Higher expression of secondary metabolic pathways related genes in developing cotton fiber is reported in previous studies [78, 79]. Importantly, *OMT* genes were reported to be involved in lignin synthesis and to be induced by inoculation of *Verticillium dahliae* in cotton [80-82]. During the inoculation of pathogens, changes in the expression patterns of phenylpropanoid related *OMT* genes were identified. These identified *OMT* genes included *GhOMT53_At*, *GhOMT58_Dt*, *GhOMT61-Dt*, and *GhOMT78_Dt* that were found significantly expressed in 12 and 48 hours postinoculation *V. dahliae* [83]. In the current study, these genes were down-regulated under abiotic stresses and in fiber development stages (Figure 4). Previous reports have evidenced that desoxyhemigossypol-6-O-methyltransferase (dHG-6-OMT) catalyzed the biosynthesis of terpenoid and provided an effective defense mechanism to cotton plant against biotic stresses including insects and pathogens [84]. In response to *V. dahliae* (V991) in CSSLs lines CCRI36 and MBI8255, diverse genes were found differentially expressed in lignin biosynthesis including *CCoAOMT*, which can adequately utilize lignin and has been characterized in several previous studies [85, 86]. Another study also reported that *CCoAOMT* was up-regulated in response to *Verticillium* pathogen in cotton and rendered cotton plants a comparable phenotypic resistance as compared to control plants [87]. A RNA-seq analysis based research identified differential expression patterns of *CCoAOMT* in response to *V.*

dahliae, confirming the effect of this *OMT* gene in the plant response to *V. dahliae* in cotton [88]. These results consequently evidenced the important role of secondary metabolic pathways and *OMT* genes in biotic stresses in cotton.

OMTs are possibly involved in plant growth, abiotic stress tolerance, and fiber development of cotton

Salinity is one of the major causes to reduce crop yield [89] and incurs up-regulation and/or down-regulation of plant genes in response [90]. The *OMT* genes have been found specific for salt stress tolerance and fruit development in tomato plant (*Solanum lycopersicum*) [91]. The SAM-dependent methyltransferases genes were identified to play important role in sweet potato (*Ipomoea batatas*) in response to salt stress [92]. In wheat, *TaCOMT-3D* contributes to stem mechanical support [93]. Another *TaCOMT* gene was also observed with constitutive expression in stem along with leaf and root [94]. The *OMT* gene (*BdCOMT1*) was strongly expressed in stem node and internode but poorly expressed in other tissues in *Brachypodium distachyon* plant [95]. The expression profiles of *OMT* gene family in the transcriptome data of TM-1 [28] and verification results through qRT-PCR also suggested that two *OMT* members *GhOMT10_Dt* and *GhOMT70_At* might contribute to salt stress tolerance in *G. hirsutum*. In the qRT-PCR verifications, *GhOMT10_Dt* and *GhOMT70_At* showed different expression profiling from 2h and 6h after 200 mM NaCl treatment (Figure 5). Probably they act differently in response to salt stress in *G. hirsutum*. Five genes including *GhOMT1_At*, *GhOMT41_At*, *GhOMT47_At*, *GhOMT17_Dt*, and *GhOMT37_Dt* had significant expressions in stem (Figure 4.A) where they might be the potential candidates to provide structural support and survival to plant in environmental stresses.

Cotton fiber quality of is an important attribute to develop elite cultivars in the presence of negative environmental factors. Studies demonstrated that *GhOMT48_At* and *GhOMT49_At* were expressed at elongation stages of a CSSL (CS-B25) and TM-1 respectively [28, 96]. In the current study, the fiber specific *OMT* genes were consistently identified across various populations and species including TM-1 (Figure 4.A) *G. arboreum* (Figure 4.B) *G. raimondii* (Figure 4.C), RILs (Figure 5.A), CSSLs (Figure 5.B, 5.C). They also showed highly similar expression patterns in different fiber development stages. The expression specificities of *GhOMT48_At*, *GhOMT49_At*, and *GhOMT49_Dt* in developing fibers were further verified through qRT-PCR studies (Figure 5.D, 5.F and 5.H). The results demonstrated that these *OMT* members could have a significant function in fiber development and fiber quality formation. But how these genes function during fiber quality formation was still open to discussion.

Lignins-like phenolics are widely studied in response to stress [97]. Recent research advancements revealed that lignin or phenolics influence fiber development at elongation and secondary cell wall synthesis stages [98]. The knock-down of Lignin-like phenolics related gene (*GhbHLH18*) in *G. hirsutum* evidenced the regulation of lignin-like phenolics pathway genes including a *COMT* and others, during cotton fiber elongation and secondary cell wall synthesis stages. The results demonstrated the roles of these genes in regulating the lignification in developing cotton fibers [7]. This study has gathered important information of *OMT* gene family which is a forward step in research to uncover the possible functions or to support previous studies in exploration the functions of *OMT* genes in plant response to salt stress and in cotton fiber development.

Conclusions

Methyltransferases are versatile class of enzymes. *OMT* contributes to diverse phenolics that are essential for plant growth and serves as protective shield against several kinds of stresses. Various bioinformatics analyses revealed that *OMT* gene family is a strong growth regulator, which not only provide protection to the plant, but also are involved in fiber elongation and secondary cell wall synthesis stages. Furthermore, expression profiling analysis based on several transcriptome data and qRT-PCR validation inferred that *GhOMT10_Dt* and *GhOMT70_At* might be the potential candidates for salt stress tolerance and that *GhOMT48_At*, *GhOMT49_At*, and *GhOMT49_Dt* might have significant influence in fiber development at elongation and secondary cell wall thickness stages of *G. hirsutum*. This proposed study concludes the important roles of *OMT* family genes in cotton fiber development and in salt stress tolerance.

Methods

Identification of *OMT* protein family members, sequences alignment, and phylogenetic tree construction

Genome data of three *Gossypium* species including *G. arboreum* (CRI), *G. raimondii* (JGI), and *G. hirsutum* (NAU) were downloaded from cotton functional genomic database (<https://cottonfgd.org/>) [99]. Genome data of *A. thaliana* (Athaliana/TAIR10,

<https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Athaliana#>) [100] and *T. cacao* (Tcacao/v2.1, <https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Tcacao#>) [101] were also downloaded for comparative analysis of *OMT* genes. The hidden Markov model profiles (PF00891 and PF01596) were downloaded from Pfam database (<https://pfam.xfam.org/>). The hmmsearch program of HMMER 3.0 software [102] was used to search for protein sequences of three *Gossypium* species with the E-value of $1e^{-5}$. *OMT* protein sequences of *A. thaliana* and *T. cacao* were also retrieved from the Phytozome database (<https://phytozome.jgi.doe.gov/pz/portal.html>) for phylogenetic analysis. The proteins with absence of required domains were manually removed. Other features of *OMT* genes including protein length (aa), and molecular weight (kDa) were characterized by using cotton functional genomic database (<http://www.cottonfgd.org/>) [99]. The full length amino acid sequences of *G. hirsutum*, *G. arboreum*, *G. raimondii*, *A. thaliana*, and *T. cacao* encoded by *OMT* genes were aligned with clustalx2 software (<http://www.clustal.org/>) [103] with default parameters for the neighbor-joining phylogenetic tree as 1000 bootstraps. Subsequently, two neighbor-joining phylogenetic trees were generated by using Mega7 [104]. The topology of both phylogenetic trees was confirmed to understand the phylogenetic relationship within the five plant species.

Nomenclature of these members was based on their chromosomal locations and numbers in each *Gossypium* species.

Chromosomal mapping and collinearity analysis

TBtool was used to perform the chromosomal mapping of the given *OMT* genes, to search the homologous pairs of *OMT* genes between genomes of the three *Gossypium* species through protein-protein blast (E-value $1e^{-5}$). Circle gene viewer model of TBtool software was used to visualize the results of collinearity between homologous gene pairs [16].

Gene structure and conserved motifs

The structure of the *OMT* genes was analyzed using the online server of Gene Structure Display (GSDS 2.0, <http://gsds.cbi.pku.edu.cn>) [105]. The conserved motifs were predicted online in MEME web based motif prediction tool version 5.0.5 (<http://meme-suite.org/>) by providing protein sequences of *OMT* genes [106].

Selection pressure, cis-regulatory elements, sub-cellular localization and gene enrichment analysis

The CDS of homologous gene pairs of *G. hirsutum* (NAU), *G. arboreum* (CRI), and *G. raimondii* (JGI) were assigned to TBtools software to estimate the Ka/Ks ratio to predict selection pressure between the genes of each pair in genomes and sub-genomes [16]. The upstream sequences (2000bp) of *OMT* genes were retrieved through cotton functional genomic database (<http://www.cottonfgd.org>) and were submitted to PlantCARE database [107] to obtain the cis-regulatory elements. Sub-cellular localization of genes was predicted using online bioinformatics tools CELLO v.2.5 and Wolf Psort with their protein sequences [24, 108]. KEGG IDs of *OMT* family genes were downloaded from cotton functional genomic database (<http://www.cottonfgd.org>), then annotation was performed by providing KEGG IDs in kyoto encyclopedia of genes and genomes database (<https://www.genome.jp/kegg/>) [109]. Gene ontology (GO) annotation IDs of *OMT* family genes were downloaded from cotton functional genomic database (<http://www.cottonfgd.org>) and were submitted to Gene ontology database (<http://geneontology.org/>) to perform GO analysis [110].

Expression profiling of *OMT* genes

The different sets of RNA sequencing data including TM-1, a genetic standard line of *G. hirsutum* (Nanjing Agricultural University, Nanjing, Jiangsu, China) (PRJNA248163) [28, 41], 69307 and 69362 (selected lines from a RIL population sGK9708×0-153, Institute of Cotton Research, Anyang, Henan, China) (PRJNA542946) [30], MBI7747, MBI7561, and MBI7285 (selected lines from CSSL population CCRI45×Hai1, SRP084203) [111, 112], and MBI9915 and MBI9749 (selected lines from CSSL population CCRI36×Hai1, SRX2843778) (Institute of Cotton Research, Anyang, Henan, China) [29, 112] were included in this study to observe the expression pattern of *OMT* family genes at different growth stages, under abiotic stress treatment stages, ovule development, and in different fiber development stages of cotton. Briefly, 69307, 0-153, MBI7747, MBI7561, MBI9915, MBI9749, and Hai1 have high fiber quality traits, while 69362, MBI7285, sGK9708, CCRI36, and CCRI45 have low fiber quality traits. Detailed information of these referenced materials is presented in Table S1.

Transcriptome data of *G. arboreum* (PRJNA179447) [31], and *G. raimondii* (PRJNA79005) [33] were also included to compare the comparative expression of these *OMT* genes.

Plant material, RNA isolation, cDNA synthesis, and qRT-PCR

Upland cotton cultivar 0-153 had elite fiber quality while sGK9708 had high yield potential and wide adaptability. They are successfully used to tag fiber quality and yield QTLs in our previous reports [113-115]. In the current study, sGK9708 and 0-153 (Table S1) were planted in April 2018 in the experimental fields of the Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, Henan. Flowers were tagged on the day of anthesis for fiber sampling in July 2018. Bolls of tagged flowers were sampled in the morning between 9:00 and 10:00 AM at 10, 15, 20, and 25 days' post anthesis (DPA). The fibers were dissected from the developing seeds right after boll picking and immediately stored at -80 °C for RNA extraction.

To examine the expression profiling of *OMT* genes under salt stress, seeds of sGK9708 cultivar were germinated in wet filter papers for 72 hours and then were transferred to hydroponic conditions. The seedlings were treated with 200mM NaCl at three leaves stage. The true leaves, stems, and roots were sampled at 0 hr, 2 hrs, and 6 hrs of the treatment. The 0h of treatment was considered as control sample to compare the expression profiling with treated samples.

Total RNA isolation was performed with the RNeasy Pure Plant Kit by (Qiagen, Beijing, China). To eliminate the genomic DNA contamination, the RNA samples were treated with DNaseI. RNA concentration and integrity was observed on Nano Drop 2000 spectrophotometer (Thermo scientific, USA) and 1% agarose gel electrophoresis. cDNAs of the RNA samples that the A260/280 ratio reached 2.00 were synthesized using PrimeScript® RT Reagent Kit (Perfect Real Time, Takara Biotechnology Co., Ltd., Dalian, China). qRT-PCR was performed with ABI 7500 fast Real-Time PCR system (Applied Biosystems, USA), with *Gh-Histone3* gene was used as reference to normalize the relative expression level. Primers pairs of five *OMT* genes were designed by using Oligo 7 [116] (Table S2). $2^{-\Delta\Delta Ct}$ method was used to calculate the gene expressions [117].

Abbreviations

aa: amino acid; DPA: days post anthesis; CCoAOMT: caffeoyl coenzyme A 3-O-methyltransferase; CMTs: C-methyltransferases; COMTs: caffeic acid 3-O-methyltransferases; NMTs: N-methyltransferases; OMT: O-methyltransferase; kDa: kilodalton; GO: gene ontology; KEGG: kyoto encyclopedia of genes and genomes; Ka: the number of non-synonymous substitutions per non-synonymous site; Ks: the number of synonymous substitutions per synonymous site; CSSLs: chromosome segment substitution lines; RILs: recombinant inbred lines; SAM: S-adenosyl-L-methionine; Ga: *Gossypium arboreum*; Gh: *Gossypium hirsutum*; Gr: *Gossypium raimondii*; qRT-PCR: quantitative real-time polymerase chain reaction

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Competing interests

Authors declare that they have no competing interests for the publication of the manuscript.

Funding

This research was funded by the National Key R&D Program of China (2017YFD0101603, 2016YFD0101401, 2016YFD0100500), the Natural Science Foundation of China (31471538 and 31371668), the Agricultural Science and Technology Innovation Program for CAAS (CAAS-ASTIP-ICRCAAS), the National High Technology Research and Development Program of China (2012AA101108 and 2009AA101104) and the Central Level of the Scientific Research Institutes for Basic R & D Special Fund Business (1610162014008). The funding agencies provided only the expenses and fees of this study. The experimental design and collection, analysis, and interpretation of data and manuscript writing were performed by the contributing authors.

Author contributions

AH and WG designed the project, QG, QZ, MSI and XD performed the formal analysis, JL, JG, YS, HS, RL, ÀL and AR provide the methodology, QZ and QG performed the software, YY and WG supervised the study, AH and WG drafted the manuscript, MA, YY and WG reviewed & edited the manuscript. All authors have read and approved the manuscript.

Acknowledgements

The authors acknowledge SciLang for its linguistic assistance during the preparation of this manuscript and the anonymous reviewers for their valuable comments and helpful suggestions which help to improve the manuscript.

References

1. Lee JJ, Woodward AW, Chen ZJ: **Gene expression changes and early events in cotton fibre development.** *Annals of botany* 2007, **100**(7):1391–1401.
2. Angelini LG, Lazzeri A, Levita G, Fontanelli D, Bozzi C: **Ramie (*Boehmeria nivea* (L.) Gaud.) and Spanish Broom (*Spartium junceum* L.) fibres for composite materials: agronomical aspects, morphology and mechanical properties.** *Industrial Crops and Products* 2000, **11**(2–3):145–161.
3. Day A, Ruel K, Neutelings G, Cronier D, David H, Hawkins S, Chabbert B: **Lignification in the flax stem: evidence for an unusual lignin in bast fibers.** *Planta* 2005, **222**(2):234–245.
4. Zhang SY, Jia TT, Zhang Z, Zou XY, Fan SM, Lei K, Jiang X, Niu DD, Yuan YL, Shang HH: **Insight into the relationship between S-lignin and fiber quality based on multiple research methods.** *Plant Physiology and Biochemistry* 2020, **147**:251–261.
5. Liu CJ: **Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly.** *Molecular Plant* 2012, **5**(2):304–317.
6. Fan L, Shi WJ, Hu WR, Hao XY, Wang DM, Yuan H, Yan HY: **Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fibers.** *Journal of Integrative Plant Biology* 2009, **51**(7):626–637.
7. Gao Z, Sun W, Wang J, Zhao C, Zuo K: **GhbHLH18 negatively regulates fiber strength and length by enhancing lignin biosynthesis in cotton fibers.** *Plant Science* 2019, **286**:7–16.
8. Boerjan W, Ralph J, Baucher M: **Lignin biosynthesis.** *Annual Review of Plant Biology* 2003, **54**(1):519–546.
9. Zhong RM, W Herbert Negrel, Jonathan Ye, Zheng-Hua: **Dual methylation pathways in lignin biosynthesis.** *The Plant Cell* 1998, **10**(12):2033–2045.
10. Ye ZH, Varner JE: **Differential expression of two O-methyltransferases in lignin biosynthesis in *Zinnia elegans*.** *Plant Physiology* 1995, **108**(2):459–467.
11. Bugos RC, Chiang VL, Campbell WH: **cDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase of aspen.** *Plant Molecular Biology* 1991, **17**(6):1203–1215.
12. Vincent D, Lapiere C, Pollet B, Cornic G, Negrone L, Zivy M: **Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation.** *Plant Physiology* 2005, **137**(3):949–960.
13. Zubieta C, Kota P, Ferrer JL, Dixon RA, Noel JP: **Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase.** *The Plant Cell* 2002, **14**(6):1265–1277.
14. Li JJ, Zhang G, Yu JH, Li YY, Huang XH, Wang WJ, Tan R, Zhou JY, Liao H: **Molecular cloning and characterization of caffeic acid 3-O-methyltransferase from the rhizome of *Ligusticum chuanxiong*.** *Biotechnology Letters* 2015, **37**(11):2295–2302.

15. Ma J, Liu F, Wang Q, Wang K, Jones DC, Zhang B: **Comprehensive analysis of TCP transcription factors and their expression during cotton (*Gossypium arboreum*) fiber early development.** *Scientific Reports* 2016, **6**:1–10.
16. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: **TBtools—an integrative toolkit developed for interactive analyses of big biological data.** *Molecular Plant* 2020, **13**(8):1194–1202.
17. Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH: **Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants.** *Genome Biology* 2019, **20**(1):38.
18. Hurst LD: **The Ka/Ks ratio: diagnosing the form of sequence evolution.** *TRENDS in Genetics* 2002, **9**(18):486–487.
19. Ma W, Zhao T, Li J, Liu B, Fang L, Hu Y, Zhang T: **Identification and characterization of the GhHsp20 gene family in *Gossypium hirsutum*.** *Scientific reports* 2016, **6**:1–13.
20. Kim J, Choi B, Cho B-K, Lim H-S, Kim JB, Natarajan S, Kwak E, Bae H: **Molecular cloning, characterization and expression of the caffeic acid O-methyltransferase (COMT) ortholog from kenaf (*Hibiscus cannabinus*).** *Plant Omics* 2013, **6**(4):246–253.
21. Binder JX, Pletscher-Frankild S, Tsafou K, Stolte C, O'Donoghue SI, Schneider R, Jensen LJ: **COMPARTMENTS: unification and visualization of protein subcellular localization evidence.** *Database* 2014, **2014**:12.
22. Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Lührmann R: **Conservation of the protein composition and electron microscopy structure of *Drosophila melanogaster* and human spliceosomal complexes.** *Molecular and cellular biology* 2009, **29**(1):281–301.
23. Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M, Lamond AI: **Directed proteomic analysis of the human nucleolus.** *Current Biology* 2002, **12**(1):1–11.
24. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier C, Nakai K: **WoLF PSORT: protein localization predictor.** *Nucleic acids research* 2007, **35**(suppl_2):W585–W587.
25. Briesemeister S, Rahnenführer J, Kohlbacher O: **Going from where to why—interpretable prediction of protein subcellular localization.** *Bioinformatics* 2010, **26**(9):1232–1238.
26. Pierleoni A, Martelli PL, Fariselli P, Casadio R: **BaCellLo: a balanced subcellular localization predictor.** *Bioinformatics* 2006, **22**(14):e408–e416.
27. Goldberg T, Hamp T, Rost B: **LocTree2 predicts localization for all domains of life.** *Bioinformatics* 2012, **28**(18):i458–i465.
28. Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T, Lian J: ***Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton.** *Nature Genetics* 2019, **51**(4):739–748.
29. Li P-t, Wang M, Lu Q-w, Ge Q, Liu A-y, Gong J-w, Shang H-h, Gong W-k, Li J-w, Song W-w: **Comparative transcriptome analysis of cotton fiber development of Upland cotton (*Gossypium hirsutum*) and Chromosome Segment Substitution Lines from *G. hirsutum* × *G. barbadense*.** *BMC genomics* 2017, **18**(1):705.
30. Zhang Z, Li J, Jamshed M, Shi Y, Liu A, Gong J, Wang S, Zhang J, Sun F, Jia F *et al*: **Genome-wide quantitative trait loci reveal the genetic basis of cotton fibre quality and yield-related traits in a *Gossypium hirsutum* recombinant inbred line population.** *Plant biotechnology journal* 2020, **18**(1):239–253.
31. Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M *et al*: **Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits.** *Nature Genetics* 2018, **50**(6):796–802.
32. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J *et al*: **Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres.** *Nature* 2012, **492**(7429):423–427.
33. Wang K, Wang Z, Li F, Ye W, Wang J, Song G, Yue Z, Cong L, Shang H, Zhu S: **The draft genome of a diploid cotton *Gossypium raimondii*.** *Nature Genetics* 2012, **44**(10):1098–1103.
34. Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT *et al*: **Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*.** *Nature* 2011, **477**(7365):419–423.
35. Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D: **Analysis of the bread wheat genome using whole-genome shotgun sequencing.** *Nature* 2012, **491**(7426):705–710.
36. Otto SP: **The evolutionary consequences of polyploidy.** *Cell* 2007, **131**(3):452–462.
37. Soltis PS, Soltis DE: **The role of hybridization in plant speciation.** *Annual Review of Plant Biology* 2009, **60**:561–588.

38. Grover CE, Gallagher JP, Szadkowski EP, Yoo MJ, Fligel LE, Wendel JF: **Homoeolog expression bias and expression level dominance in allopolyploids.** *New Phytologist* 2012, **196**(4):966–971.
39. Leach LJ, Belfield EJ, Jiang C, Brown C, Mithani A, Harberd NP: **Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat.** *BMC Genomics* 2014, **15**(1):276.
40. Cheng F, Wu J, Fang L, Sun S, Liu B, Lin K, Bonnema G, Wang X: **Biased Gene Fractionation and Dominant Gene Expression among the Subgenomes of Brassica rapa.** *PLoS ONE* 2012, **7**(5):e36442.
41. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM *et al.*: **Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement.** *Nature Biotechnology* 2015, **33**(5): 531–537.
42. Cheng F, Wu J, Cai X, Liang J, Freeling M, Wang X: **Gene retention, fractionation and subgenome differences in polyploid plants.** *Nature plants* 2018, **4**(5):258-268.
43. Renny-Byfield S, Gallagher JP, Grover CE, Szadkowski E, Page JT, Udall JA, Wang X, Paterson AH, Wendel JF: **Ancient gene duplicates in *Gossypium* (cotton) exhibit near-complete expression divergence.** *Genome Biology and Evolution* 2014, **6**(3):559–571.
44. Lu N, Ma W, Han D, Liu Y, Wang Z, Wang N, Yang G, Qu G, Wang Q, Zhao K: **Genome-wide analysis of the *Catalpa bungei* caffeic acid O-methyltransferase (COMT) gene family: identification and expression profiles in normal, tension, and opposite wood.** *PeerJ* 2019, **7**:e6520.
45. Van Zee JP, Schlueter JA, Schlueter S, Dixon P, Sierra CAB, Hill CA: **Paralog analyses reveal gene duplication events and genes under positive selection in *Ixodes scapularis* and other ixodid ticks.** *BMC Genomics* 2016, **17**(1):241.
46. Conant GC, Wolfe KH: **Turning a hobby into a job: how duplicated genes find new functions.** *Nature Reviews Genetics* 2008, **9**(12):938–950.
47. Ounaroon A, Decker G, Schmidt J, Lottspeich F, Kutchan TM: **(R, S)-Reticuline 7-O-methyltransferase and (R, S)-norcoclaurine 6-O-methyltransferase of *Papaver somniferum*—cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy.** *The plant journal* 2003, **36**(6):808-819.
48. Ge Q, Cui Y, Li J, Gong J, Lu Q, Li P, Shi Y, Shang H, Liu A, Deng X *et al.*: **Disequilibrium evolution of the Fructose-1, 6-bisphosphatase gene family leads to their functional biodiversity in *Gossypium* species.** *BMC Genomics* 2020.
49. Lamaoui M, Jemo M, Datla R, Bekkaoui F: **Heat and drought stresses in crops and approaches for their mitigation.** *Frontier in Chemistry* 2018, **6**:26.
50. Rao KM, Raghavendra A, Reddy KJ: **Physiology and molecular biology of stress tolerance in plants.** Dordrecht: Springer Science & Business Media; 2006.
51. Xu P, Liu Z, Fan X, Gao J, Zhang X, Zhang X, Shen X: **De novo transcriptome sequencing and comparative analysis of differentially expressed genes in *Gossypium aridum* under salt stress.** *Gene* 2013, **525**(1):26–34.
52. Sazegari S, Niazi A, Ahmadi FS: **A study on the regulatory network with promoter analysis for *Arabidopsis* DREB-genes.** *Bioinformatics* 2015, **11**(2):101–106.
53. Meier S, Bastian R, Donaldson L, Murray S, Bajic V, Gehring C: **Co-expression and promoter content analyses assign a role in biotic and abiotic stress responses to plant natriuretic peptides.** *BMC Plant Biology* 2008, **8**(1):24.
54. Maleck K, Levine A, Eulgem T, Morgan A, Schmidl J, Lawton K, Dangl J, Dietrich R: **An *Arabidopsis* promoter element shared among genes co-regulated during systemic acquired disease resistance.** *Nature Genetics* 2000, **26**:403–410.
55. Eulgem T, Rushton PJ, Robatzek S, Somssich IE: **The WRKY superfamily of plant transcription factors.** *Trends in Plant Science* 2000, **5**(5):199–206.
56. Huang T, Nicodemus J, Zarka DG, Thomashow MF, Wisniewski M, Duman JG: **Expression of an insect (*Dendroides canadensis*) antifreeze protein in *Arabidopsis thaliana* results in a decrease in plant freezing temperature.** *Plant Molecular Biology* 2002, **50**(3):333–344.
57. Hara K, Yagi M, Kusano T, Sano H: **Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding.** *Molecular and General Genetics* 2000, **263**(1):30–37.
58. Rizhsky L, Davletova S, Liang H, Mittler R: **The zinc finger protein *Zat12* is required for cytosolic ascorbate peroxidase 1 expression during oxidative stress in *Arabidopsis*.** *Journal of Biological Chemistry* 2004, **279**(12):11736–11743.

59. Pnueli L, Hallak-Herr E, Rozenberg M, Cohen M, Goloubinoff P, Kaplan A, Mittler R: **Molecular and biochemical mechanisms associated with dormancy and drought tolerance in the desert legume *Retama raetam***. *The Plant Journal* 2002, **31**(3):319–330.
60. Seki M, Ishida J, Narusaka M, Fujita M, Nanjo T, Umezawa T, Kamiya A, Nakajima M, Enju A, Sakurai T: **Monitoring the expression pattern of around 7,000 *Arabidopsis* genes under ABA treatments using a full-length cDNA microarray**. *Functional & Integrative Genomics* 2002, **2**(6):282–291.
61. Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R: **When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress**. *Plant Physiology* 2004, **134**(4):1683–1696.
62. Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K: **Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression**. *The Plant Cell* 1997, **9**(10):1859–1868.
63. Thomashow MF: **Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway**. *Plant Physiology* 2010, **154**(2):571–577.
64. Zhang Y, Sun T, Liu S, Dong L, Liu C, Song W, Liu J, Gai S: **MYC cis-Elements in PsMPT Promoter Is Involved in Chilling Response of *Paeonia suffruticosa***. *PLoS One* 2016, **11**(5):e0155780.
65. Zhao Y: **Auxin biosynthesis and its role in plant development**. *Annual Review of Plant Biology* 2010, **61**:49–64.
66. Ramawat KG, Mérillon J-M: **Natural products: phytochemistry, botany and metabolism of alkaloids, phenolics and terpenes**. Heidelberg, Germany: Springer; 2013.
67. Cheyner V, Comte G, Davies KM, Lattanzio V, Martens S: **Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology**. *Plant Physiology and Biochemistry* 2013, **72**:1–20.
68. Laura A, Alvarez-Parrilla E, Gonzalez-Aguilar GA: **Fruit and vegetable phytochemicals: Chemistry, nutritional value and stability**: John Wiley & Sons; 2009.
69. Naikoo MI, Dar MI, Raghieb F, Jaleel H, Ahmad B, Raina A, Khan FA, Naushin F: **Role and Regulation of Plants Phenolics in Abiotic Stress Tolerance: An Overview**. In: *Plant Signaling Molecules*. Elsevier; 2019: 157–168.
70. Dixon RAP, N. L.: **Stress-Induced Phenylpropanoid Metabolism**. *Plant Cell* 1995, **7**(7):1085–1097.
71. Handa N, Kohli SK, Sharma A, Thukral AK, Bhardwaj R, Abd_Allah EF, Alqarawi AA, Ahmad P: **Selenium modulates dynamics of antioxidative defence expression, photosynthetic attributes and secondary metabolites to mitigate chromium toxicity in *Brassica juncea* L. plants**. *Environmental and Experimental Botany* 2019, **161**:180–192.
72. Smirnov OE, Kosyan AM, Kosyk OI, Taran NY: **Response of Phenolic Metabolism Induced by Aluminium Toxicity in *Fagopyrum Esculentum* Moench. Plants**. *Ukrainian Biochemical Journal* 2015, **87**(6):129–135.
73. Weretilnyk EA, Alexander KJ, Drebenstedt M, Snider JD, Summers PS, Moffatt BA: **Maintaining methylation activities during salt stress. The involvement of adenosine kinase**. *Plant Physiology* 2001, **125**(2):856–865.
74. Ancillotti C, Bogani P, Biricolti S, Calistri E, Checchini L, Ciofi L, Gonnelli C, Del Bubba M: **Changes in polyphenol and sugar concentrations in wild type and genetically modified *Nicotiana langsdorffii* Weinmann in response to water and heat stress**. *Plant Physiology Biochemistry* 2015, **97**:52–61.
75. Sasayama D, Azuma T, Itoh K: **Involvement of cell wall-bound phenolic acids in decrease in cell wall susceptibility to expansins during the cessation of rapid growth in internodes of floating rice**. *Journal of Plant Physiology* 2011, **168**(2):121–127.
76. Fan L, Linker R, Gepstein S, Tanimoto E, Yamamoto R, Neumann PM: **Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics**. *Plant Physiology* 2006, **140**(2):603–612.
77. Wang J, Pichersky E: **Identification of specific residues involved in substrate discrimination in two plant O-methyltransferases**. *Archives of Biochemistry Biophysics* 1999, **368**(1):172–180.
78. Gou JY, Wang LJ, Chen SP, Hu WL, Chen XY: **Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis**. *Cell Res* 2007, **17**(5):422–434.
79. Al-Ghazi Y, Bourrot S, Arioli T, Dennis ES, Llewellyn DJ: **Transcript profiling during fiber development identifies pathways in secondary metabolism and cell wall structure that may contribute to cotton fiber quality**. *Plant Cell Physiology* 2009, **50**(7):1364–1381.

80. Grimmig B, Kneusel R, Junghanns K, Matern U: **Expression of Bifunctional Caffeoyl-CoA 3-O-Methyltransferase in Stress Compensation and Lignification.** *Plant Biology* 1999, **1**(3):299–310.
81. Ni W, Fahrendorf T, Ballance GM, Lamb CJ, Dixon RA: **Stress responses in alfalfa (*Medicago sativa* L.). XX. Transcriptional activation of phenylpropanoid pathway genes in elicitor-induced cell suspension cultures.** *Plant Molecular Biology* 1996, **30**(3):427–438.
82. Cui Y, BELL AA, Joost O, Magill C: **Expression of potential defense response genes in cotton.** *Physiological Molecular Plant Pathology* 2000, **56**(1):25–31.
83. Li C, He Q, Zhang F, Yu J, Li C, Zhao T, Zhang Y, Xie Q, Su B, Mei L: **Melatonin enhances cotton immunity to *Verticillium* wilt via manipulating lignin and gossypol biosynthesis.** *The Plant Journal* 2019, **100**(4):784–800.
84. Liu J, Benedict CR, Stipanovic RD, Bell AA: **Purification and characterization of S-Adenosyl-L-Methionine: desoxyhemigossypol-6-O-methyltransferase from cotton plants. An enzyme capable of methylating the defense terpenoids of cotton.** *Plant Physiology* 1999, **121**(3):1017–1024.
85. Wang W, Li R, Zhu Q, Tang X, Zhao Q: **Transcriptomic and physiological analysis of common duckweed *Lemna minor* responses to NH₄⁺ toxicity.** *BMC Plant Biology* 2016, **16**(1):92.
86. Li P-t, Chen T-t, Lu Q-w, Ge Q, Gong W-k, Liu A-y, Gong J-w, Shang H-h, Deng X-y, Li J-w: **Transcriptomic and biochemical analysis of upland cotton (*Gossypium hirsutum*) and a chromosome segment substitution line from *G. hirsutum* × *G. barbadense* in response to *Verticillium dahliae* infection.** *BMC Plant Biology* 2019, **19**(1):1–24.
87. Tang Y, Zhang ZN, Lei Y, Hu G, Liu JF, Hao MY, Chen AM, Peng QZ, Wu JH: **Cotton WATs Modulate SA Biosynthesis and Local Lignin Deposition Participating in Plant Resistance Against *Verticillium dahliae*.** *Frontiers in Plant Science* 2019, **10**:526.
88. Xu L, Zhu L, Tu L, Liu L, Yuan D, Jin L, Long L, Zhang X: **Lignin metabolism has a central role in the resistance of cotton to the wilt fungus *Verticillium dahliae* as revealed by RNA-Seq-dependent transcriptional analysis and histochemistry.** *Journal of Experimental Botany* 2011, **62**(15):5607–5621.
89. Tuteja N: **Mechanisms of high salinity tolerance in plants.** *Methods in Enzymology* 2007, **428**:419–438.
90. Joshi A, Dang HQ, Vaid N, Tuteja N: **Isolation of high salinity stress tolerant genes from *Pisum sativum* by random overexpression in *Escherichia coli* and their functional validation.** *Plant Signaling Behavior* 2009, **4**(5):400–412.
91. Liu DD, Sun XS, Liu L, Shi HD, Chen SY, Zhao DK: **Overexpression of the Melatonin Synthesis-Related Gene SICOMT1 Improves the Resistance of Tomato to Salt Stress.** *Molecules* 2019, **24**(8):1514.
92. Liu DG, He SZ, Song XJ, Zhai H, Liu N, Zhang DD, Ren ZT, Liu QC: **lbSIMT1, a novel salt-induced methyltransferase gene from *Ipomoea batatas*, is involved in salt tolerance.** *Plant Cell Tissue and Organ Culture* 2015, **120**(2):701–715.
93. Wang M, Zhu X, Wang K, Lu C, Luo M, Shan T, Zhang Z: **A wheat caffeic acid 3-O-methyltransferase TaCOMT-3D positively contributes to both resistance to sharp eyespot disease and stem mechanical strength.** *Scientific Reports* 2018, **8**(1):1–14.
94. Ma QH: **The expression of caffeic acid 3-O-methyltransferase in two wheat genotypes differing in lodging resistance.** *Journal of Experimental Botany* 2009, **60**(9):2763–2771.
95. Wu X, Wu J, Luo Y, Bragg J, Anderson O, Vogel J, Gu YQ: **Phylogenetic, Molecular, and Biochemical Characterization of Caffeic Acid o-Methyltransferase Gene Family in *Brachypodium distachyon*.** *International Journal of Plant Genomics* 2013, **2013**(5-6):423189.
96. Hsu C-Y, II MAA, Miao Q, Saha S, Jenkins JN, Ayubov MS, Abdurakhmonov IY, Peterson DG, Ma D-P: **Transcriptome Analysis of Ten Days Post Anthesis Elongating Fiber in the Upland Cotton (*Gossypium hirsutum*) Chromosome Substitution Line CS-B25.** *American Journal of Plant Sciences* 2018, **9**:1334–1361.
97. Bhardwaj R, Handa N, Sharma R, Kaur H, Kohli S, Kumar V, Kaur P: **Lignins and abiotic stress: An overview.** In: Ahmad P, Wani M (eds) *Physiological Mechanisms and Adaptation Strategies in Plants Under Changing Environment.* Springer, New York, NY. 2014: 267–296.
98. Han LB, Li YB, Wang HY, Wu XM, Li CL, Luo M, Wu SJ, Kong ZS, Pei Y, Jiao GL *et al.*: **The dual functions of WLIM1a in cell elongation and secondary wall formation in developing cotton fibers.** *Plant Cell* 2013, **25**(11):4421–4438.
99. Zhu T, Liang C, Meng Z, Sun G, Meng Z, Guo S, Zhang R: **CottonFGD: an integrated functional genomics database for cotton.** *BMC plant biology* 2017, **17**(1):1–9.

100. Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C: **Whole-genome sequencing of multiple *Arabidopsis thaliana* populations.** *Nature genetics* 2011, **43**(10):956.
101. Argout X, Salse J, Aury J-M, Guiltinan MJ, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova SN: **The genome of *Theobroma cacao*.** *Nature genetics* 2011, **43**(2):101–108.
102. Finn RD, Clements J, Eddy SR: **HMMER web server: interactive sequence similarity searching.** *Nucleic acids research* 2011, **39**(suppl_2):W29–W37.
103. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R: **Clustal W and Clustal X version 2.0.** *bioinformatics* 2007, **23**(21):2947–2948.
104. Kumar S, Stecher G, Tamura K, evolution: **MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.** *Molecular biology* 2016, **33**(7):1870–1874.
105. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G: **GSDS 2.0: an upgraded gene feature visualization server.** *Bioinformatics* 2015, **31**(8):1296–1297.
106. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching.** *Nucleic acids research* 2009, **37**(suppl_2):W202–W208.
107. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.** *Nucleic acids research* 2002, **30**(1):325–327.
108. Yu CS, Chen YC, Lu CH, Hwang JK: **Prediction of protein subcellular localization.** *Proteins: Structure, Function, Bioinformatics* 2006, **64**(3):643–651.
109. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic acids research* 2000, **28**(1):27–30.
110. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, Hub A, Group WPW: **AmiGO: online access to ontology and annotation data.** *Bioinformatics* 2009, **25**(2):288–289.
111. Lu Q, Shi Y, Xiao X, Li P, Gong J, Gong W, Liu A, Shang H, Li J, Ge Q *et al.*: **Transcriptome Analysis Suggests That Chromosome Introgression Fragments from Sea Island Cotton (*Gossypium barbadense*) Increase Fiber Strength in Upland Cotton (*Gossypium hirsutum*).** *G3 (Bethesda)* 2017, **7**(10):3469–3479.
112. Shi Y, Li W, Li A, Ge R, Zhang B, Li J, Liu G, Li J, Liu A, Shang H: **Constructing a high-density linkage map for *Gossypium hirsutum* × *Gossypium barbadense* and identifying QTLs for lint percentage.** *Journal of Integrative Plant Biology* 2015, **57**(5):450–467.
113. Jamshed M, Jia F, Gong J, Palanga KK, Shi Y, Li J, Shang H, Liu A, Chen T, Zhang Z *et al.*: **Identification of stable quantitative trait loci (QTLs) for fiber quality traits across multiple environments in *Gossypium hirsutum* recombinant inbred line population.** *BMC Genomics* 2016, **17**:197.
114. Sun F-D, Zhang J-H, Wang S-F, Gong W-K, Shi Y-Z, Liu A-Y, Li J-W, Gong J-W, Shang H-H, Yuan Y-L: **QTL mapping for fiber quality traits across multiple generations and environments in upland cotton.** *Molecular Breeding* 2012, **30**(1):569–582.
115. Zhang Z, Li J, Jamshed M, Shi Y, Liu A, Gong J, Wang S, Zhang J, Sun F, Jia F *et al.*: **Genome-wide quantitative trait loci reveal the genetic basis of cotton fibre quality and yield-related traits in a *Gossypium hirsutum* recombinant inbred line population.** *Plant Biotechnol J* 2020, **18**(1):239–253.
116. Rychlik W: **OLIGO 7 primer analysis software.** In: *PCR primer design*. Springer; 2007: 35–59.
117. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method.** *Methods* 2001, **25**(4):402–408.

Table

Table 1. Predicted Subcellular localization of *OMT* genes of *G. hirsutum*

Gene ID	CELLO	C_Reliability	Wolf Psort	P_Reliability	Gene ID	CELLO	C_Reliability	Wolf Psort	P_Reliability
1OMT51_At	Cp	3.712	Cp	1	GhOMT10_Dt	Cp	3.384	Cp	6
1OMT33_At	Cp	3.597	Cp	9.5	GhOMT7_Dt	Cp	4.717	Cp	11
1OMT74_At	Cp	2.985	Cp	7	GhOMT8_Dt	Cp	4.717	Cp	7
1OMT40_At	Cp	3.897	Cp	12	GhOMT9_Dt	Cp	4.735	Cp	9
1OMT68_At	Pp/Cp	2.094/2.624	Cp	6	GhOMT79_Dt	Cp	3.973	Cp	10
1OMT71_At	Cp	4.545	Cp	5	GhOMT81_Dt	Cp	3.924	Cp	6.5
1OMT78_At	Cp	4.122	Cp	13.5	GhOMT78_Dt	Cp	2.531	Cp	2
1OMT81_At	Cp	2.707	Cp	5	GhOMT54_Dt	Pp/Cp	1.967/2.566	Cp	8
hOMT3_At	Cp	4.033	Cp	2	GhOMT76_Dt	Cp	3.273	Mc	8
1OMT10_At	Cp	4.822	Cp	6	GhOMT32_Dt	Cp	4.607	Cp	12
1OMT12_At	Cp	4.154	Cp	4	GhOMT30_Dt	Cp	3.946	Cp	1
hOMT6_At	Cp	4.234	Cp	8	GhOMT31_Dt	Cp	4.695	Cp	6
1OMT76_At	Cp	4.275	Cp	11.5	GhOMT49_Dt	Pp	3.275	Cp	8
1OMT32_At	Cp	4.675	Cp	8	GhOMT57_Dt	Cp	4.391	Cp	7
1OMT30_At	Cp	4.743	Cp	6	GhOMT77_Dt	Cp	2.352	Cp	3
1OMT49_At	Pp	3.107	Pp	11	GhOMT58_Dt	Cp	4.675	Cp	5
1OMT77_At	Cp	3.838	Cp	7	GhOMT63_Dt	Cp	3.572	Cp	10
1OMT57_At	Cp	2.229	Cp	9	GhOMT62_Dt	Cp	2.968	Cp	4
1OMT52_At	Cp	1.577	Cp	10	GhOMT61_Dt	Cp	3.224	Cp	4
1OMT53_At	Pp/Cp	1.995/1.743	Cp	6.5	GhOMT17_Dt	Cp	4.112	Cp	2
1OMT62_At	Cp	3.577	Cp	2	GhOMT13_Dt	Cp	4.124	Cp	7
1OMT65_At	Cp	3.508	Cp	8	GhOMT70_Dt	Cp	4.385	Cp	13
1OMT24_At	Cp	4.529	Cp	8	GhOMT1_Dt	Cp	4.927	Cp	10
1OMT14_At	Cp	4.404	Cp	12	GhOMT55_Dt	Cp	2.781	Cp	1
1OMT29_At	Cp	4.069	Cp	1	GhOMT82_Dt	OM	2.226	Chp	9
1OMT41_At	Cp	4.655	Cp	6	GhOMT45_Dt	Pp	3.910	Cp	6
1OMT70_At	Cp	3.408	Cp	8	GhOMT46_Dt	Pp	2.490	Cp	10
hOMT1_At	Cp	4.921	Cp	7	GhOMT47_Dt	Pp/Cp	1.961/2.316	Cp	10
1OMT55_At	IM/Cp	1.807/2.478	Cp	1	GhOMT48_Dt	Pp	4.276	Cp	8
1OMT82_At	OM	2.374	Chp	9.5	GhOMT72_Dt	Cp	4.474	Cp	11.5
1OMT45_At	Pp	4.340	Cp	7	GhOMT60_Dt	Cp	4.312	Cp	7
1OMT47_At	Pp/Cp	2.121/2.358	Cp	12	GhOMT59_Dt	Cp	4.534	Cp	6
1OMT48_At	Pp	4.154	Chp	6	GhOMT5_Scaf	Cp	4.641	Cp	11
hOMT2_At	Cp	4.870	Cp	5	GhOMT11_Scaf	Cp	4.885	Cp	4
1OMT72_At	Cp	4.493	Cp	13.5	GhOMT68_Scaf	Cp	3.973	Cp	8
1OMT60_At	Cp	4.637	Cp	5	GhOMT52_Scaf	Cp	2.277	Cp	4
1OMT51_Dt	Cp	3.636	Cp	2	GhOMT46_Scaf	Cp	2.154	Cp	11
1OMT37_Dt	Cp	4.092	Cp	6	GhOMT15_Scaf	Cp	3.576	Cp	9
1OMT35_Dt	Cp	4.228	Cp	4	GhOMT75_Scaf	Cp	2.090	Cp	5
1OMT33_Dt	Cp	3.130	Cp	8	GhOMT73_Scaf	Cp	3.123	Cp	6
1OMT71_Dt	Cp	4.475	Cp	11.5					
hOMT6_Dt	Cp	4.619	Cp	8					

Cp: Cytoplasmic, Pp: Periplasmic, OM: outer membrane, IM: inner membrane, Chp: Chloroplast, Mc: Mitochondria
C_Reliability: Lower reliability values_show the stronger possibility of predicted localization
P_Reliability: Higher reliability values_show the stronger possibility of predicted localization

Figures

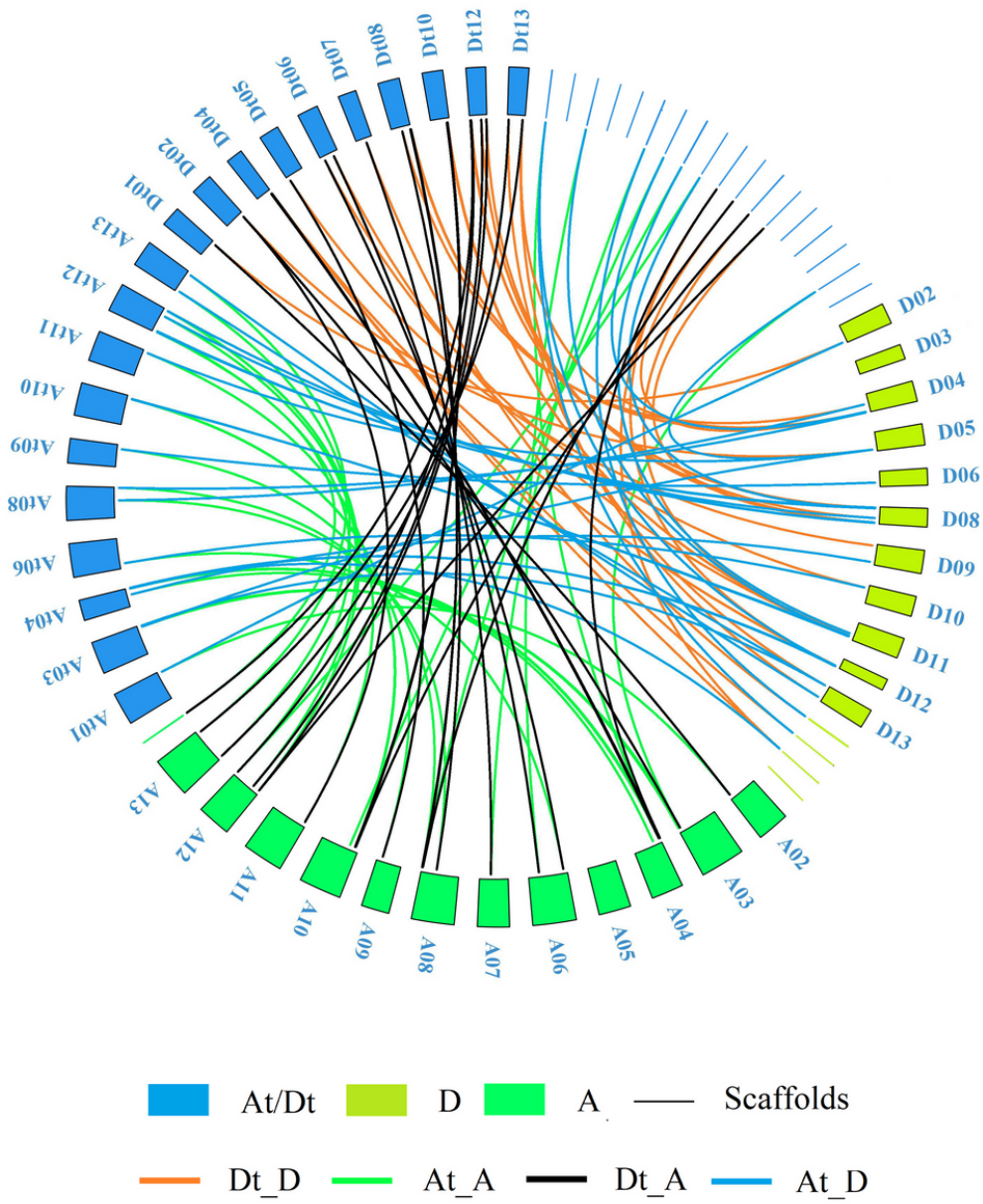


Figure 1

Collinearity analysis of OMT genes between AtDt (*G. hirsutum*), A (*G. arboreum*), and D (*G. raimondii*) genomes.

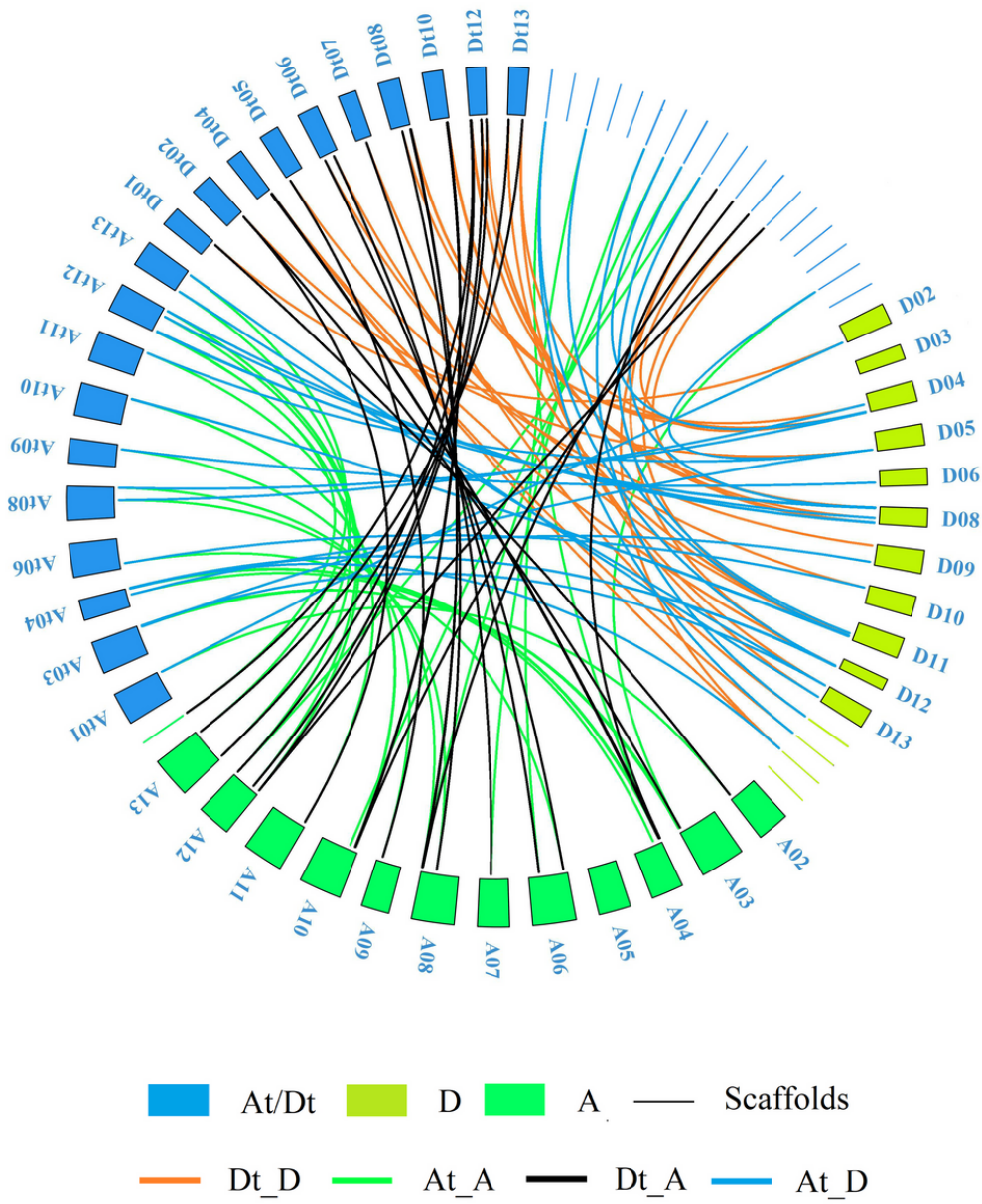


Figure 1

Collinearity analysis of OMT genes between AtDt (*G. hirsutum*), A (*G. arboreum*), and D (*G. raimondii*) genomes.

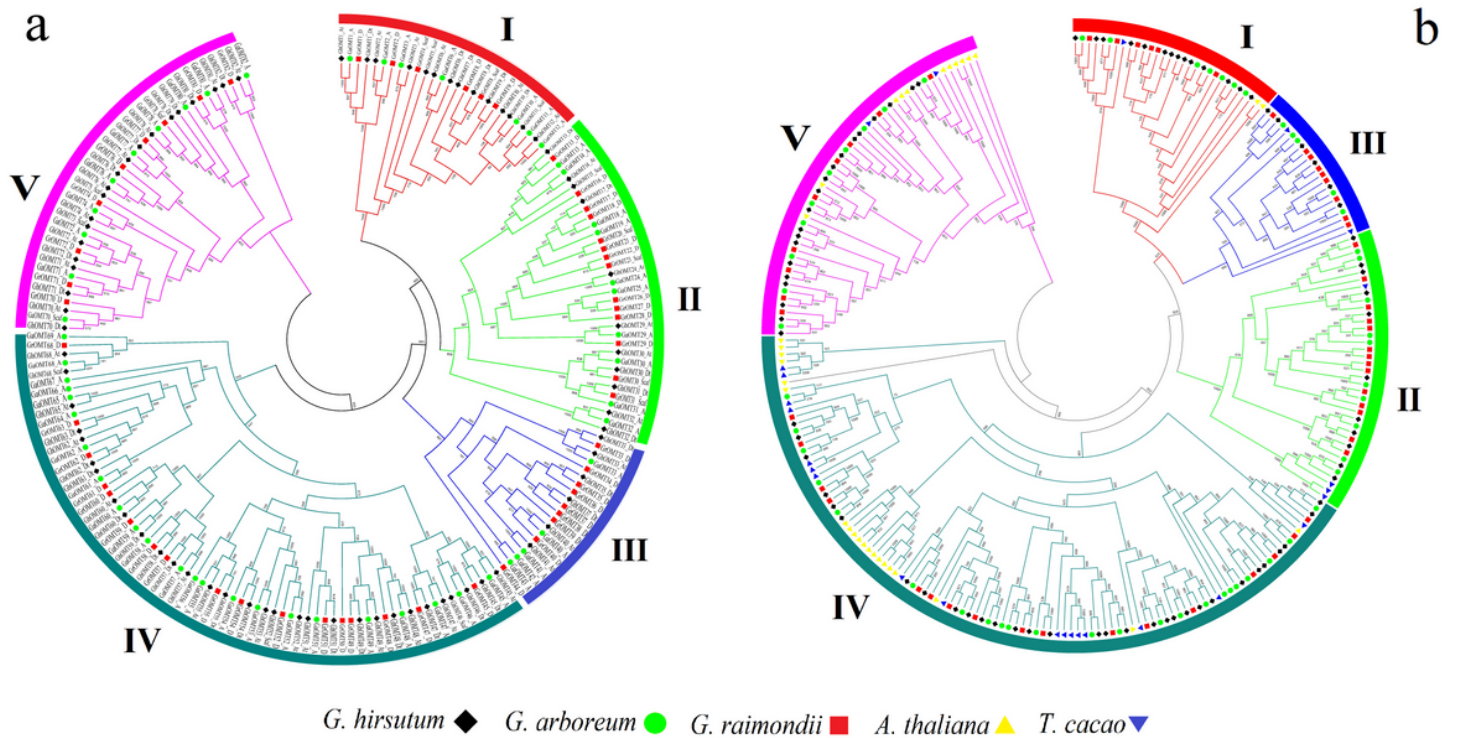


Figure 2

Phylogenetic analysis of OMT genes in *Gossypium*, *A. thaliana*, and *T. cacao* species. (a) Neighbor-joining phylogenetic tree of 192 OMT genes of *G. hirsutum*, *G. arboreum*, and *G. raimondii*. (b) Neighbor-joining phylogenetic tree of 251 OMT sequences of *G. hirsutum*, *G. arboreum*, *G. raimondii*, *A. thaliana*, and *T. cacao*. I, II, III, IV and V indicate the five groups of phylogenetic tree. Shapes with different colors represent OMT members of *G. hirsutum*, *G. arboreum*, *G. raimondii*, *A. thaliana*, and *T. cacao*.

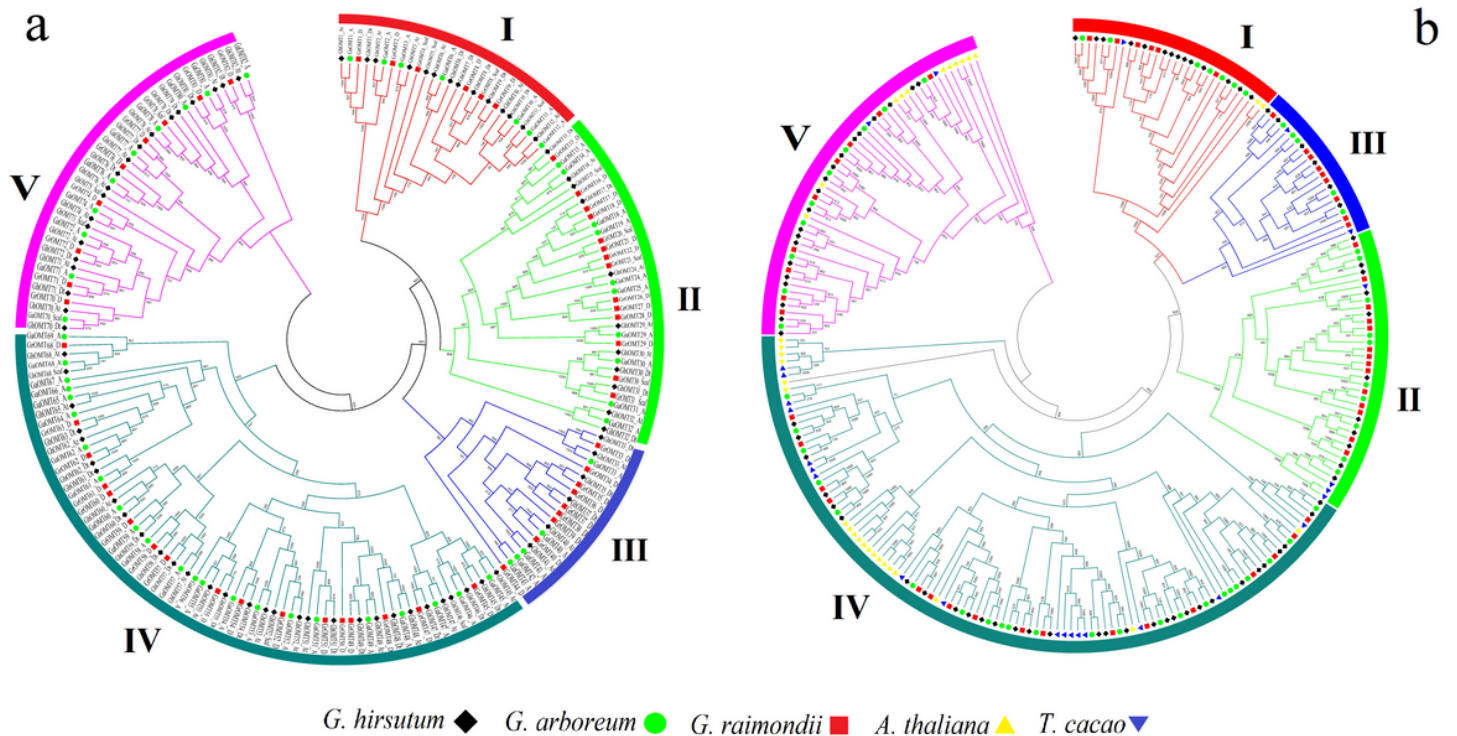


Figure 2

Phylogenetic analysis of OMT genes in *Gossypium*, *A. thaliana*, and *T. cacao* species. (a) Neighbor-joining phylogenetic tree of 192 OMT genes of *G. hirsutum*, *G. arboreum*, and *G. raimondii*. (b) Neighbor-joining phylogenetic tree of 251 OMT sequences of *G. hirsutum*, *G. arboreum*, *G. raimondii*, *A. thaliana*, and *T. cacao*. I, II, III, IV and V indicate the five groups of phylogenetic tree. Shapes with different colors represent OMT members of *G. hirsutum*, *G. arboreum*, *G. raimondii*, *A. thaliana*, and *T. cacao*.

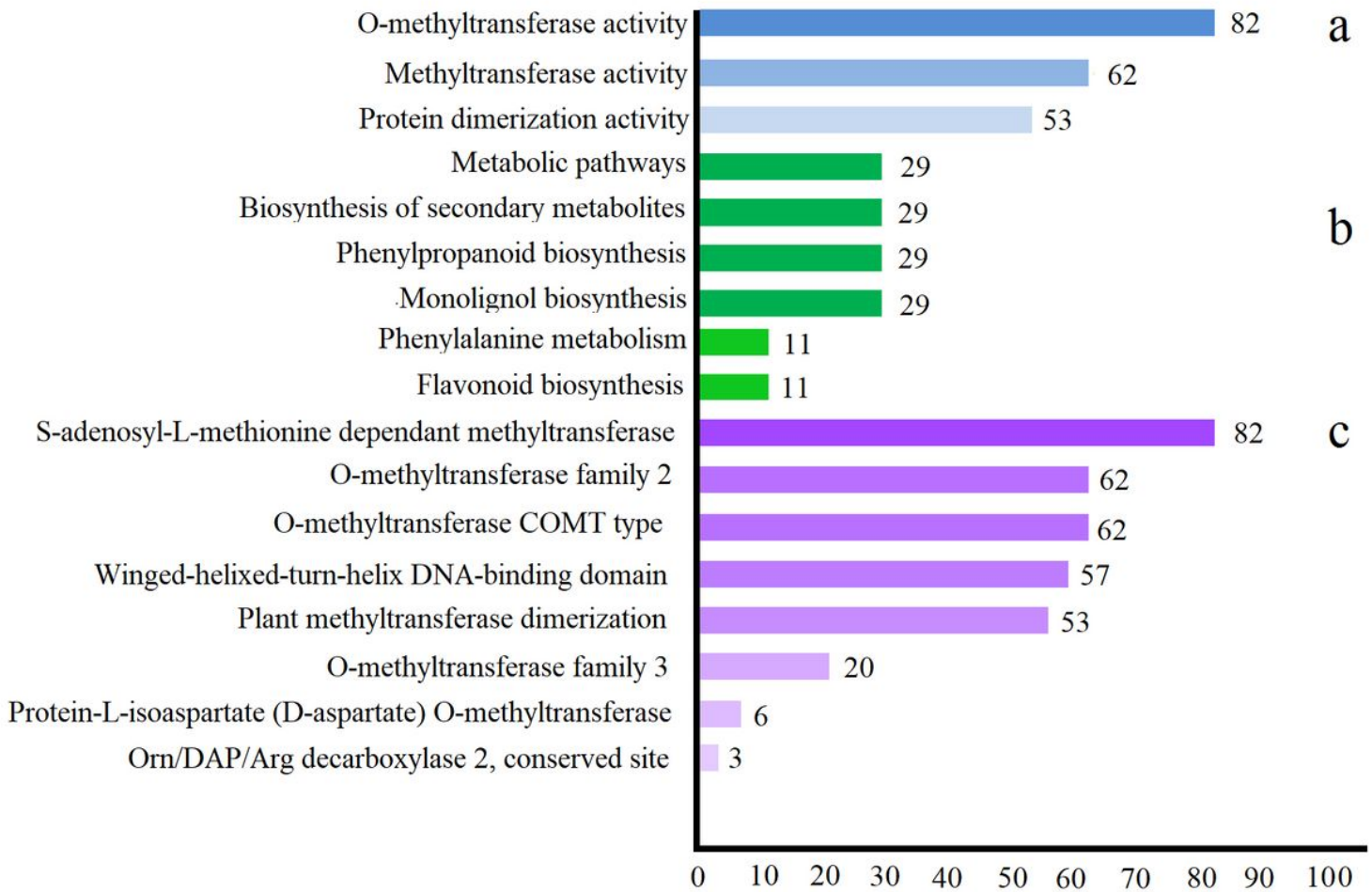


Figure 3

Enrichment analysis of OMT family genes in *G. hirsutum*. a. The functional annotations of Gene Ontology. b. The functional annotations of KEGG pathways. c. The functional annotations of Interpro. The scales indicate the enriched gene number in respective categories.

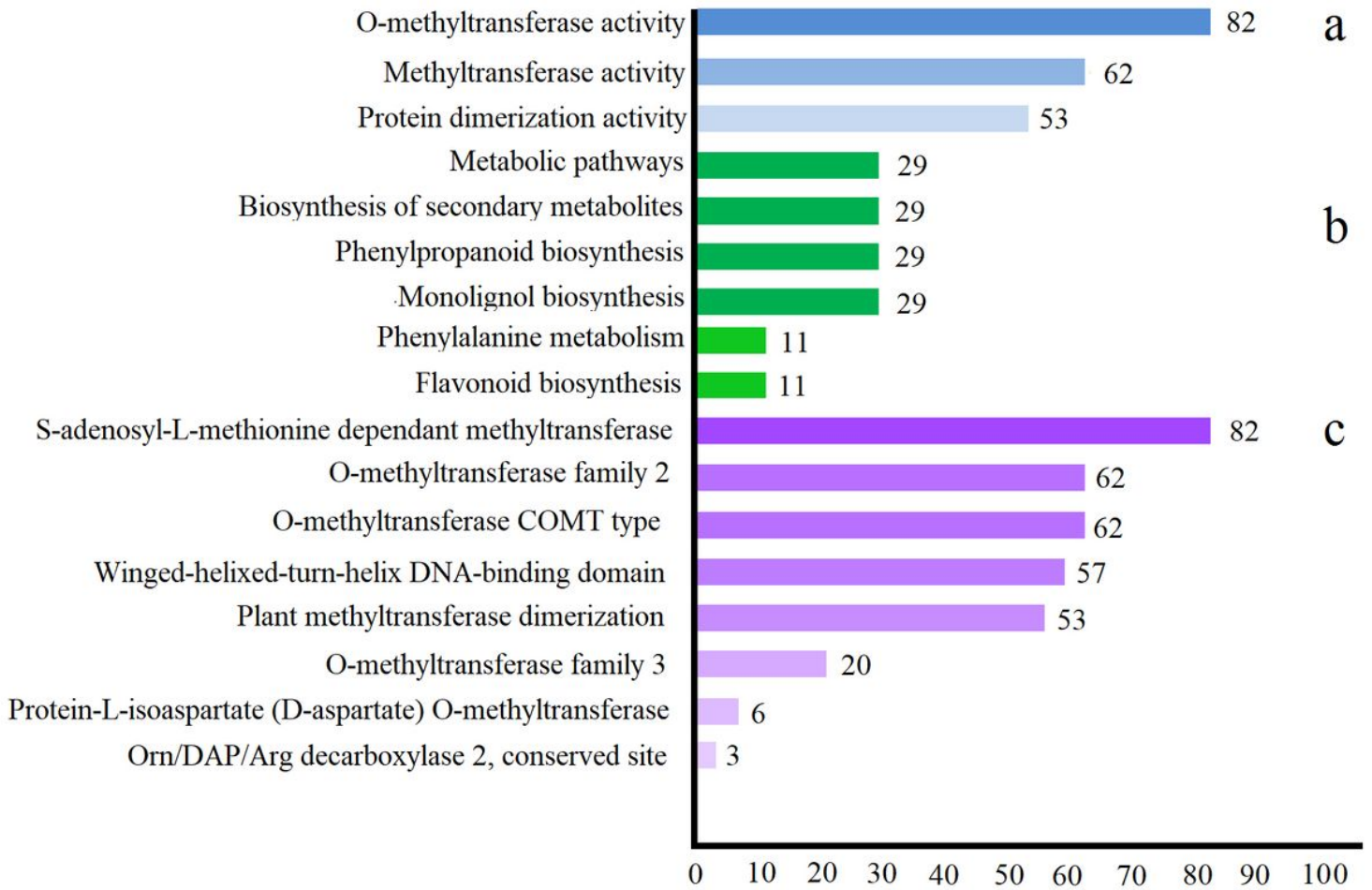


Figure 3

Enrichment analysis of OMT family genes in *G. hirsutum*. a. The functional annotations of Gene Ontology. b. The functional annotations of KEGG pathways. c. The functional annotations of Interpro. The scales indicate the enriched gene number in respective categories.

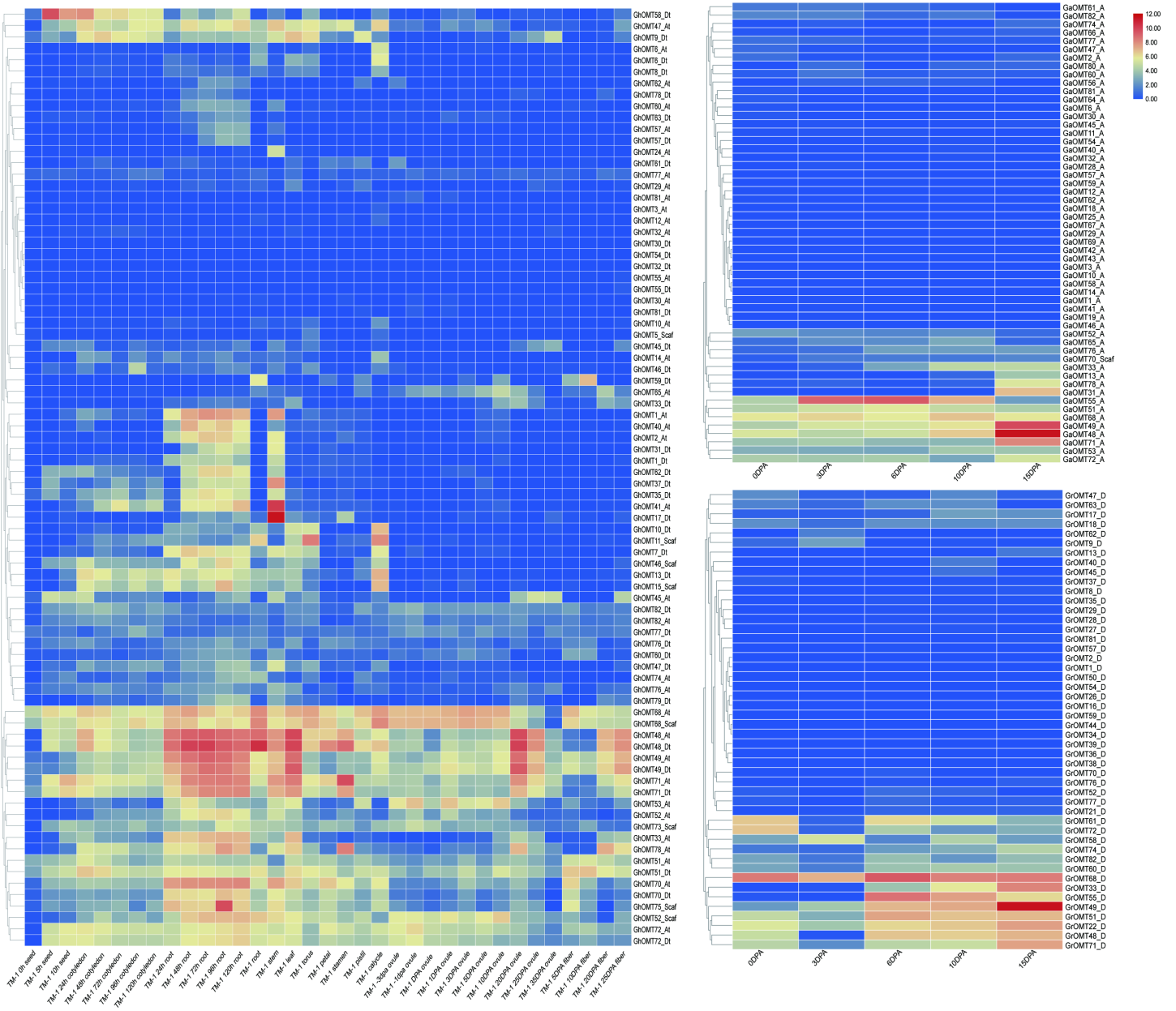


Figure 4

Transcriptome analysis of OMT family genes in different growth, ovule and fiber development stages. A: Transcriptome heatmap of GhOMT family genes in standard genetic cultivar TM-1 of *G. hirsutum* at different growth, ovule and fiber development stages [28, 41]. B: Transcriptome heatmap of GaOMT family genes in *G. arboreum* [41]. C: Transcriptome heatmap of GrOMT family genes in *G. raimondii* [33].

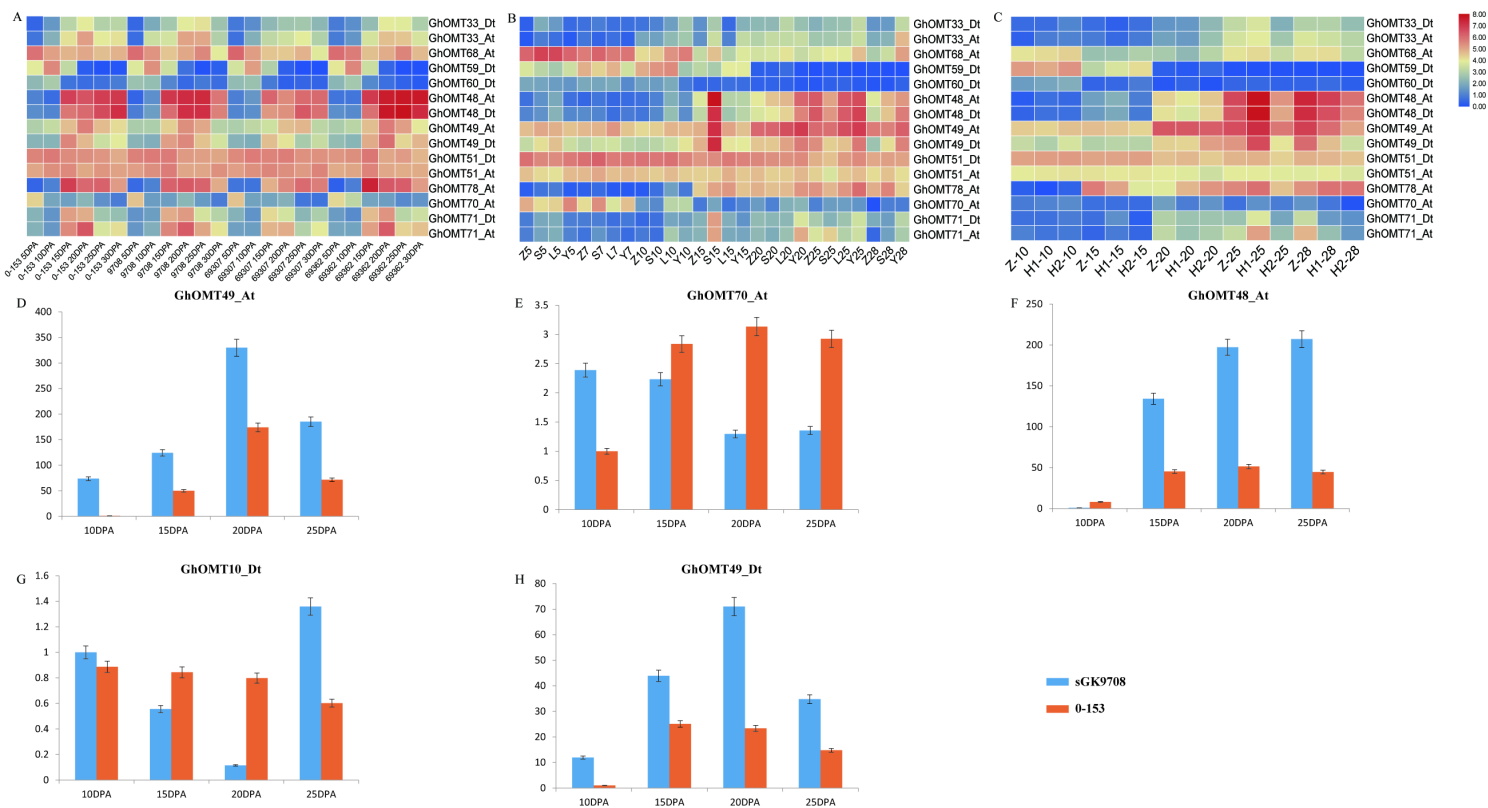


Figure 5

Specific responses of selected OMT genes in fiber development. A: Transcriptome heatmap of selected GhOMT genes in RIL lines and their parents [30]. B: Transcriptome heatmap of selected GhOMT genes in CSSLs of CCRI45 and Hai1 [111, 112]; Z, S, L, and Y represent CCRI45, MBI7561, MBI7747, and MBI7285, respectively; 5, 7, 10, 15, 20, 25, and 28 represent different DPA. C: Transcriptome heatmap of selected GhOMT genes in CSSLs of CCRI36 and Hai1 [29, 112]; Z, H1, and H2 represent CCRI36, MBI9915, and MBI9749; 10, 15, 20, 25, and 28 represent different DPA. D-H: qRT-PCR verification results of GhOMT49_At, GhOMT70_At, GhOMT48_At, GhOMT10_Dt, GhOMT49_Dt in developing fibers of sGK9708 and 0-153 at 10, 15, 20, and 25 DPAs.

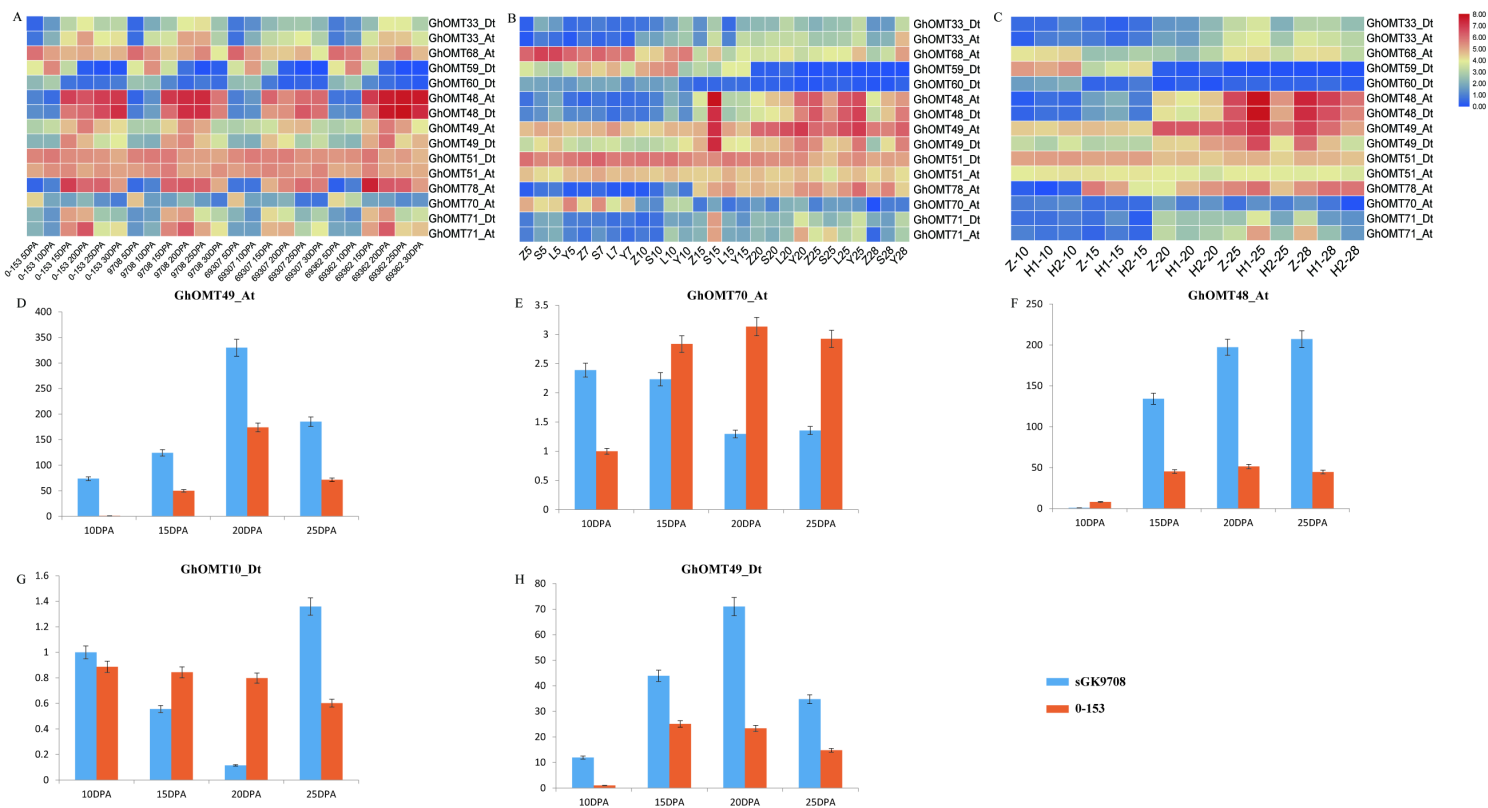


Figure 5

Specific responses of selected OMT genes in fiber development. A: Transcriptome heatmap of selected GhOMT genes in RIL lines and their parents [30]. B: Transcriptome heatmap of selected GhOMT genes in CSSLs of CCRI45 and Hai1 [111, 112]; Z, S, L, and Y represent CCRI45, MBI7561, MBI7747, and MBI7285, respectively; 5, 7, 10, 15, 20, 25, and 28 represent different DPA. C: Transcriptome heatmap of selected GhOMT genes in CSSLs of CCRI36 and Hai1 [29, 112]; Z, H1, and H2 represent CCRI36, MBI9915, and MBI9749; 10, 15, 20, 25, and 28 represent different DPA. D-H: qRT-PCR verification results of GhOMT49_At, GhOMT70_At, GhOMT48_At, GhOMT10_Dt, GhOMT49_Dt in developing fibers of sGK9708 and 0-153 at 10, 15, 20, and 25 DPAs.

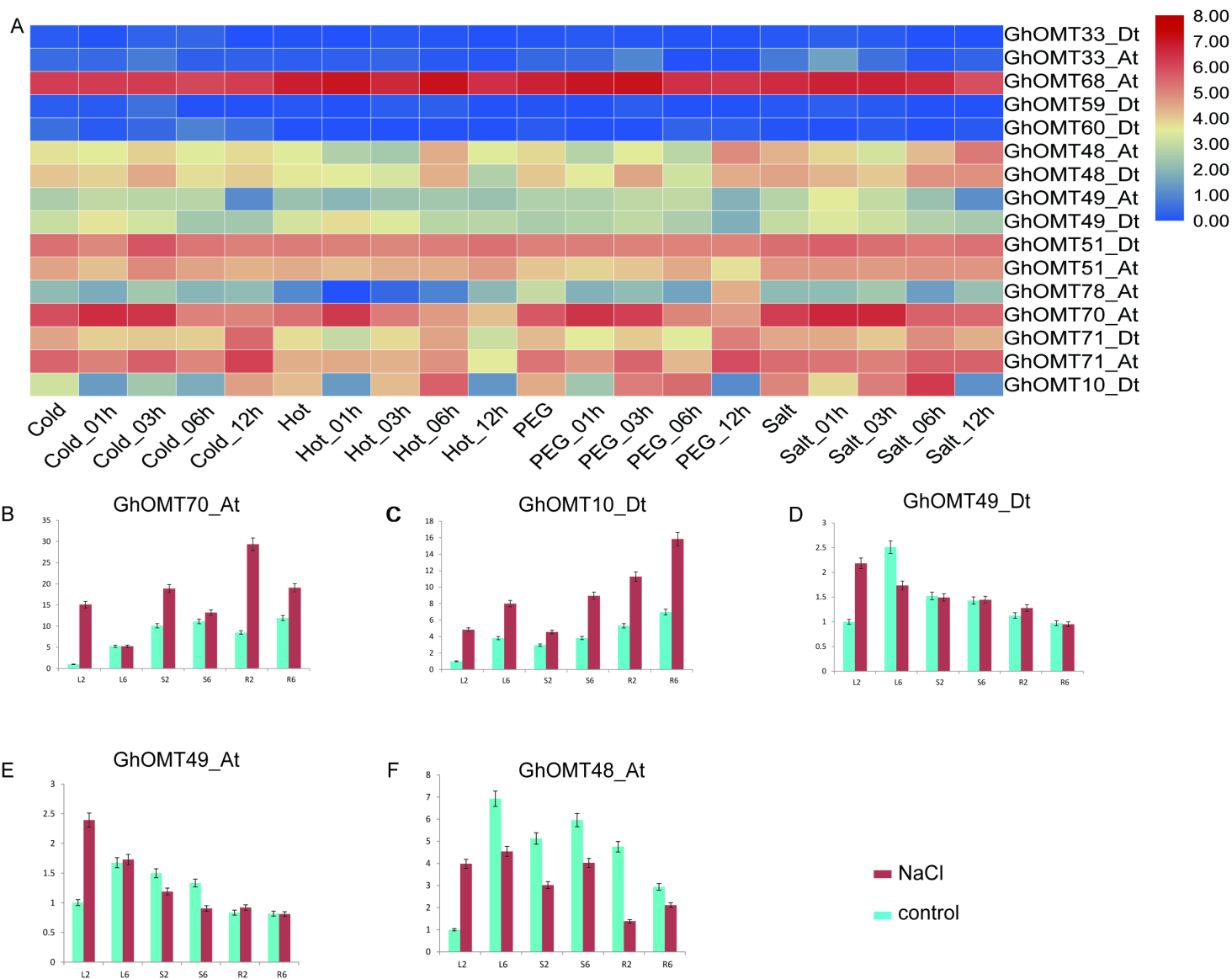


Figure 6

Specific responses of selected OMTs in salt stress treatment. A: Transcriptome heatmap of selected GhOMT genes in cold, hot, osmotic and salt treatments [28, 41]. B-F: qRT-PCR verification results of GhOMT70_At, GhOMT10_Dt, GhOMT49_Dt, GhOMT49_At, and GhOMT48_At in salt treatments of sGK9708.

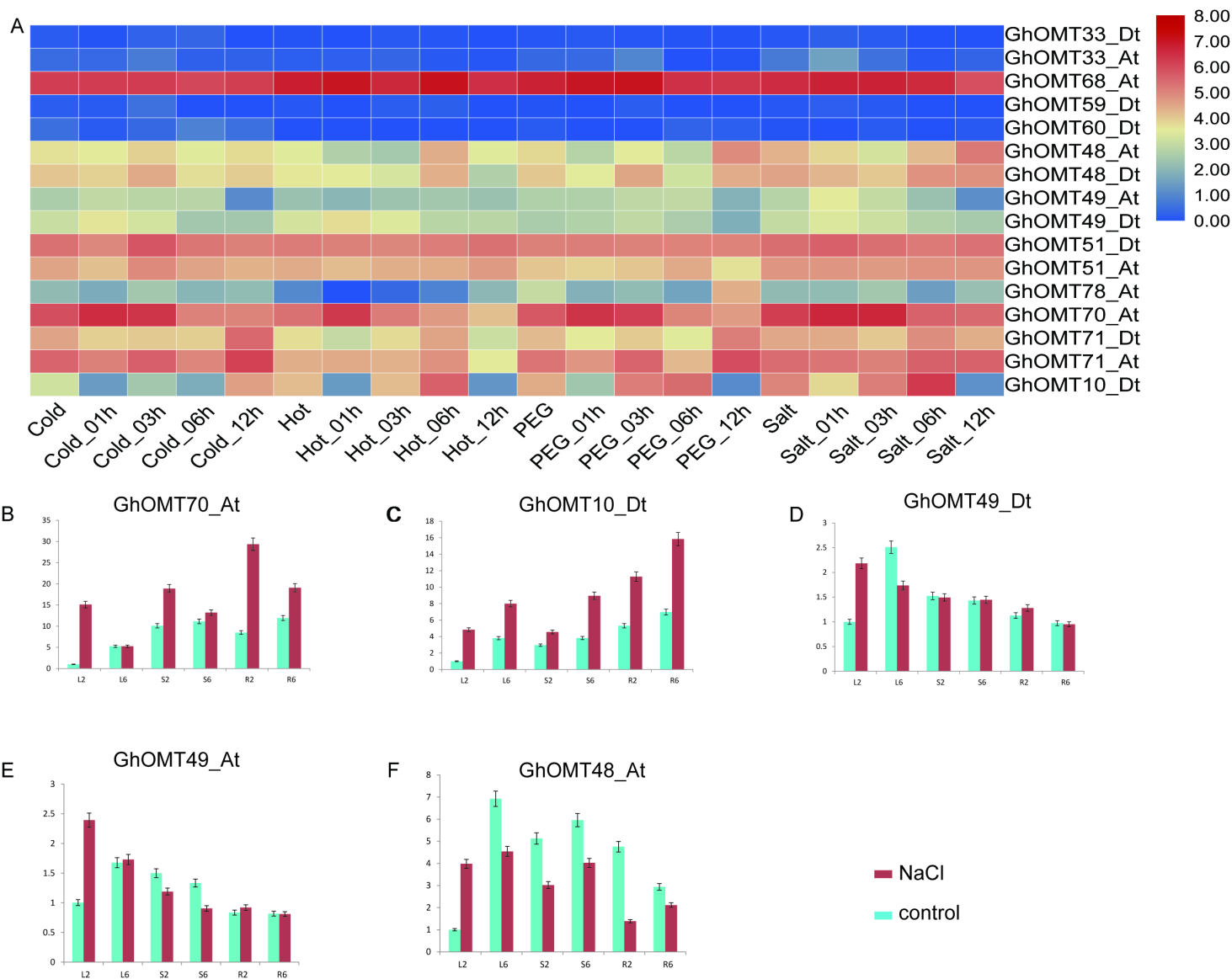


Figure 6

Specific responses of selected OMTs in salt stress treatment. A: Transcriptome heatmap of selected GhOMT genes in cold, hot, osmotic and salt treatments [28, 41]. B-F: qRT-PCR verification results of GhOMT70_At, GhOMT10_Dt, GhOMT49_Dt, GhOMT49_At, and GhOMT48_At in salt treatments of sGK9708.

Supplementary Files

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

- [TableS1.DetailInformationofplantmaterials.docx](#)
- [TableS1.DetailInformationofplantmaterials.docx](#)
- [TableS2.SequencesofPrimerPairs.docx](#)
- [TableS3.BasicinformationandanalysesofOMTs.xlsx](#)
- [TableS3.BasicinformationandanalysesofOMTs.xlsx](#)
- [TableS4.Analysesofduplicationevents.xlsx](#)
- [TableS4.Analysesofduplicationevents.xlsx](#)
- [TableS5.StructuralfeaturesofOMTgenes.xlsx](#)

- [TableS5.StructuralfeaturesofOMTgenes.xlsx](#)
- [FigureS1.ChromosomaldistributionofOMTsFinal.jpg](#)
- [FigureS1.ChromosomaldistributionofOMTsFinal.jpg](#)
- [FigureS2.IdentificationofmotifsofOMTsfinal.png](#)
- [FigureS2.IdentificationofmotifsofOMTsfinal.png](#)
- [FigureS3.IdentificationofcisregulatoryelementsofOMTs.png](#)
- [FigureS3.IdentificationofcisregulatoryelementsofOMTs.png](#)