

# Disequilibrium evolution of the Fructose-1,6-bisphosphatase gene family leads to their functional biodiversity in *Gossypium* species

**Qun Ge(Former Corresponding Author)**

Cotton Research Institute of Chinese Academy of Agricultural sciences

**Yànli Cūi**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Jùnwén Lǐ**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Jǔwǔ Gōng**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Quánwěi Lú**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Péngtāo Lǐ**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Yùzhēn Shí**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Hǎihóng Shāng**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Àiyīng Liú**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Xiǎoyīng Dèng**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Jìngtāo Pān**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Qúanjiā Chén**

Xinjiang Agricultural University

**Youlu Yuan**

Chinese Academy of Agricultural Sciences Cotton Research Institute

**Wankui Gong(New Corresponding Author)** (✉ [gongwankui@caas.cn](mailto:gongwankui@caas.cn))

<https://orcid.org/0000-0002-8556-0669>

**Keywords:** Cotton, Fructose-1, 6-bisphosphatase, Evolution, Translocation, Expression patterns

**Posted Date:** April 1st, 2020

**DOI:** <https://doi.org/10.21203/rs.2.21774/v2>

**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 Genomics on June 1st, 2020. See the published version at <https://doi.org/10.1186/s12864-020-6773-z>.

# Abstract

**Background:** Fructose-1,6-bisphosphatase (FBP) is a key enzyme in the plant sucrose synthesis pathway, in the Calvin cycle, and plays an important role in photosynthesis regulation in green plants. However, no systemic analysis of FBPs has been reported in *Gossypium* species.

**Results:** A total of 41 *FBP* genes from four *Gossypium* species were identified and analyzed. These *FBP* genes were sorted into two groups and seven subgroups. Results revealed that *FBP* family genes were under purifying selection pressure that rendered FBP family members as being conserved evolutionarily, and there was no tandem or fragmental DNA duplication in *FBP* family genes. Collinearity analysis revealed that a *FBP* gene was located in a translocated DNA fragment and the whole *FBP* gene family was under disequilibrium evolution that led to a faster evolutionary progress of the members in *G. barbadense* and in  $A_t$  subgenome than those in other *Gossypium* species and in the  $D_t$  subgenome, respectively, in this study. Through RNA-seq analyses and qRT-PCR verification, different *FBP* genes had diversified biological functions in cotton fiber development (two genes in 0 DPA and 1DPA ovules and four genes in 20–25 DPA fibers), in plant responses to *Verticillium* wilt onset (two genes) and to salt stress (eight genes).

**Conclusion:** The FBP gene family displayed a disequilibrium evolution pattern in *Gossypium* species, which led to diversified functions affecting not only fiber development, but also responses to *Verticillium* wilt and salt stress. All of these findings provide the foundation for further study of the function of *FBP* genes in cotton fiber development and in environmental adaptability.

## Background

Fructose-1,6-bisphosphatase (FBP, EC 3.1.3.11) catalyzes the decomposition of fructose 1,6-diphosphate (F-1,6-P<sub>2</sub>) into 6-phosphate fructose (F-6-P) and inorganic phosphorus (Pi) [1, 2]. It is ubiquitous across organisms and is a key enzyme in the Calvin cycle and the gluconeogenesis pathway [3, 4]. These reactions are involved in carbon fixation and sucrose metabolism and are present in the chloroplast stroma and the cytosol of green plants [5]. In most higher plants, FBP exists in three possible forms including a monomer, dimer, and tetramer, among which only the tetramer has catalytic activity[6].

In higher plants, based on their different catalytic mechanisms and independent evolutionary phylogensis, FBPs can be classified into two groups, cytosolic FBPs (cyFBPs) and chloroplast FBPs (cpFBPs). cyFBP plays an important regulatory role in the gluconeogenesis pathway and the synthesis of sucrose, while cpFBP is involved in the reduction of the pentose phosphate pathway [3, 4, 7]. cyFBP and sucrose phosphate synthase (SPS) are the main rate-limiting enzymes in the sucrose synthesis pathway [8]. 6-phosphate fructose is an essential monosaccharide for sucrose synthesis, and cyFBP and 6-phosphate fructokinase (PFK), pyrophosphate, and 1,6-diphosphate fructose transferase (PFP) jointly regulate the formation of fructose-6-phosphate. cyFBP can be inhibited by the metabolic product AMP, 2,6-diphosphoric fructose (F-2,6-P<sub>2</sub>), and also by Mg<sup>2+</sup> and Ca<sup>2+</sup>, while cpFBP is not sensitive to either

AMP or fructose-1,6-diphosphate [2]. Studies of FBP over expressions show that it can increase photosynthetic capacity, sucrose synthesis, and promote sugar accumulation, thereby accelerating plant growth.

Recently, some *FBP* genes have been cloned in several species such as *Beta vulgaris*, *Spinaciaoleracea*, *Glycine max*, *Arabidopsis thaliana*, *Pisumsativum*, *G. hirsutum* and *Pyropia haitanensis*, and other plants [5, 9-14]. The main research activities on these *FBP* genes have included identifying their functions in plant photosynthesis and glucose metabolism through molecular bioinformatic analysis and over-expression [11, 13-18]. In a transgenic study in *A. thaliana*, antisense transcripts were applied to inhibit the expression of a *cyFBP* gene. The decreased expression of the *FBP* gene resulted in decreased sucrose synthesis, accumulated intermediate metabolites, and eventually blocked photosynthesis [11]. In another study in *A. thaliana*, over-expression of a *cyFBP* gene caused an increase in sucrose synthesis and promoted plant growth in transgenic plants [15]. Inhibiting the expression of this gene in *Solanum tuberosum* could also reduce sucrose synthesis during the photosynthetic process [16]. In rice, loss of *cyFBP* reduced photosynthetic sucrose synthesis and delayed plant growth [17]. When *cpFBP* was inhibited in tomato, only small changes in carbohydrate metabolism were observed, but this inhibition caused a significant decrease in fruit size [18]. The different response modes of *PhcpFBP* mRNA levels in *Pyropia haitanensis* indicated that *cpFBP* also plays an important role in response to abiotic stresses such as high temperature and drought [14]. The different expression level of *GhFBP* at different times during cotton fiber development indicated that it plays a key role in the early stage of fiber secondary cell wall development[13].

Cotton is an important economic crop in the world, and cotton fiber is an important raw natural material for the textile industry. Cotton fiber is developed from the differentiation of a single ectodermic epidermal cell, and the fiber formation process can be divided into four distinct but partially overlapping periods: initiation, elongation (primary wall formation), secondary wall thickening, and dehydration maturity [19]. Many methods, including QTL identification [20-22], GWAS analysis [23-26], and functional gene identification [27-29], have been used to tackle the problems of fiber development and fiber quality formation. Studies have revealed that fiber development is a very complex process, with a large number of metabolic pathways providing material support, and thousands of specific genes being involved in expression regulation. At the same time, *Verticillium* wilt, which has the nickname "cotton cancer," is currently one of the most serious diseases that restricts cotton production and affects fiber quality [65]. A high concentration of saline stress also negatively affects the growth, development, and fiber quality of cotton [30, 31].

Although a few functional studies of *FBP* genes in some plant species have revealed that *FBP* genes could have certain impacts on various biological activities, *FBP* behavior is still poorly understood. Specifically, how *FBP* genes function at the whole genome level, especially in *Gossypium* species, remains unclear. The completion of whole genome sequencing databases for two important diploid cotton species *G. raimondii* [32, 33] and *G. arboreum* [34], and two domesticated tetraploid species *G. hirsutum* [35-38] and *G. barbadense* [37-39], provides brand-new platforms for functional genomic

studies. In this study, we identified 41 *FBP* family members in the genomes of these four cotton species and 73 *FBP* members in nine other species. Intensive bioinformatic analyses, including physicochemical properties, chromosomal localization, evolutionary relationships and gene structure, conserved motifs and *FBP* domain features, and functional expression analyses including transcriptomic and quantitative RT-PCR (qRT-PCR) were performed. The results indicated that *FBP* genes were involved in plant responses to biotic and abiotic stresses, as well as cotton fiber formation. This study provides a foundation for functional verification of the *FBP* genes of cotton in the future and useful information for the improvement of cultivars with excellent fiber quality and broad environmental adaptability.

## Results

### Identification of FBP family members

A total of 41 *FBP* genes from four *Gossypium* species, including 14 in *G. hirsutum* (*GhFBP*), 15 in *G. barbadense* (*GbFBP*), 6 in *G. arboreum* (*GaFBP*), and 7 in *G. raimondii* (*GrFBP*), were identified in this report (Supplementary file 1). The number of *FBP* genes in the tetraploid genomes of *G. hirsutum* and *G. barbadense* (AD genome) was almost double those in the diploid genomes of *G. raimondii* (D genome) and *G. arboreum* (A genome). These two tetraploid *Gossypium* genomes arose from a natural hybridization between two ancestors of diploid *G. raimondii* and *G. arboreum* [36, 38, 40].

In addition, in order to elucidate the evolutionary and phylogenetic relationship of these *FBP* genes, we identified 73 *FBP* family genes in nine other species, including 4 in *Arabidopsis thaliana*, 5 in *Theobroma cacao*, 12 in *populus trichocarpa*, 12 in *Glycine max*, 11 in *Zea mays*, 5 in *Vitis vinifera*, 6 in *Selaginella moellendorffii*, 11 in *Physcomitrella patens*, and 7 in *Oryza sativa* (Supplementary file 1).

### Phylogenetic analysis of the FBP Gene Family

To elucidate the evolutionary relationship of the identified FBP proteins between *Gossypium* and other species, the amino acid sequences of all the FBP proteins were aligned to identify their phylogenetic similarities with orthologs using the neighbor-joining model from MEGA 7, and a phylogenetic tree was thus constructed as shown in Figure 1a. According to their evolutionary relationships, 114 FBP proteins were divided into 2 groups: cytosolic FBPs, (*cyFBPs*) which included 40 members; and chloroplast FBPs, (*cpFBPs*) which included 74 members [7, 14]. The result of phylogenetic analysis indicated that FBPs had a closer evolutionary relationship between the four *Gossypium* species as compared with other species. The phylogenetic results also indicated that between all the other species, cocoa had the closest evolutionary relationship to the examined cotton species [36, 38]. Further phylogenetic analysis of *FBPs* from the four cotton species indicated that the *cyFBPs* were assorted into three subgroups, while the sorted into *cpFBPs* four subgroups (Figure 1b). Each subgroup of *Gossypium FBPs* consisted of six members, including one from the A genome (*G. arboreum*) and one from the D genome (*G. raimondii*), two from *G. hirsutum*, and two from *G. barbadense*. As both of *G. hirsutum* and *G. barbadense* are comprised of  $A_t$  and  $D_t$  subgenomes, each subgenome provided one member in each sub-group of the

*FBP* family. There is only one subgroup in cpFBPs that had 5 *FBPs*, but there was no *FBP* from *G. arboreum* identified in these analyses (Figure 1b).

### Gene structure and protein domain of FBP family members

The length of amino acid (aa) sequences of FBP proteins ranged from 341 to 608, 341 to 412, 341 to 428, and 341 to 606 in *G. arboreum*, *G. raimondii*, *G. hirsutum*, and *G. barbadense*, respectively. The *cyFBP* group had 18 members (42.87%), which had a uniform length of 341 aa with only two exceptions, namely Gorai.005G080300.1 and GB\_A02G1288.1. The *cpFBP* group had 23 members (57.13%), which had a varied length of aa sequences (Figure 1, Supplementary file 2). The PI values of the four cotton FBPs ranged from 5.00 to 7.68.

In total, 10 motifs were identified in the FBP family in the four *Gossypium* species, with each FBP containing 7 to 9 motifs in general (Figure 2a, Figure S1). The significant difference between *cyFBPs* and *cpFBPs* was that motif 5 was identified exclusively in *cyFBPs*, while motif 9 was exclusively present in *cpFBPs*. Each phylogenetic subgroup had a similar composition and arrangement of motifs, which was highly consistent with the results of phylogenetic analysis. The results also showed some minor variance in motif composition and arrangement between the subgroups (Figure 2a).

Gene structure analysis also showed consistent results to our phylogenetic and protein motif analyses (Figure 2b). The exon number of FBP genes ranged from 3 to 12. *cyFBPs* had 11 to 12 exons, while *cpFBPs* only had 3–5 exons. The gene structure of each subgroup was almost the same, which indicated conserved evolution patterns for *FBP* family members. The *cyFBP* gene structures could be further divided into three types (Figure 2). Both subgroups *cyFBP* 1 and *cyFBP* 2 had 12 exons and 11 introns, with a varied distribution between them. Subgroup *cyFBP* 3 had 11 exons and 10 introns. In contrast to *cyFBPs*, *cpFBPs* had much fewer exons. The *cpFBP* genes could be sorted into four subgroups. Subgroups *cpFBP* 1 and *cpFBP* 2 had 4 exons and 3 introns, with different distributions between them. Subgroup *cpFBP* 4 had 8 exons and 7 introns, while subgroup *cpFBP* 3 had a varied number of exons and introns, and the exon number of this subgroup ranged from 3 to 8. The results also indicated that only FBP genes from *G. raimondii* had UTR structures. This indicated that *cyFBPs* had more complicated gene structures than *cpFBPs* had.

To further understand how *FBP* genes function, the composition and distribution of *cis*-regulatory elements (CRE) were identified in the 5' untranslated regions 2000 bp upstream of each gene from the PlantCare website (Figure 3). The results indicated that the composition and distribution of CREs varied significantly across the whole *FBP* gene family. It also could be seen that the CREs had a high congruency with the results of gene structure, protein domain, and phylogenetic analyses. Each subcategory of *FBP* genes had identical or similar compositions and distributions of CREs in their 5' upstream regions (Figure 3).

Further analysis indicated that the 5' up-stream regions of FBP genes contained almost all of the following categories of CREs: constitutive, inducible and tissue-specific. The constitutive CREs include

typical basic components such as TATA-Boxes and CAAT-Boxes. Inducible CREs included photo-responsive elements, ATCC-motifs, Box 4, I-Boxes, Sp1, TCCC-motifs, GAG-motifs, gibberellin response elements (GARE-motifs), P-Boxes, abscisic acid responsive elements (ABREs), salicylic acid reaction elements, TCA-elements, anaerobic induction elements (AREs), stress-responsive elements, TC-rich repeats, and MYB binding site (MBS). In addition, the GARE-motif was exclusively identified in the promoter region of one subcategory of genes including *GH\_A02G0701.1*, *GH\_D02G0715.1*, *GB\_A02G0693.1*, *GB\_D02G0741.1*, *GH\_A02G1268.1*, and *GB\_A02G1288.1*.

### **Distribution and collinearity analysis of the *FBP* gene family *Gossypium* species**

In the genome of *G. arboreum*, *FBP* genes were identified on chromosomes A02, A03, A04, A10, A11, and A12, while in the genome of *G. raimondii*, *FBP* genes were identified on chromosomes D02, D05, D07, D08, D11, and D12. In the tetraploid genomes of *G. hirsutum* and *G. barbadense*, *FBP* genes had similar distribution on chromosomes A<sub>t</sub>02, A<sub>t</sub>04, A<sub>t</sub>10, A<sub>t</sub>11, A<sub>t</sub>12, D<sub>t</sub>02, D<sub>t</sub>03, D<sub>t</sub>04, D<sub>t</sub>10, D<sub>t</sub>11, and D<sub>t</sub>12. Homologous analysis indicated that a homologous gene identified on A03 of *G. arboreum* was identified on chromosome A<sub>t</sub>02 in *G. hirsutum* and *G. barbadense*.

Tandem and fragmental DNA duplication provides major forces that drive the formation of gene families [41, 42] as well as whole genome evolution. In the current study, the duplication events of cotton *FBP* genes were analyzed. Although the results did not support any tandem repeat events occurring during the evolution of the cotton *FBP* gene family, collinearity analysis showed that in these two diploid species the *FBP* genes were perfectly chromosome-pair-wise homologous (Figure 4a). Meanwhile, in the two tetraploid species, each *FBP* gene from one species (*hirsutum* or *barbadense*) had two homologous genes in both the A<sub>t</sub> and D<sub>t</sub> subgenomes in its counterpart species (*barbadense* or *hirsutum*) (Figure 4d). Collinearity analysis between diploid and tetraploid species indicated that in *G. hirsutum* each gene had two homologous genes in the two diploid species (Figure 4b), while in *G. barbadense*, two *FBP* genes on GbA<sub>t</sub>02 did not have homologous genes in *raimondii* and one *FBP* gene at GbD<sub>t</sub>12 did not have a homologous gene in *arboreum* (Figure 4c).

### **Analysis of selection pressure of *FBP* genes in four cotton species**

Calculating non-synonymous (Ka) and synonymous (Ks) substitution rates is a useful method for assessing sequence variation of protein orthologous in different species or taxa with unknown evolutionary states [43]. The value of Ka/Ks represents the ratio between Ka and Ks of two homologous protein-coding genes. Ka/Ks > 1 indicates that a gene has been positively selected, while a Ka/Ks = 1 indicates that a gene has been neutrally selected, and a Ka/Ks < 1 indicates that a gene has been selectively purified [43]. The Ka/Ks values of homologous *FBP* genes between *G. arboreum* and *G. raimondii* ranged from 0.05 to 0.62, while those between *G. hirsutum* and *G. arboreum* or *G. raimondii* ranged from 0 to 0.8. Those between *G. barbadense* and *G. arboreum* or *G. raimondii* ranged from 0 to 0.6, and the values between A<sub>t</sub> and D<sub>t</sub> paralogous genes in *G. hirsutum* and *G. barbadense* ranged 0.07 to

0.76 and 0.02 to 0.52, respectively (supplementary file 3). These results indicated that the *FBP* genes in these four *Gossypium* species were under purifying selection.

### **FBP gene expression in fiber development and in response to biotic and abiotic stresses**

To explore the potential function of *FBP* genes in the growth and development of cotton fibers, we downloaded cotton fiber transcriptome data from the NCBI SRA database and reanalyzed the expression profiling of *FBP* genes. The results of *FBP* gene expression analysis showed that the homologous genes *GH\_A02G0701.1* and *GH\_D02G0715.1* from *G. hirsutum*, and *GB\_A02G0693.1* and *GB\_D02G0741.1* from *G. Barbadense* had higher FPKM values in developing fibers at 20 days post-anthesis (DPA) and 25 DPA (supplementary file 4). The homologous genes *GH\_A02G1268.1* and *GB\_A02G1288.1* had high expression FPKM values in the early stage of the fiber development (0 DPA and 1 DPA ovule) (Figure 6a, b). The expression of *GH\_D02G0715.1* and *GH\_A02G0701.1* in the secondary cell wall synthesis stage of fiber development through qRT-PCR validation assays were consistent with *in silico* transcriptome analysis (Figure 6c, d).

In plant response to *Verticillium* wilt stress, the FPKM values of the *FBP* gene family members that were extracted from the previously mentioned transcriptome data showed that the homologous genes *GH\_A04G1526.1* and *GH\_D04G1869.1* had much higher expression values at 24 and 48 hours after inoculation (HAI) with *Verticillium dahliae*, with their highest peaks being reached at 24 HAI (Figure 7a, supplementary file 4). These results suggested a certain biological function of *FBP* genes in plant responses to *Verticillium* wilt stress.

The results of qRT-PCR analysis showed that both *GH\_A04G1526.1* and *GH\_D04G1869.1* had different expression behaviors in root tissues between susceptible and resistant cultivars at different developmental stages of *Verticillium dahliae* after inoculation. In the VW tolerant cultivar Jimian 11(J11), both *GH\_A04G1526.1* and *GH\_D04G1869.1* had immediate responses to inoculation with *Verticillium dahliae* and their expression levels reached a maximum at 12 HAI. The levels then dropped rapidly and maintained fairly low expression levels (Figure 7b and 7c). In the VW susceptible cultivar ZZM, *GH\_A04G1526.1* and *GH\_D04G1869.1* acted differently, with *GH\_A04G1526.1* slightly increasing its expression after inoculation up to 48 HAI, followed by its expression increasing rapidly and reaching a peak at 72 HAI (Figure 7b), while *GH\_D04G1869.1* maintained low expression throughout the entire experimental procedure (Figure 7c). These different responses suggested that *GH\_A04G1526.1* might take part in resistant reactions, while *GH\_D04G1869.1* participated in susceptible reactions to *Verticillium* wilt in cotton.

The responses of *FBP* genes to salt stress were also evaluated using RNA transcriptome data analysis [44] under salt stress (Figure 8, supplementary file 4). Our transcriptome analysis indicated that six members of the *FBP* family, *GH\_A10G2530.1*, *GH\_D10G2661.1*, *GH\_A11G3741.1*, *GH\_D11G3768.1*, *GH\_A02G1268.1*, and *GH\_D03G0740.1*, had significantly higher responsive expression to salt stress treatments in foliage and two members, *GH\_A04G1526.1* and *GH\_D04G1869.1*, had significantly higher responsive expression in roots (Figure 8). In the salt susceptible cultivar CCR112, the tested genes that

had expressions in foliage had similar expression tendencies in responses to salt pressure. Their expressions were significantly inhibited within three hours after salt stress was imposed. This inhibition continued and reached its highest at 12 hours after the initiation of stress. After this time, as time proceeded, the plant began to develop some sorts of “adaption” mechanisms, and their expression recovered to a certain level. In the salt tolerant semi-wild species MAR85, the inhibition of these genes was to a much smaller extent. It could be seen from our results that the expression levels of these genes at 12 h from salt resistant material were almost double those from the salt sensitive materials. These expression differences between two cultivars reached significant level at least in one treatment stage (Figure 8). Both *GH\_A04G1526.1* and *GH\_D04G1869.1* had significant higher responsive expressions in root tissues of CCRI12 than in root tissues of MAR85 (Figure 8).

## Discussion

### Cotton *FBP* genes identification

FBPase decomposes fructose 1,6-diphosphate to 6-phosphate fructose and inorganic phosphate. It provides a key substrate for sucrose synthesis in the cytoplasm, participates in the regulation of the sucrose synthesis pathway, and affects the distribution of carbohydrates. cpFBP is mainly involved in the Calvin cycle and the synthesis of starch [45], and cyFBP is mainly involved in the synthesis of sucrose [3]. Studies have revealed that two reactions with FBP isozymes are in the branching points from which their metabolites flow from the Calvin cycle into the process of either starch or sucrose synthesis. In *A. thaliana*, an *FBP* mutant resulted in the decrease in soluble sugar content and starch accumulation, and a significant increase in SOD activity [46]. The mutant also led to developmental changes including an increase in the number of root vascular bundles [46]. The cloning of the *FBP* gene has also been reported in several plant species including cotton, but there has been no further report in cotton regarding *FBP* genes or gene family.

Our study suggested that the *FBP* gene family maintained a certain stability in different diploid species including *A. thaliana* (in which four *FBP* genes were identified), *T. cacao* (five *FBP* genes), *G. arboreum* (seven *FBP* genes), and *G. raimondii* (seven *FBP* genes). There were no identified tandem repeats or segmental duplications in *FBP* family across these species (Figure 1 and 4). The fluctuation of the environment forces organisms living in it to continuously evolve to adapt to changes [47]. Gene duplication is a basic process in genome evolution [48]. A duplicated gene may face several evolutionary fates, either becoming degraded by loss of function mutations, shared due to gene dosage effects or becoming sub-/neo-functionalized [48]. Yet the exact evolutionary events functioning on a gene are always episodic [47]. Polyploidization is one major episodic evolutionary event that can provide whole genome level genetic variation for plant adaptive evolution [49]. Even though the mechanisms of how polyploidy genomes establish their genetic structure are still unclear, such duplicated genes in polyploidy species may face the same evolutionary fates [48]. Our *FBP* collinearity analysis suggested that there was no tandem or segmental duplications in this *FBP* gene family in *Gossypium* species, despite the whole genome replication event due to polyploidization, which gave rise to the allotetraploid AD genome

species (*G. hirsutum* and *G. barbadense*) from a natural hybridization between ancestors of *G. arboreum* or *G. herbaceum* (A group) and *G. raimondii* (D group) 1.7 to 1.9 million years ago [36, 38, 50, 51]. Afterwards, these species, which were involved in this hybridization event, independently evolved and were domesticated in different geographical areas [38]. Taken the findings in the current study that this *FBP* gene family was subject to purifying selection pressure based on their lower Ka/Ks values and that the high similarities in gene structures and *cis* regulatory elements in the 5' upstream regions of each subgroup of the *FBP* gene family, we concluded that the *FBP* gene family in *Gossypium* genomes was evolutionarily conserved.

A noticeable fact of the collinearity analysis in our study was that some potential collinearity of *FBP* gene pairs between *G. barbadense* and *G. raimondii*, between *G. barbadense* and *G. arboreum* and between *G. hirsutum* and *G. raimondii*, were lost (Figure 4c). These findings were supported by their independent evolutions after phylogenesis of each species. These results also implied that the *FBP* gene members in *G. barbadense* may have undergone a faster evolution than those on the rest of the *Gossypium* species in this study. This result was highly consistent with previous findings [52-54]. Moreover, we also demonstrated additional structural rearrangements in the A<sub>t</sub> subgenome rather than in the D<sub>t</sub> subgenome of *G. hirsutum* [53]. Evidence showed asymmetric subgenome domestication for directional selection towards fiber length [54]. *NAC* family genes in the A<sub>t</sub> and D<sub>t</sub> subgenomes displayed asymmetric molecular evolution in terms of gene loss rates, evolutionary rates, and expression levels [52]. Recent evidence has demonstrated a faster evolution rate in allotetraploid cotton than in diploid cotton, in the A<sub>t</sub> subgenome rather than in the D<sub>t</sub> subgenome, and in *G. barbadense* than in *G. hirsutum* [38].

Evidence has demonstrated that both *G. hirsutum* and *G. barbadense* were from one transoceanic hybridization event between two ancestors of diploid species *G. arboreum* or *G. herbaceum* and *G. raimondii* [50, 51]. Through evolutionary analysis of the genomic sequence, we also demonstrated that after diploid cotton differentiated from a common ancestor of cocoa, and five or more genome-wide replication events occurred [36]. Actually, although this allotetraploid cotton species had conserved gene order and collinearity between the A and D subgenomes [24], there were at least 9 translocations and 28 inversions identified in *G. hirsutum* draft genome [36]. A recent study confirmed these translocations between the A genome and the D genome in diploid cotton [38]. Previous comparative studies demonstrated that, in *Gossypium*, chromosomes D03 of *G. raimondii*, A02 of *G. arboreum*, GhA<sub>t</sub>03 of *G. hirsutum*, and GbA<sub>t</sub>03 of *G. barbadense* were homologous [38]. A reciprocal translocation event between chromosome A01 and A02 in diploid *G. arboreum* was estimated before the ancestor of the extant domesticated tetraploid cotton was formed [36, 38]. Our results revealed that the genes *Gorai.002G055400.1* in D02, *Ga02G0778.1* in A02, *GH\_A02G1268.1* in GhA<sub>t</sub>02, *GH\_D03G0740.1* in GhD<sub>t</sub>03, *GB\_A02G1288.1* in GbA<sub>t</sub>02, and *GB\_D03G0699.1* in GbD<sub>t</sub>03 were homologous genes (Figure 4). This fact demonstrated that the gene *Ga02G0778.1* in A02 was located on a DNA fragment that was translocated from A01 to A02 (Figure 6a, 6b). How these translocations affect the genes expression patterns is therefore still unknown. We did, however, demonstrate a significant expression profiling

difference between these orthologous genes regardless of whether they were on the translocated fragment.

### Functional biodiversity of *FBP* genes in *Gossypium* species

FBPs in our study were mainly sorted into two groups based on their subcellular compartmentalization, which conditioned their respective functions and regulatory mechanisms [7]. cpFBP is a key carbon-metabolizing enzyme that catalyzes the conversion of fructose-1,6-bisphosphate to fructose-6-phosphate in the Calvin cycle and eventually to starch, while cyFBP a key enzyme that catalyzes the conversion of triose phosphates to sucrose in the sucrose-biosynthesis pathway [7]. In the animal kingdom, FBP was found to be involved in various disease reactions, while in plants, it has been demonstrated to be involved in sugar partitioning [7], and thus, participates in plant growth [15, 17], as well as plant responses to biotic and abiotic stresses [14]. In the current study, we identified elevated expression of *FBP* genes during *Gossypium* fiber development (Figure 6), in response to salt treatment (Figure 8), and in response to *V. dahliae* inoculation (Figure 7).

Cotton fibers are formed by a single ovule outer bead epidermal cell through a specialized process of initial differentiation, elongation, thickening, and dehydration to form mature epidermal fibers [19, 50, 55, 56]. The onset of cotton fiber differentiation usually occurs near the day of flowering (0 DPA). These initially differentiated fibrocytes develop into fiber cells. The secondary expansion of epidemic cells is initiated after 7 to 10 DPA, and these cells gradually develop into fuzz that is attached to the surface of a seed [55-57]. Factors that influence the initiation, differentiation, and development of fibers include phytohormones, such as IAA and GA, and brassinolactone, as well as transcription factors [58-61]. The arrangement of cellulose deposition plays an important role in the formation of fiber strength and fineness [62, 63]. It has been demonstrated that the sucrose content in the developing fibers of immature-fiber-mutants are significantly lower than that of TM-1 [64]. This could be explained by the finding that sucrose, which is imported into developing fiber cells, is the major carbon source for cellulose synthesis and energy production there. Sucrose, together with K<sup>+</sup> and malate, also provides the major active osmotic solutes required to maintain cell turgor for fiber elongation [65]. It has also been demonstrated that suppressing *GhSCP2D* expression in cotton fiber eventually activates sucrose transporter genes such as SUTs and SWEETs in early fiber developmental stages [66]. The transcription factor GhMYB212 was also demonstrated to be involved in sucrose transport into elongating fiber cells [26]. A sufficient and sustainable supply of this sugar is essential for fiber development. Here we demonstrated high expression of some *cyFBP* genes, including *GH\_A02G0701.1*, *GH\_D02G0715.1*, and *GH\_A02G1268.1*, in developing cotton fibers, which hinted that these *cyFBP* genes probably assumed the responsibility of maintaining the supply of sucrose in developing fiber cells, especially during late secondary cell wall thickness stages.

*Verticillium* wilt is the most important disease restricting cotton production due to its serious impacts on cotton yield and fiber quality [67]. Bioinformatic analysis revealed that the 5'-upstream regulatory regions of *cyFBP* genes in the current study, in particular in *GH\_A04G1526.1* and *GH\_D04G1869.1* (Figure 3),

harbor pathogen-induced, *cis*-acting elements, including W-boxes, GCC-motifs, TC-rich repeats, JERE-motifs, Gstl-motifs and D-motif components, G boxes, GT-1, GCC boxes and S boxes [68-71]. These findings hinted that these cyFBP genes affect a certain impact on plant responses to the onset of *Verticillium* wilt. It is generally recognized that the responses of cotton plant to *Verticillium* wilt are mainly accomplished by thickening the cell wall and blocking the catheter, which includes gel formation, epidermal lignification, and internal tissue cork modification to form an invasion plugging body, which prevents the spread of the pathogen in the plant and confines the invading *V. daliae* to a restricted local area [72]. After infection with *V. daliae*, corresponding changes in cotton plant metabolism also include the biosynthetic enhancement of lignin, phytoalexin, tannin, polysaccharide, gossypol, and lipids [69, 73]. Lignification of the cell wall enhances its mechanical resistance to the penetration of fungal appressoria [74]. The lignin-based disease resistance mechanisms are very important in the processes of plant pathogenic microorganism interactions [74]. Lignification has been extensively studied in cotton resistance [75-77], and tolerance to salt stress [78]. Actually, evidence suggests that increasing metabolizable sugar content should positively influence the abundance of lignins. Researchers have speculated that there also may be a mechanism in plants that senses changes in sugar content and regulates lignin biosynthesis [79].

In the current study, we identified significant differences in the inhibition of both *cpFBP* and *cyFBP* gene expression levels between CRI12, an upland cotton cultivar, and MAR85, a semi-wild relative of upland cotton (Figure 8). Previous studies revealed that a *cpFBP* gene from *Oriza sativa* was more sensitive to salt stress treatments than that from its close halophytic relative, *Porteresia coarctata* [80, 81]. The *cpFBP* genes from the two plants were determined to have five amino acid residue differences [82]. Xiao et al. [14] detected a significant inhibition of *PhcpFBP* gene expression at the early stages of high temperature stress, while with persistent high temperature stress, the expression of *PhcpFBP* gradually resumed a normal level. Hark et al.[83] determined that the *cyFBP* gene was not expressed continuously in severe water stress, although a low level of enzyme activity was still present (33% of control). These studies indicated that both the *cpFBP* and *cyFBP* genes might be involved in plant responses to salt stress in one way or another. At the same time, the genes that have elevated expressions in response to inoculation with *V.daliae* also showed responsive expression to salt treatment (Figure 8). This implies that the pathways these genes catalyze might be correlated to plant responses to both salt stress and *V. daliae* inoculation.

## Conclusions

In this study, a total of 41 *FBP* genes were identified in four cotton species, distributed to six chromosomes in diploid cotton and 11 chromosomes in tetraploid cotton, and could be classified into two classes. There were large translocations between A02 and A03 chromosomes in tetraploid cotton and there was an inversion translocation on the A02 and D02 homologous chromosomes in diploid cotton, meaning the collinearity of the *FBP* genes changed. Through our results from RNA-seq and qRT-PCR, we know that the different *FBP* genes had different functions, not only affecting fiber development, but also the response to *Verticillium* wilt and salt stress. Additionally, GARE-motif may affect gene

expression during the fiber development. Genes of the same subclass have the same motif, gene structure, and similar *cis*-acting elements, meaning they have similar functions. All of these results provide a foundation for further study on the function of the *FBP* genes in cotton fiber development and environmental adaptability.

## Methods

### Plant growth and tissue sampling

#### Fiber sample

The *G. hirsutum* cultivars ZR014121 and EZ60 were grown in an experimental farm in Anyang, Henan, China. The flowers were tagged on the day of anthesis and that day was regarded as 0 DPA. Flowers and developing bolls were sampled on 0, 5, 10, 15, 20, 25, and 28 DPA from the tagged flowers and the ovules and developing fibers were harvested from the developing bolls and samples were immediately dipped in liquid nitrogen and stored at -80°C for future use.

#### Root samples after *Verticillium dahliae* inoculation

The *G. hirsutum* cultivars Jimian 11 (J11) and ZZM were planted in a green house in the laboratory. At the two-leaf stage, the cotton plants were inoculated with strain vd080 of *Verticillium dahliae* by root-dips. Roots were sampled at 0 hour (0 H), 6 H, 12 H, 24 H, 36 H, 48 H, 72 H, and 96 H after inoculation. Three replicates were collected and immediately stored at -80°C.

### Transcriptome Analyses and Quantitative Real-Time PCR

Transcriptome data of cotton fiber development of *G. hirsutum* TM-1 and *G. barbadense* Hai7124 were downloaded (accession number PRJNA490626) from the Sequence Read Archive (SRA) of the NCBI database (<https://www.ncbi.nlm.nih.gov/>) [38]. Raw read data (Sra) were converted to Fastq data using Fastq-dump, and Fastq data were filtered using Trimmomatic-0.36 to remove residues and low quality reads. The filtered data were then assembled from scratch using Trinity, the transcripts were assembled, and the reads were then posted back to the reference genome to calculate expression. After FPKM values were obtained, a heat map was generated using MEV 4.0 software [84].

The expression levels of the members of the *G. hirsutum* *FBP* gene family were extracted and re-analyzed from the transcriptome data of MBI8255 and CCRI36 after *Verticillium dahliae* inoculation using high-throughput sequencing methods described in a previous study [85]. The transcriptome data from CRI12 and MAR85 in salt stress were kindly provided by Dr. Liu's Lab [44]. The expression heat map of each member of the *G. hirsutum* *FBP* gene family was created using MEV 4.0 software [84].

Total RNA samples of cotton fiber and roots were extracted using an RNAPrep Pure Plant Kit from Tiangen Biotechnology Co., Ltd. (Beijing, China) and were visualized on 1% agarose gels. cDNA was synthesized using a FastQuant RT Kit (Tiangen Biotech (Beijing) Co., Ltd.). Primer premier 5.0 software

was used to design gene specific primers for fluorescent quantitative PCR based on full-length gene cDNA. The *GhHistone3* gene was used as an internal reference gene. qRT-PCR was performed using a LightCycler® 480 II Real-Time PCR Instrument (Basel Roche, Switzerland). The expression levels of the genes were calculated using the  $2^{-\Delta\Delta CT}$  method, with three independent PCR amplifications [86].

### Identification of cotton *FBP* genes

The latest versions of the genomes of four cotton species were downloaded from CottonFGD (<https://cottonfgd.org/>) [87], including *G. hirsutum* (<https://cottonfgd.org/about/download/assembly/genome.Ghir.ZJU.fa.gz>), *G. barbadense* (<https://cottonfgd.org/about/download/assembly/genome.Gbar.ZJU.fa.gz>), *G. arboreum* (<https://cottonfgd.org/about/download/assembly/genome.Garb.CRI.fa.gz>), *G. raimondii* (<https://cottonfgd.org/about/download/assembly/genome.Grai.JGI.fa.gz>). Some of the genomes for comparative analysis, including *T. cacao* [23], *Oryza sativa* [25], *A. thaliana* [26], *P. trichocarpa*, *V. vinifera*, *S. moellendorffii*, *Z. mays*, *G. max*, and *P. patens*, were downloaded from the ePhytozome 2.1 database (<http://www.phytozome.net>). A hidden Markov model (HMM) profile (PF00316) was downloaded from EMBL-EBI (<http://pfam.xfam.org/family/PF00316/hmm>). The hmmsearch program in HMMER 3.0 software [88] was then used to search for protein sequences of these four *Gossypium* species using Ces.hmm as a model and --cut\_ga as a parameter. The relative molecular weight and the theoretical isoelectric point prediction of the obtained amino acid sequences of FBP proteins were performed using the ExPASy (<http://expasy.org/>) online database.

### *FBP* gene family system evolution, gene structure, and chromosomal distribution analysis

Multi-sequence alignments of all FBP protein sequences were performed using Clustal X[89], and phylogenetic trees of related proteins were constructed using the MEGA software proximity method (version 6.06) (Neighbor-Joining, NJ) [90], with the calibration parameter Bootstrap being set to 1000. The structure of the candidate genes was analyzed using the online software GSDS 2.0 (<http://gsds.cbi.pku.edu.cn/>) [91]. The MEME online tool (<http://meme-suite.org/>) was used to analyze the conserved motifs of the amino acid sequences (Motif), in which the parameter settings were as follows: the minimum length of the conserved motif was six, the maximum length was fifty, and the maximum motif number was ten.

### Collinearity and selection pressure analysis

To determine whether the *FBP* gene family expanded through segmental duplication or tandem duplication events, a collinear analysis was completed with an all-to-all BLAST array (*E*-value of  $1e-5$ ) in the MCScan program[92]. Selection pressure analysis was performed by calculation of the *K<sub>a</sub>* (non-synonymous substitution rate) and *K<sub>s</sub>* (synonymous substitution rate) values of homologous genes using KaKs\_Calculator 2.0 software.

### Prediction of *cis*-acting elements in cotton FBP gene promoters

Two thousand bp upstream sequences of the candidate genes were extracted from the genome sequences of the four *Gossypium* species. *cis*-acting elements were predicted using online software PLANT CARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)[93].

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

The latest versions of the genome of four cotton species were downloaded from CottonFGD (<https://cottonfgd.org/>), including *G. hirsutum* (<https://cottonfgd.org/about/download/assembly/genome.Ghir.ZJU.fa.gz>), *G. barbadense* (<https://cottonfgd.org/about/download/assembly/genome.Gbar.ZJU.fa.gz>), *G. arboreum* (<https://cottonfgd.org/about/download/assembly/genome.Garb.CRI.fa.gz>), *G. raimondii* (<https://cottonfgd.org/about/download/assembly/genome.Grai.JGI.fa.gz>). Some of the genomes for comparison analysis, including *T. cacao*, *Oryza sativa*, *A. thaliana*, *P. trichocarpa*, *V. vinifera*, *S. moellendorffii*, *Z. mays*, *G. max* and *P. patens* were downloaded from the ePhytozome 2.1 database (<http://www.phytozome.net>).

All extra data generated or analyzed during this study are included in this published article.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This study was supported in part by grants from the National Key R&D Program of China (2017YFD0101600, 2016YFD0100500, and 2016YFD0101401), the Agricultural Science and Technology Innovation Program for CAAS (CAAS-ASTIP-ICRCAAS), and the Natural Science Foundation of China (31801404).

### Author Contributions

QC, YY and WG initiated the research. QG, JL, JG, and QL designed the analysis and experiments. QG, YC, and JP performed the functional analysis and molecular experiments. YS, HS, AL, XD, and PL performed

genomic analyses. QG and WG drafted the manuscript. QC, YY, and WG finalized the manuscript. All authors contributed in the interpretation of results and approved the final manuscript.

## Acknowledgements

Not applicable

## Abbreviations

aa: Amino acid; DPA: Day post anthesis; Ga: *Gossypium arboreum*; Gb: *Gossypium barbadense*; Gh: *Gossypium hirsutum*; Gr: *Gossypium raimondii*; Ka: Nonsynonymous substitution rate; Ks: Synonymous substitution rate; FBP: Fructose-1,6-bisphosphatase; qRT-PCR: Quantitative real-time polymerase chain reaction; SPS: sucrose phosphate synthase.

## References

1. Al-Robaiy, S. and K. Eschrich, *Rat muscle fructose-1,6-bisphosphatase: cloning of the cDNA, expression of the recombinant enzyme, and expression analysis in different tissues*. Biol Chem, 1999. **380**(9): p. 1079-1085.
2. Zimmermann, G., G.J. Kelly, and E. Latzko, *Efficient purification and molecular properties of spinach chloroplast fructose 1,6-bisphosphatase*. Eur J Biochem, 1976. **70**(2): p. 361-367.
3. Daie, J., *Cytosolic fructose-1,6-bisphosphatase: A key enzyme in the sucrose biosynthetic pathway*. Photosynth Res, 1993. **38**(1): p. 5-14.
4. Nel, W.T., S. E., *Plant fructose-1,6-bisphosphatases: characteristics and properties*. Int J Biochem, 1992. **24**(8): p. 1267-1283.
5. Anderson, L.E., R. Yousefzai, M.R. Ringenberg, and A.A. Carol, *Both chloroplastic and cytosolic fructose bisphosphatase isozymes are present in the pea leaf nucleus*. Plant Science, 2004. **166**(3): p. 721-730.
6. Schnarrenberger C and M. W, *The Calvin cycle-a historical perspective*. Photosynthetica, 1997. **33**: p. 331-345.
7. Serrato, A.J., J. de Dios Barajas-Lopez, A. Chueca, and M. Sahrawy, *Changing sugar partitioning in FBPase-manipulated plants*. J Exp Bot, 2009. **60**(10): p. 2923-2931.
8. Huber, S.C. and J.L. Huber, *Role of sucrose-phosphate synthase in sucrose metabolism in leaves*. Plant Physiol, 1992. **99**(4): p. 1275-1278.
9. Hur, Y., E.A. Unger, and A.C. Vasconcelos, *Isolation and characterization of a cDNA encoding cytosolic fructose-1,6-bisphosphatase from spinach*. Plant Mol Biol, 1992. **18**(4): p. 799-802.
10. Khayat, E., C. Harn, and J. Daie, *Purification and Light-Dependent Molecular Modulation of the Cytosolic Fructose-1,6-Bisphosphatase in Sugarbeet Leaves*. Plant Physiol, 1993. **101**(1): p. 57-64.

11. Strand, A., R. Zrenner, S. Trevanion, M. Stitt, P. Gustafsson, and P. Gardestrom, *Decreased expression of two key enzymes in the sucrose biosynthesis pathway, cytosolic fructose-1,6-bisphosphatase and sucrose phosphate synthase, has remarkably different consequences for photosynthetic carbon metabolism in transgenic Arabidopsis thaliana*. Plant J, 2000. **23**(6): p. 759-70.
12. Jang, H.K., S.W. Lee, Y.H. Lee, and T.R. Hahn, *Purification and characterization of a recombinant pea cytoplasmic fructose-1,6-bisphosphatase*. Protein Expr Purif, 2003. **28**(1): p. 42-48.
13. Xu W, W.C., Xu X, Cai C, and G. W, *Cloning and expression analysis of a novel gene encoding fructose-1,6-bisphosphatase in cotton*. Cotton Science, 2013. **25**: p. 549-556.
14. Xiao, H., C. Chen, Y. Xu, D. Ji, and C. Xie, *Cloning and expression analysis of the chloroplast fructose-1,6-bisphosphatase gene from Pyropia haitanensis*. Acta Oceanologica Sinica, 2014. **33**(4): p. 92-100.
15. Cho, M.H., A. Jang, S.H. Bhoo, J.S. Jeon, and T.R. Hahn, *Manipulation of triose phosphate/phosphate translocator and cytosolic fructose-1,6-bisphosphatase, the key components in photosynthetic sucrose synthesis, enhances the source capacity of transgenic Arabidopsis plants*. Photosynth Res, 2012. **111**(3): p. 261-268.
16. Zrenner, R., K.P. Krause, P. Apel, and U. Sonnewald, *Reduction of the cytosolic fructose-1,6-bisphosphatase in transgenic potato plants limits photosynthetic sucrose biosynthesis with no impact on plant growth and tuber yield*. Plant J, 1996. **9**(5): p. 671-681.
17. Lee, S.K., J.S. Jeon, F. Bornke, L. Voll, J.I. Cho, C.H. Goh, S.W. Jeong, Y.I. Park, S.J. Kim, S.B. Choi, A. Miyao, H. Hirochika, G. An, M.H. Cho, S.H. Bhoo, U. Sonnewald, and T.R. Hahn, *Loss of cytosolic fructose-1,6-bisphosphatase limits photosynthetic sucrose synthesis and causes severe growth retardations in rice (Oryza sativa)*. Plant Cell Environ, 2008. **31**(12): p. 1851-1863.
18. Obiadalla-Ali, H., A.R. Fernie, A. Lytovchenko, J. Kossmann, and J.R. Lloyd, *Inhibition of chloroplastic fructose 1,6-bisphosphatase in tomato fruits leads to decreased fruit size, but only small changes in carbohydrate metabolism*. Planta, 2004. **219**(3): p. 533-540.
19. BASRA A S, M.C.P., *Development of cotton fiber[M]*. International Review Cytology, 1984. **89**: p. 65-113.
20. Zhang, Z., J. Li, M. Jamshed, Y. Shi, A. Liu, J. Gong, S. Wang, J. Zhang, F. Sun, F. Jia, Q. Ge, L. Fan, Z. Zhang, J. Pan, S. Fan, Y. Wang, Q. Lu, R. Liu, X. Deng, X. Zou, X. Jiang, P. Liu, P. Li, M.S. Iqbal, C. Zhang, J. Zou, H. Chen, Q. Tian, X. Jia, B. Wang, N. Ai, G. Feng, Y. Wang, M. Hong, S. Li, W. Lian, B. Wu, J. Hua, C. Zhang, J. Huang, A. Xu, H. Shang, W. Gong, and Y. Yuan, *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, 2019.
21. Wang, F., J. Zhang, Y. Chen, C. Zhang, J. Gong, Z. Song, J. Zhou, J. Wang, C. Zhao, M. Jiao, A. Liu, Z. Du, Y. Yuan, S. Fan, and J. Zhang, *Identification of candidate genes for key fibre-related QTLs and derivation of favourable alleles in Gossypium hirsutum recombinant inbred lines with G. barbadense introgressions*. Plant Biotechnol J, 2019.

22. Deng, X., J. Gong, A. Liu, Y. Shi, W. Gong, Q. Ge, J. Li, H. Shang, Y. Wu, and Y. Yuan, *QTL mapping for fiber quality and yield-related traits across multiple generations in segregating population of CCRI 70*. Journal of Cotton Research, 2019. **2**(1).
23. Du, X., G. Huang, S. He, Z. Yang, G. Sun, X. Ma, N. Li, X. Zhang, J. Sun, M. Liu, Y. Jia, Z. Pan, W. Gong, Z. Liu, H. Zhu, L. Ma, F. Liu, D. Yang, F. Wang, W. Fan, Q. Gong, Z. Peng, L. Wang, X. Wang, S. Xu, H. Shang, C. Lu, H. Zheng, S. Huang, T. Lin, Y. Zhu, and F. Li, *Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits*. Nat Genet, 2018. **50**(6): p. 796-802.
24. Ma, Z., S. He, X. Wang, J. Sun, Y. Zhang, G. Zhang, L. Wu, Z. Li, Z. Liu, G. Sun, Y. Yan, Y. Jia, J. Yang, Z. Pan, Q. Gu, X. Li, Z. Sun, P. Dai, Z. Liu, W. Gong, J. Wu, M. Wang, H. Liu, K. Feng, H. Ke, J. Wang, H. Lan, G. Wang, J. Peng, N. Wang, L. Wang, B. Pang, Z. Peng, R. Li, S. Tian, and X. Du, *Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield*. Nat Genet, 2018. **50**(6): p. 803-813.
25. Su, J., L. Li, C. Zhang, C. Wang, L. Gu, H. Wang, H. Wei, Q. Liu, L. Huang, and S. Yu, *Genome-wide association study identified genetic variations and candidate genes for plant architecture component traits in Chinese upland cotton*. Theor Appl Genet, 2018. **131**(6): p. 1299-1314.
26. Sun, Z., X. Wang, Z. Liu, Q. Gu, Y. Zhang, Z. Li, H. Ke, J. Yang, J. Wu, L. Wu, G. Zhang, C. Zhang, and Z. Ma, *A genome-wide association study uncovers novel genomic regions and candidate genes of yield-related traits in upland cotton*. Theor Appl Genet, 2018. **131**(11): p. 2413-2425.
27. Li, Y., N.N. Wang, Y. Wang, D. Liu, Y. Gao, L. Li, and X.B. Li, *The cotton XLIM protein (GhXLIM6) is required for fiber development via maintaining dynamic F-actin cytoskeleton and modulating cellulose biosynthesis*. Plant J, 2018. **96**(6): p. 1269-1282.
28. Zhang, J., G.Q. Huang, D. Zou, J.Q. Yan, Y. Li, S. Hu, and X.B. Li, *The cotton (Gossypium hirsutum) NAC transcription factor (FSN1) as a positive regulator participates in controlling secondary cell wall biosynthesis and modification of fibers*. New Phytol, 2018. **217**(2): p. 625-640.
29. Zhao, B., J.F. Cao, G.J. Hu, Z.W. Chen, L.Y. Wang, X.X. Shangguan, L.J. Wang, Y.B. Mao, T.Z. Zhang, J.F. Wendel, and X.Y. Chen, *Core cis-element variation confers subgenome-biased expression of a transcription factor that functions in cotton fiber elongation*. New Phytol, 2018. **218**(3): p. 1061-1075.
30. M. Ashraf, S.A., *Influence of sodium chloride on ion accumulation, yield components and fiber characteristics in salt-tolerant and salt-sensitive lines of cotton (Gossypium hirsutum L.)*. Field Crops Research 2000. **66** p. 115-127.
31. Abdelraheem, A., N. Esmaili, M. O'Connell, and J. Zhang, *Progress and perspective on drought and salt stress tolerance in cotton*. Industrial Crops and Products, 2019. **130**: p. 118-129.
32. Paterson, A.H., J.F. Wendel, H. Gundlach, H. Guo, J. Jenkins, D. Jin, D. Llewellyn, K.C. Showmaker, S. Shu, J. Udall, M.J. Yoo, R. Byers, W. Chen, A. Doron-Faigenboim, M.V. Duke, L. Gong, J. Grimwood, C. Grover, K. Grupp, G. Hu, T.H. Lee, J. Li, L. Lin, T. Liu, B.S. Marler, J.T. Page, A.W. Roberts, E. Romanel, W.S. Sanders, E. Szadkowski, X. Tan, H. Tang, C. Xu, J. Wang, Z. Wang, D. Zhang, L. Zhang, H. Ashrafi, F. Bedon, J.E. Bowers, C.L. Brubaker, P.W. Chee, S. Das, A.R. Gingle, C.H. Haigler, D. Harker, L.V.

- Hoffmann, R. Hovav, D.C. Jones, C. Lemke, S. Mansoor, M. ur Rahman, L.N. Rainville, A. Rambani, U.K. Reddy, J.K. Rong, Y. Saranga, B.E. Scheffler, J.A. Scheffler, D.M. Stelly, B.A. Triplett, A. Van Deynze, M.F. Vaslin, V.N. Waghamare, S.A. Walford, R.J. Wright, E.A. Zaki, T. Zhang, E.S. Dennis, K.F. Mayer, D.G. Peterson, D.S. Rokhsar, X. Wang, and J. Schmutz, *Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres*. Nature, 2012. **492**(7429): p. 423-427.
33. Wang, K., Z. Wang, F. Li, W. Ye, J. Wang, G. Song, Z. Yue, L. Cong, H. Shang, S. Zhu, C. Zou, Q. Li, Y. Yuan, C. Lu, H. Wei, C. Gou, Z. Zheng, Y. Yin, X. Zhang, K. Liu, B. Wang, C. Song, N. Shi, R.J. Kohel, R.G. Percy, J.Z. Yu, Y.X. Zhu, J. Wang, and S. Yu, *The draft genome of a diploid cotton Gossypium raimondii*. Nat Genet, 2012. **44**(10): p. 1098-1103.
34. Li, F., G. Fan, K. Wang, F. Sun, Y. Yuan, G. Song, Q. Li, Z. Ma, C. Lu, C. Zou, W. Chen, X. Liang, H. Shang, W. Liu, C. Shi, G. Xiao, C. Gou, W. Ye, X. Xu, X. Zhang, H. Wei, Z. Li, G. Zhang, J. Wang, K. Liu, R.J. Kohel, R.G. Percy, J.Z. Yu, Y.X. Zhu, J. Wang, and S. Yu, *Genome sequence of the cultivated cotton Gossypium arboreum*. Nat Genet, 2014. **46**(6): p. 567-572.
35. Li, F., G. Fan, C. Lu, G. Xiao, C. Zou, R.J. Kohel, Z. Ma, H. Shang, X. Ma, J. Wu, X. Liang, G. Huang, R.G. Percy, K. Liu, W. Yang, W. Chen, X. Du, C. Shi, Y. Yuan, W. Ye, X. Liu, X. Zhang, W. Liu, H. Wei, S. Wei, G. Huang, X. Zhang, S. Zhu, H. Zhang, F. Sun, X. Wang, J. Liang, J. Wang, Q. He, L. Huang, J. Wang, J. Cui, G. Song, K. Wang, X. Xu, J.Z. Yu, Y. Zhu, and S. Yu, *Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution*. Nat Biotechnol, 2015. **33**(5): p. 524-530.
36. Zhang, T., Y. Hu, W. Jiang, L. Fang, X. Guan, J. Chen, J. Zhang, C.A. Sasaki, B.E. Scheffler, D.M. Stelly, A.M. Hulse-Kemp, Q. Wan, B. Liu, C. Liu, S. Wang, M. Pan, Y. Wang, D. Wang, W. Ye, L. Chang, W. Zhang, Q. Song, R.C. Kirkbride, X. Chen, E. Dennis, D.J. Llewellyn, D.G. Peterson, P. Thaxton, D.C. Jones, Q. Wang, X. Xu, H. Zhang, H. Wu, L. Zhou, G. Mei, S. Chen, Y. Tian, D. Xiang, X. Li, J. Ding, Q. Zuo, L. Tao, Y. Liu, J. Li, Y. Lin, Y. Hui, Z. Cao, C. Cai, X. Zhu, Z. Jiang, B. Zhou, W. Guo, R. Li, and Z.J. Chen, *Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement*. Nat Biotechnol, 2015. **33**(5): p. 531-537.
37. Wang, M., P. Wang, M. Lin, Z. Ye, G. Li, L. Tu, C. Shen, J. Li, Q. Yang, and X. Zhang, *Evolutionary dynamics of 3D genome architecture following polyploidization in cotton*. Nat Plants, 2018. **4**(2): p. 90-97.
38. Hu, Y., J. Chen, L. Fang, Z. Zhang, W. Ma, Y. Niu, L. Ju, J. Deng, T. Zhao, J. Lian, K. Baruch, D. Fang, X. Liu, Y.L. Ruan, M.U. Rahman, J. Han, K. Wang, Q. Wang, H. Wu, G. Mei, Y. Zang, Z. Han, C. Xu, W. Shen, D. Yang, Z. Si, F. Dai, L. Zou, F. Huang, Y. Bai, Y. Zhang, A. Brodt, H. Ben-Hamo, X. Zhu, B. Zhou, X. Guan, S. Zhu, X. Chen, and T. Zhang, *Gossypium barbadense and Gossypium hirsutum genomes provide insights into the origin and evolution of allotetraploid cotton*. Nat Genet, 2019. **51**(4): p. 739-748.
39. Liu, X., B. Zhao, H.J. Zheng, Y. Hu, G. Lu, C.Q. Yang, J.D. Chen, J.J. Chen, D.Y. Chen, L. Zhang, Y. Zhou, L.J. Wang, W.Z. Guo, Y.L. Bai, J.X. Ruan, X.X. Shangguan, Y.B. Mao, C.M. Shan, J.P. Jiang, Y.Q. Zhu, L. Jin, H. Kang, S.T. Chen, X.L. He, R. Wang, Y.Z. Wang, J. Chen, L.J. Wang, S.T. Yu, B.Y. Wang, J. Wei,

- S.C. Song, X.Y. Lu, Z.C. Gao, W.Y. Gu, X. Deng, D. Ma, S. Wang, W.H. Liang, L. Fang, C.P. Cai, X.F. Zhu, B.L. Zhou, Z. Jeffrey Chen, S.H. Xu, Y.G. Zhang, S.Y. Wang, T.Z. Zhang, G.P. Zhao, and X.Y. Chen, *Gossypium barbadense genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites*. Sci Rep, 2015. **5**: p. 14139.
40. WENDEL, J.F., *New World tetraploid cottons contain Old World cytoplasm*. Proc. Nati. Acad. Sci, 1989. **Vol. 86**, pp. **4132-4136**.
41. Cannon, S.B., A. Mitra, A. Baumgarten, N.D. Young, and G. May, *The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana*. BMC Plant Biol, 2004. **4**: p. 10.
42. Schaper, E. and M. Anisimova, *The evolution and function of protein tandem repeats in plants*. New Phytol, 2015. **206**(1): p. 397-410.
43. Hurst, D., *The Ka\_Ks ratio\_diagnosing the form of sequence evolution*. TRENDS in Genetics 2002. **18**(9).
44. Xu, Y., R.O. Magwanga, X. Cai, Z. Zhou, X. Wang, Y. Wang, Z. Zhang, D. Jin, X. Guo, Y. Wei, Z. Li, K. Wang, and F. Liu, *Deep Transcriptome Analysis Reveals Reactive Oxygen Species (ROS) Network Evolution, Response to Abiotic Stress, and Regulation of Fiber Development in Cotton*. Int J Mol Sci, 2019. **20**(8).
45. Chueca, A., M. Sahrawy, E.A. Pagano, and J. Lopez Gorge, *Chloroplast fructose-1,6-bisphosphatase: structure and function*. Photosynth Res, 2002. **74**(3): p. 235-249.
46. Rojas-Gonzalez, J.A., M. Soto-Suarez, A. Garcia-Diaz, M.C. Romero-Puertas, L.M. Sandalio, A. Merida, I. Thormahlen, P. Geigenberger, A.J. Serrato, and M. Sahrawy, *Disruption of both chloroplastic and cytosolic FBPase genes results in a dwarf phenotype and important starch and metabolite changes in Arabidopsis thaliana*. J Exp Bot, 2015. **66**(9): p. 2673-2689.
47. Guo, H., Y. Jiao, X. Tan, X. Wang, X. Huang, H. Jin, and A.H. Paterson, *Gene duplication and genetic innovation in cereal genomes*. Genome Res, 2019. **29**(2): p. 261-269.
48. Lan, X. and J.K. Pritchard, *Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals*. Science, 2016. **352**(6288): p. 1009-1013.
49. Liu, S., Y. Liu, X. Yang, C. Tong, D. Edwards, I.A. Parkin, M. Zhao, J. Ma, J. Yu, S. Huang, X. Wang, J. Wang, K. Lu, Z. Fang, I. Bancroft, T.J. Yang, Q. Hu, X. Wang, Z. Yue, H. Li, L. Yang, J. Wu, Q. Zhou, W. Wang, G.J. King, J.C. Pires, C. Lu, Z. Wu, P. Sampath, Z. Wang, H. Guo, S. Pan, L. Yang, J. Min, D. Zhang, D. Jin, W. Li, H. Belcram, J. Tu, M. Guan, C. Qi, D. Du, J. Li, L. Jiang, J. Batley, A.G. Sharpe, B.S. Park, P. Ruperao, F. Cheng, N.E. Waminal, Y. Huang, C. Dong, L. Wang, J. Li, Z. Hu, M. Zhuang, Y. Huang, J. Huang, J. Shi, D. Mei, J. Liu, T.H. Lee, J. Wang, H. Jin, Z. Li, X. Li, J. Zhang, L. Xiao, Y. Zhou, Z. Liu, X. Liu, R. Qin, X. Tang, W. Liu, Y. Wang, Y. Zhang, J. Lee, H.H. Kim, F. Denoeud, X. Xu, X. Liang, W. Hua, X. Wang, J. Wang, B. Chalhoub, and A.H. Paterson, *The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes*. Nat Commun, 2014. **5**: p. 3930.
50. JF W, A.S., T S,, *An Unusual Ribosomal DNA Sequence from Gossypium gossypioides Reveals Ancient, Cryptic, Intergenomic Introgression*. Mol. Phylogenet. Evol., 1995. **4**(3): p. 298-313.

51. JF, W., *New World tetraploid cottons contain Old World cytoplasm*. Proc Natl Acad Sci, 1989. **86**: p. 4132-4136.
52. Fan, K., F. Li, J. Chen, Z. Li, W. Lin, S. Cai, J. Liu, and W. Lin, *Asymmetric Evolution and Expansion of the NAC Transcription Factor in Polyploidized Cotton*. Front Plant Sci, 2018. **9**: p. 47.
53. Fang, L., X. Guan, and T. Zhang, *Asymmetric evolution and domestication in allotetraploid cotton (*Gossypium hirsutum* L.)*. The Crop Journal, 2017. **5**(2): p. 159-165.
54. Wang, X., H. Guo, J. Wang, T. Lei, T. Liu, Z. Wang, Y. Li, T.H. Lee, J. Li, H. Tang, D. Jin, and A.H. Paterson, *Comparative genomic de-convolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation*. New Phytol, 2016. **209**(3): p. 1252-1263.
55. Kim, H.J. and B.A. Triplett, *Cotton Fiber Growth in Planta and in Vitro. Models for Plant Cell Elongation and Cell Wall Biogenesis*. Plant Physiology, 2001. **127**(4): p. 1361-1366.
56. Lee, J.J., A.W. Woodward, and Z.J. Chen, *Gene expression changes and early events in cotton fibre development*. Ann Bot, 2007. **100**(7): p. 1391-1401.
57. Guan, X., Q. Song, and Z.J. Chen, *Polyploidy and small RNA regulation of cotton fiber development*. Trends Plant Sci, 2014. **19**(8): p. 516-528.
58. Beasley, C.A., *Hormonal regulation of growth in unfertilized cotton ovules*. Science, 1973. **179**(4077): p. 1003-1005.
59. Suo, J.F., X.O. Liang, L. Pu, Y.S. Zhang, and Y.B. Xue, *Identification of GhMYB109 encoding a R2R3 MYB transcription factor that expressed specifically in fiber initials and elongating fibers of cotton (*Gossypium Hirsutum* L.)*. Biochimica Et Biophysica Acta-Genes and Expression, 2003. **1630**(1): p. 25-34.
60. Luo, M., Y.H. Xiao, X.B. Li, X.F. Lu, W. Deng, D. Li, L. Hou, M.Y. Hu, Y. Li, and Y. Pei, *GhDET2, a steroid 5 alpha-reductase, plays an important role in cotton fiber cell initiation and elongation*. Plant Journal, 2007. **51**(3): p. 419-430.
61. Sun, W., Z. Gao, J. Wang, Y. Huang, Y. Chen, J. Li, M. Lv, J. Wang, M. Luo, and K. Zuo, *Cotton fiber elongation requires the transcription factor GhMYB212 to regulate sucrose transportation into expanding fibers*. New Phytol, 2019. **222**(2): p. 864-881.
62. Xu, W.L., D.J. Zhang, Y.F. Wu, L.X. Qin, G.Q. Huang, J. Li, L. Li, and X.B. Li, *Cotton PRP5 gene encoding a proline-rich protein is involved in fiber development*. Plant Molecular Biology, 2013. **82**(4-5): p. 353-365.
63. Wang, J.A., H.Y. Wang, P.M. Zhao, L.B. Han, G.L. Jiao, Y.Y. Zheng, S.J. Huang, and G.X. Xia, *Overexpression of a Profilin (*GhPFN2*) Promotes the Progression of Developmental Phases in Cotton Fibers*. Plant and Cell Physiology, 2010. **51**(8): p. 1276-1290.
64. Cheng Wang, Y.L., Wentin Xu, Tianzhen Zhang and Wangzhen Guo, *Aberrant phenotype and transcriptome expression during fiber cell wall thickening caused by the mutation of the Im gene in immature fiber (im) mutant in *Gossypium hirsutum* L.* BMC Genomics, 2014. **15**.
65. Wang, L. and Y.L. Ruan, *Unraveling mechanisms of cell expansion linking solute transport, metabolism, plasmodesmal gating and cell wall dynamics*. Plant Signal Behav, 2010. **5**(12): p. 1561-

1564.

66. Zhang, Z., Y.L. Ruan, N. Zhou, F. Wang, X. Guan, L. Fang, X. Shang, W. Guo, S. Zhu, and T. Zhang, *Suppressing a Putative Sterol Carrier Gene Reduces Plasmodesmal Permeability and Activates Sucrose Transporter Genes during Cotton Fiber Elongation*. *Plant Cell*, 2017. **29**(8): p. 2027-2046.
67. Zhang, J., S. Sanogo, R. Flynn, J.B. Baral, S. Bajaj, S.E. Hughs, and R.G. Percy, *Germplasm evaluation and transfer of Verticillium wilt resistance from Pima (Gossypium barbadense) to Upland cotton (G. hirsutum)*. *Euphytica*, 2011. **187**(2): p. 147-160.
68. Ouriel Faktor, G.L., Richard A. Dixon, Chris J. Lamb, *The G-box and H-box in a 39 bp region of a French bean*. *The Plant Journal*, 1997. **11**: p. 1105-1113.
69. Rushton, P.J., A. Reinstadler, V. Lipka, B. Lippok, and I.E. Somssich, *Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling*. *Plant Cell*, 2002. **14**(4): p. 749-762.
70. Brown, R.L., K. Kazan, K.C. McGrath, D.J. Maclean, and J.M. Manners, *A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of Arabidopsis*. *Plant Physiol*, 2003. **132**(2): p. 1020-1032.
71. Park, H.C., M.L. Kim, Y.H. Kang, J.M. Jeon, J.H. Yoo, M.C. Kim, C.Y. Park, J.C. Jeong, B.C. Moon, J.H. Lee, H.W. Yoon, S.H. Lee, W.S. Chung, C.O. Lim, S.Y. Lee, J.C. Hong, and M.J. Cho, *Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor*. *Plant Physiol*, 2004. **135**(4): p. 2150-2161.
72. DeVay, J.E., Weir, B. L., Wakeman, R. J., and Stapleton, J. J. , *Effects of Verticillium dahliae Infection of Cotton Plants (Gossypium hirsutum) on Potassium Levels in Leaf Petioles*. *Plant Dis.*, 1997. **81**: p. 1089-1092.
73. Fang W, Z.S., Ji D., *Studies on the Inheritance of Verticillium dahliae Resistance in G. hisutum and G. barbadense*. *Cotton Science*, 2003. **15**: p. 3- 7.
74. Huckelhoven, R., *Cell wall-associated mechanisms of disease resistance and susceptibility*. *Annu Rev Phytopathol*, 2007. **45**: p. 101-127.
75. Zhang, Y., L. Wu, X. Wang, B. Chen, J. Zhao, J. Cui, Z. Li, J. Yang, L. Wu, J. Wu, G. Zhang, and Z. Ma, *The cotton laccase gene GhLAC15 enhances Verticillium wilt resistance via an increase in defence-induced lignification and lignin components in the cell walls of plants*. *Mol Plant Pathol*, 2019. **20**(3): p. 309-322.
76. Xu, L., L. Zhu, L. Tu, L. Liu, D. Yuan, L. Jin, L. Long, and X. Zhang, *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*. *J Exp Bot*, 2011. **62**(15): p. 5607-5621.
77. Guo, J. and M.H. Wang, *Characterization of the phenylalanine ammonia-lyase gene (SIPAL5) from tomato (Solanum lycopersicum L.)*. *Mol Biol Rep*, 2009. **36**(6): p. 1579-1585.
78. Hu, Q., L. Min, X. Yang, S. Jin, L. Zhang, Y. Li, Y. Ma, X. Qi, D. Li, H. Liu, K. Lindsey, L. Zhu, and X. Zhang, *Laccase GhLac1 Modulates Broad-Spectrum Biotic Stress Tolerance via Manipulating Phenylpropanoid Pathway and Jasmonic Acid Synthesis*. *Plant Physiol*, 2018. **176**(2): p. 1808-1823.

79. Rogers, L.A., C. Dubos, I.F. Cullis, C. Surman, M. Poole, J. Willment, S.D. Mansfield, and M.M. Campbell, *Light, the circadian clock, and sugar perception in the control of lignin biosynthesis*. J Exp Bot, 2005. **56**(416): p. 1651-1563.
80. Bose, J., R. Munns, S. Shabala, M. Gilliam, B. Pogson, and S.D. Tyerman, *Chloroplast function and ion regulation in plants growing on saline soils: lessons from halophytes*. J Exp Bot, 2017. **68**(12): p. 3129-3143.
81. Ghosh, S., S. Bagchi, and A. Lahiri Majumder, *Chloroplast fructose-1,6-bisphosphatase from Oryza differs in salt tolerance property from the Porteresia enzyme and is protected by osmolytes*. Plant Sci, 2001. **160**(6): p. 1171-1181.
82. Chatterjee, J., B. Patra, R. Mukherjee, P. Basak, S. Mukherjee, S. Ray, S. Bhattacharyya, S. Maitra, K. GhoshDastidar, S. Ghosh, S. Sengupta, and A.L. Majumder, *Cloning, characterization and expression of a chloroplastic fructose-1,6-bisphosphatase from Porteresia coarctata conferring salt-tolerance in transgenic tobacco*. Plant Cell, Tissue and Organ Culture (PCTOC), 2013. **114**(3): p. 395-409.
83. Harn, C.H., Jaleh Daie I and Jang Ryol Liu, *Regulation of Cytosolic Fructose-1,6-Bisphosphatase Under Water-Stressed Leaves of Sugar Beet: Protein Modification is not a Mechanism for Coarse Control*. J. Plant Biol, 1997. **40** (4) p. 261-266.
84. Saeed, A.I., V.S. , and J.W.e.a. , *TM4\_ A Free, Open-Source System for Microarray Data Management and Analysis*. BioTechniques, 2003. **34**: p. 374-378.
85. Li, P.T., M.H.O. Rashid, T.T. Chen, Q.W. Lu, Q. Ge, W.K. Gong, A.Y. Liu, J.W. Gong, H.H. Shang, X.Y. Deng, J.W. Li, S.Q. Li, X.H. Xiao, R.X. Liu, Q. Zhang, L. Duan, X.Y. Zou, Z. Zhang, X. Jiang, Y. Zhang, R.H. Peng, Y.Z. Shi, and Y.L. Yuan, *Transcriptomic and biochemical analysis of upland cotton (Gossypium hirsutum) and a chromosome segment substitution line from G. hirsutum x G. barbadense in response to Verticillium dahliae infection*. BMC Plant Biol, 2019. **19**(1): p. 19.
86. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-408.
87. Zhu, T., C. Liang, Z. Meng, G. Sun, Z. Meng, S. Guo, and R. Zhang, *CottonFGD: an integrated functional genomics database for cotton*. BMC Plant Biol, 2017. **17**(1): p. 101.
88. Finn, R.D., J. Clements, and S.R. Eddy, *HMMER web server: interactive sequence similarity searching*. Nucleic Acids Res, 2011. **39**(Web Server issue): p. 29-37.
89. Julie D. Thompson, T.J.G., and Des G. Higgins, *Multiple Sequence Alignment Using ClustalW and ClustalX*. Current Protocols in Bioinformatics, 2003. **2**: p. 1-22.
90. Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar, *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0*. Mol Biol Evol, 2013. **30**(12): p. 2725-9.
91. Hu, B., J. Jin, A.Y. Guo, H. Zhang, J. Luo, and G. Gao, *GSDS 2.0: an upgraded gene feature visualization server*. Bioinformatics, 2015. **31**(8): p. 1296-1297.
92. Wang, Y., H. Tang, J.D. Debarry, X. Tan, J. Li, X. Wang, T.H. Lee, H. Jin, B. Marler, H. Guo, J.C. Kissinger, and A.H. Paterson, *MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity*. Nucleic Acids Res, 2012. **40**(7): p. e49.

93. Lescot, M., P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouze, and S. Rombauts, *PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences*. *Nucleic Acids Res*, 2002. **30**(1): p. 325-7.

## Supplementary Materials

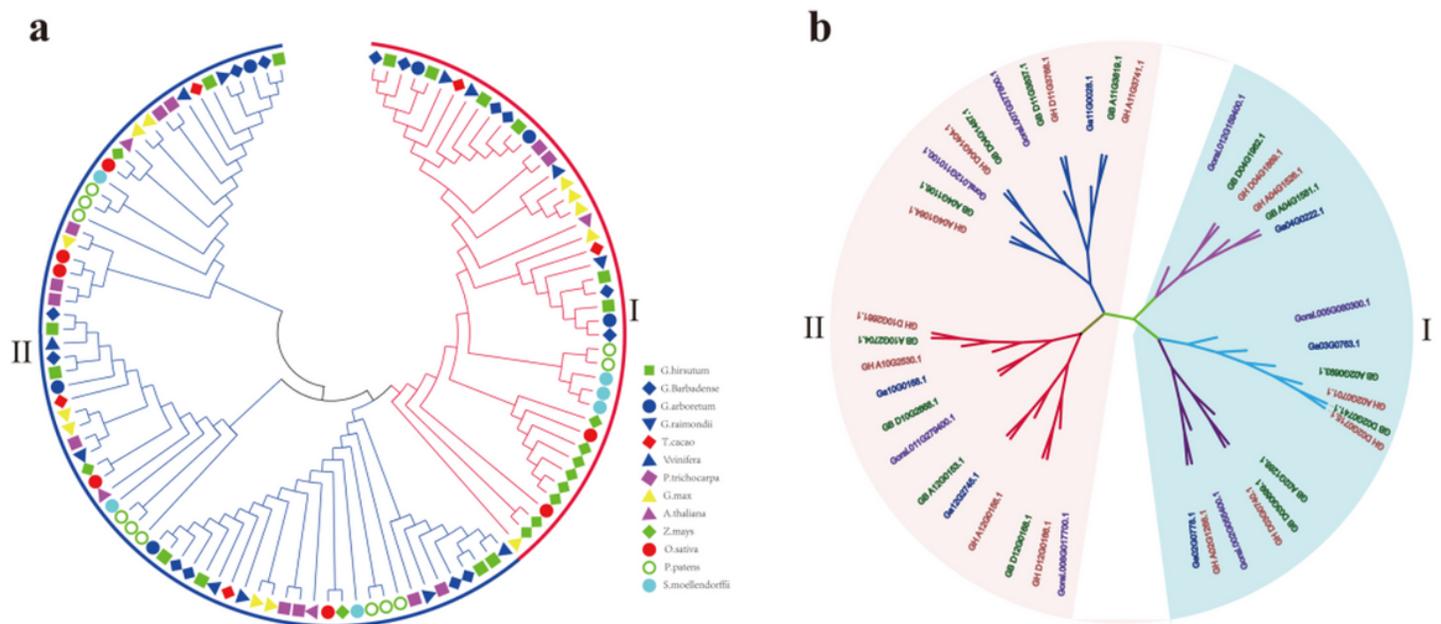
The following are available online at [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Supplementary file 1: Multiple sequence alignment involving 114 sequences from the genomes of the following eight species: *G. hirsutum*, *G. arboreum*, *G. barbadense*, *G. raimondii*, *A. thaliana*, *T. cacao*, *P. trichocarpa*, *G. max*, *Z. mays*, *V. vinifera*, *S. moellendorffii*, *P. patens*, and *O. sativa*.

Supplementary file 2: Amino acid compositions of *FBP* genes.

Supplementary file 3: Ka/Ks ratio of *FBP* homologous genes; Supplementary file 4: The FPKM value of *FBP* genes in tetraploid cotton species.

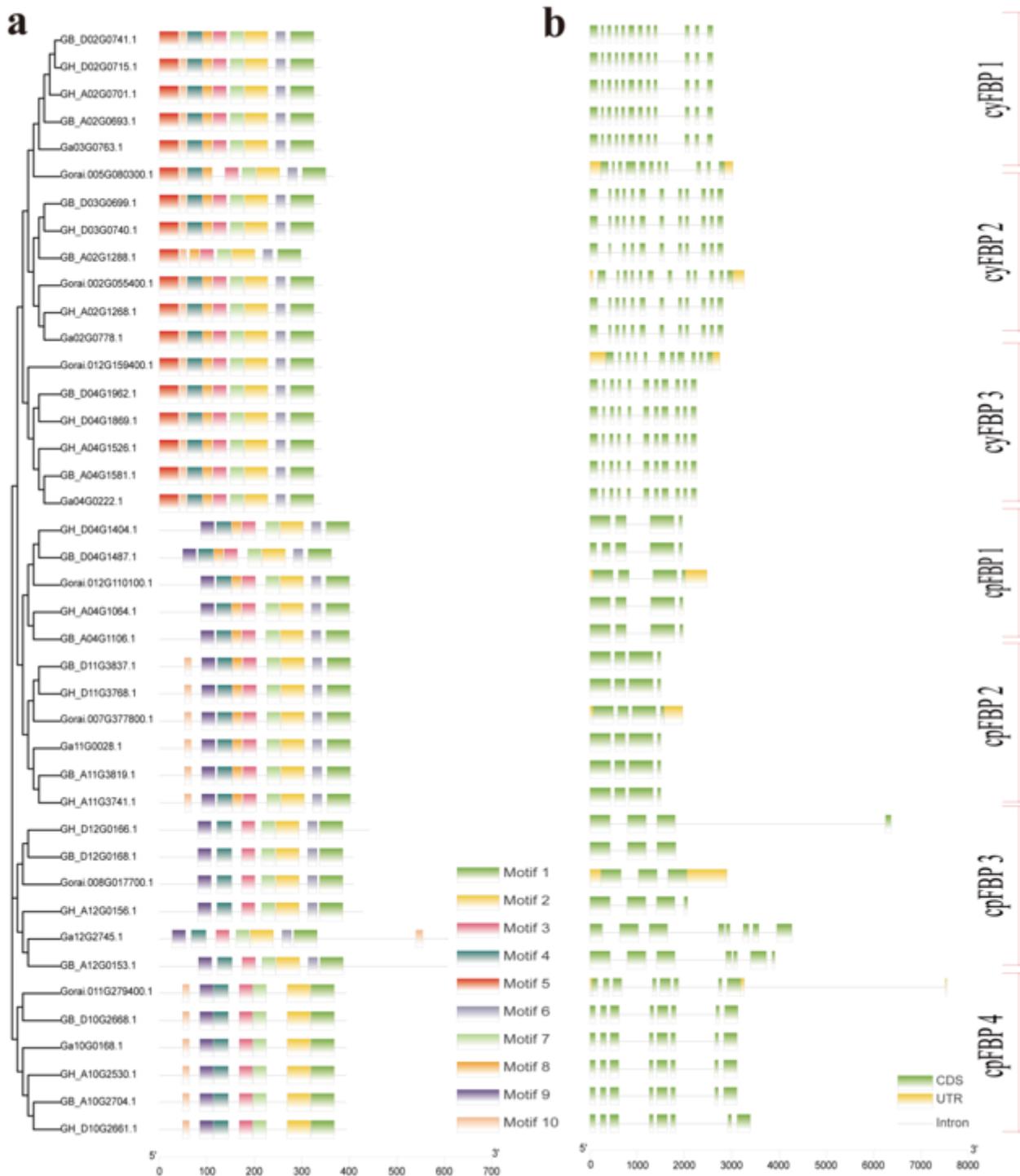
Figure S1. Logos of 10 motifs for four cotton species according to the MEME suite.

## Figures



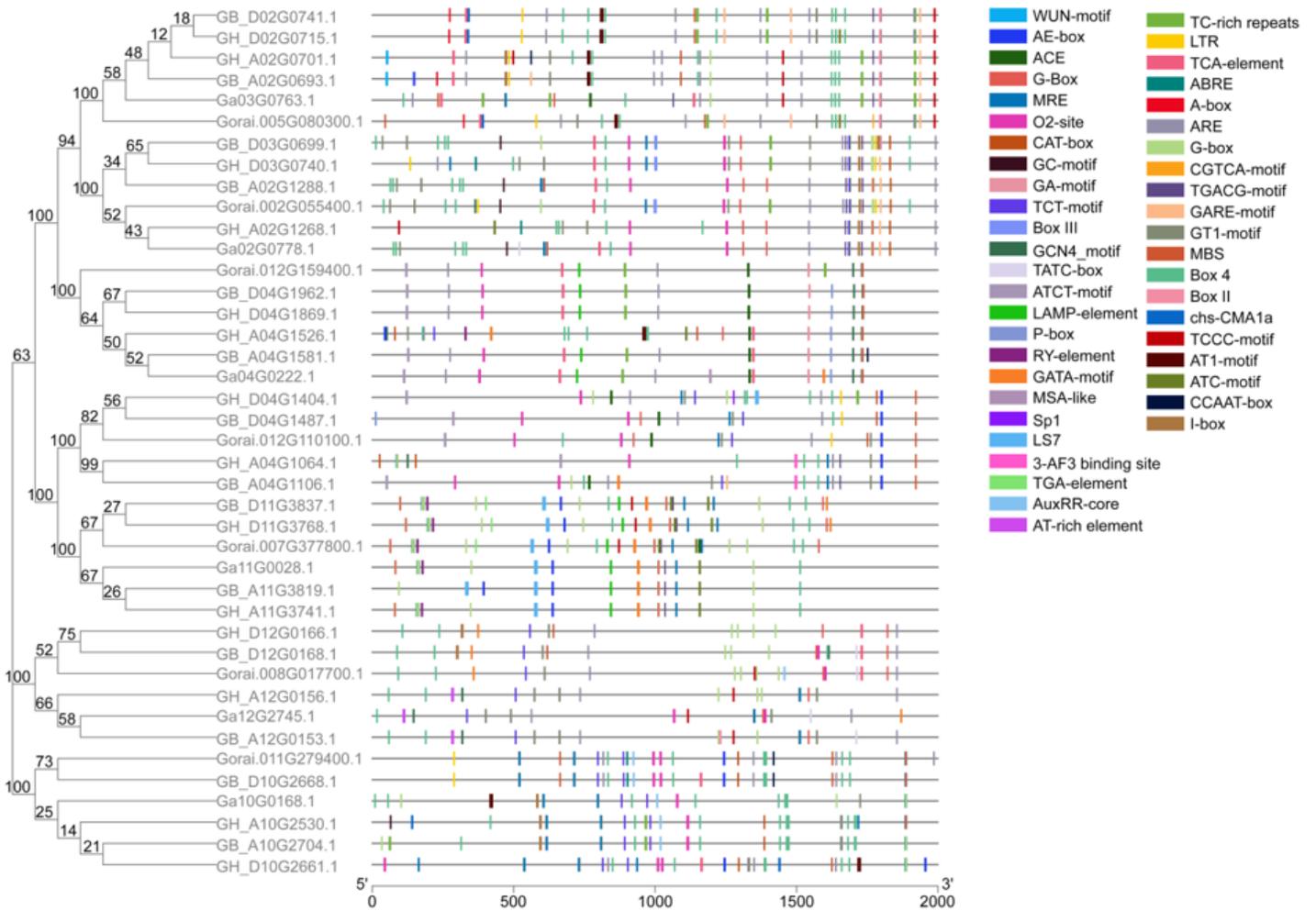
**Figure 1**

Phylogenetic trees of FBPs. (a) Phylogenetic tree of 114 FBPs from 13 species, including *G. hirsutum*, *G. barbadense*, *G. arboreum*, *G. raimondii*, *A. thaliana*, *T. cacao*, *P. trichocarpa*, *G. max*, *Z. mays*, *V. vinifera*, *S. moellendorffii*, *P. patens* and *O. sativa*; (b) Phylogenetic tree of 41 FBPs from four *Gossypium* species. I represent cyFBPs and II represent cpFBPs.



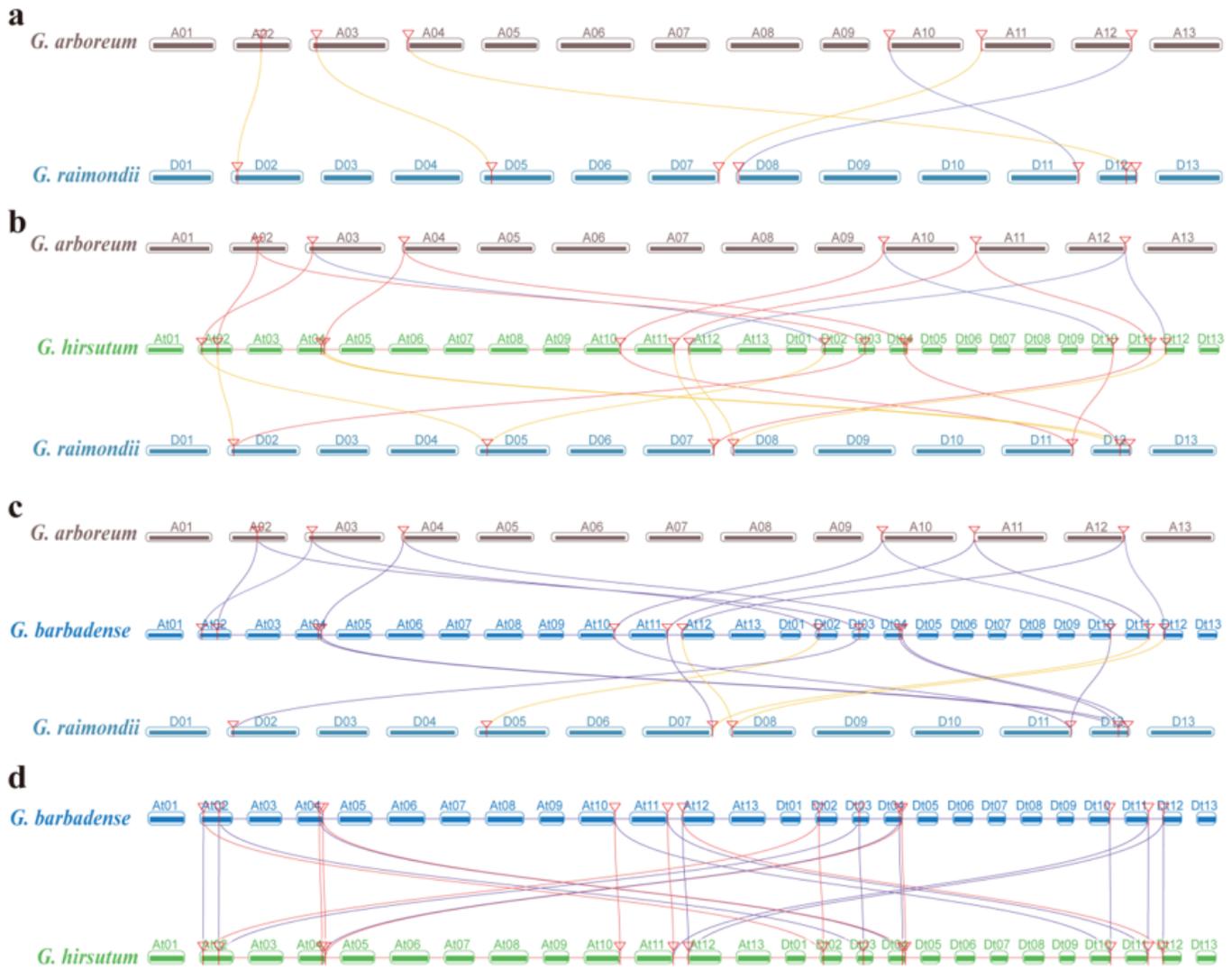
**Figure 2**

Phylogenetic relationship, motifs and gene structure of FBP members in four cottons.



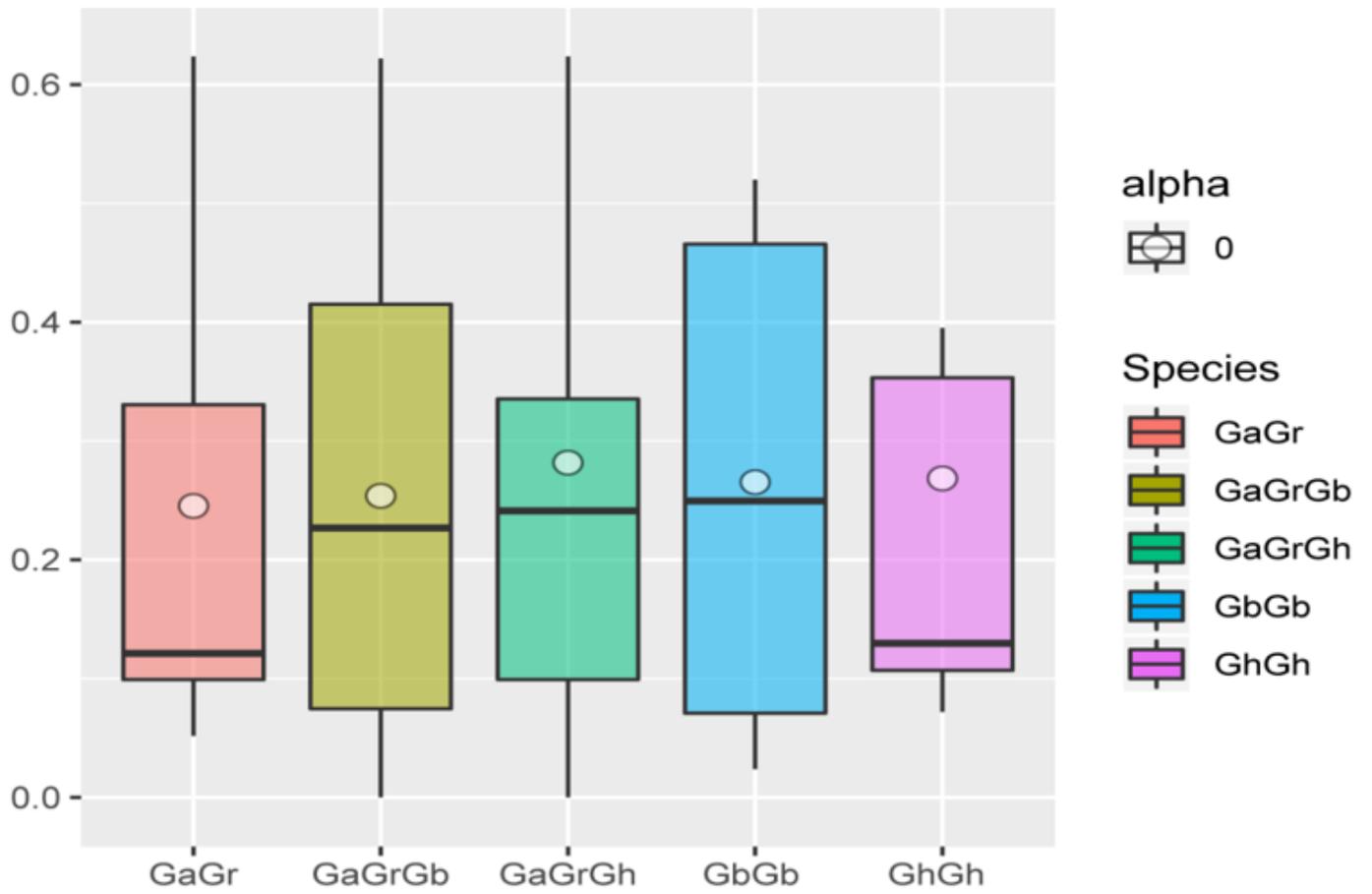
**Figure 3**

Cis-acting elements analysis of cotton FBP genes.



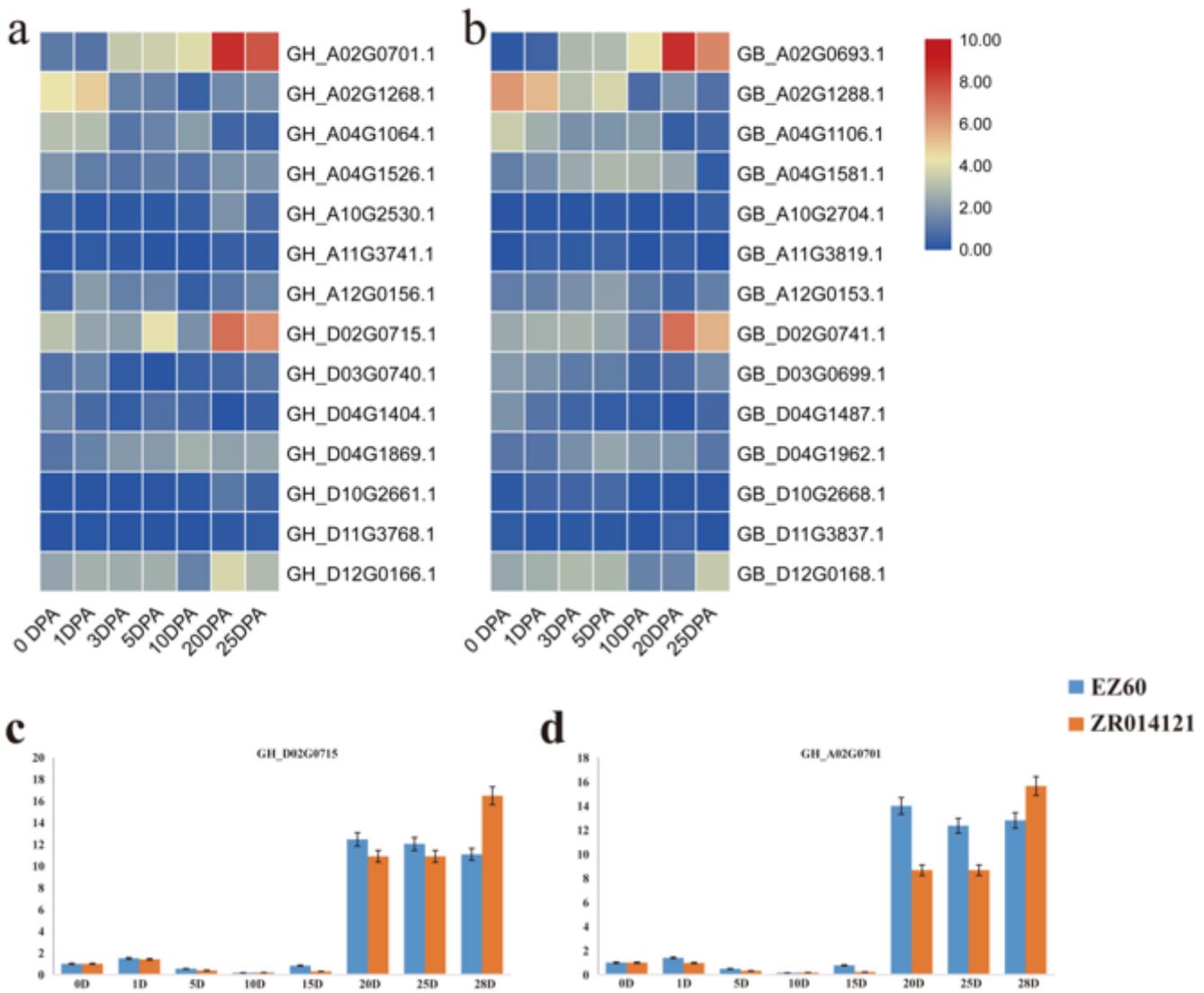
**Figure 4**

Collinearity of FBP genes between different cotton species. (a) between *G. raimondii* and *G. arboreum*; (b) between *G. raimondii*, *G. arboreum* and *G. hirsutum*; (c) between *G. raimondii*, *G. arboreum* and *G. barbadense*. (d) between *G. hirsutum* and *G. barbadense*.



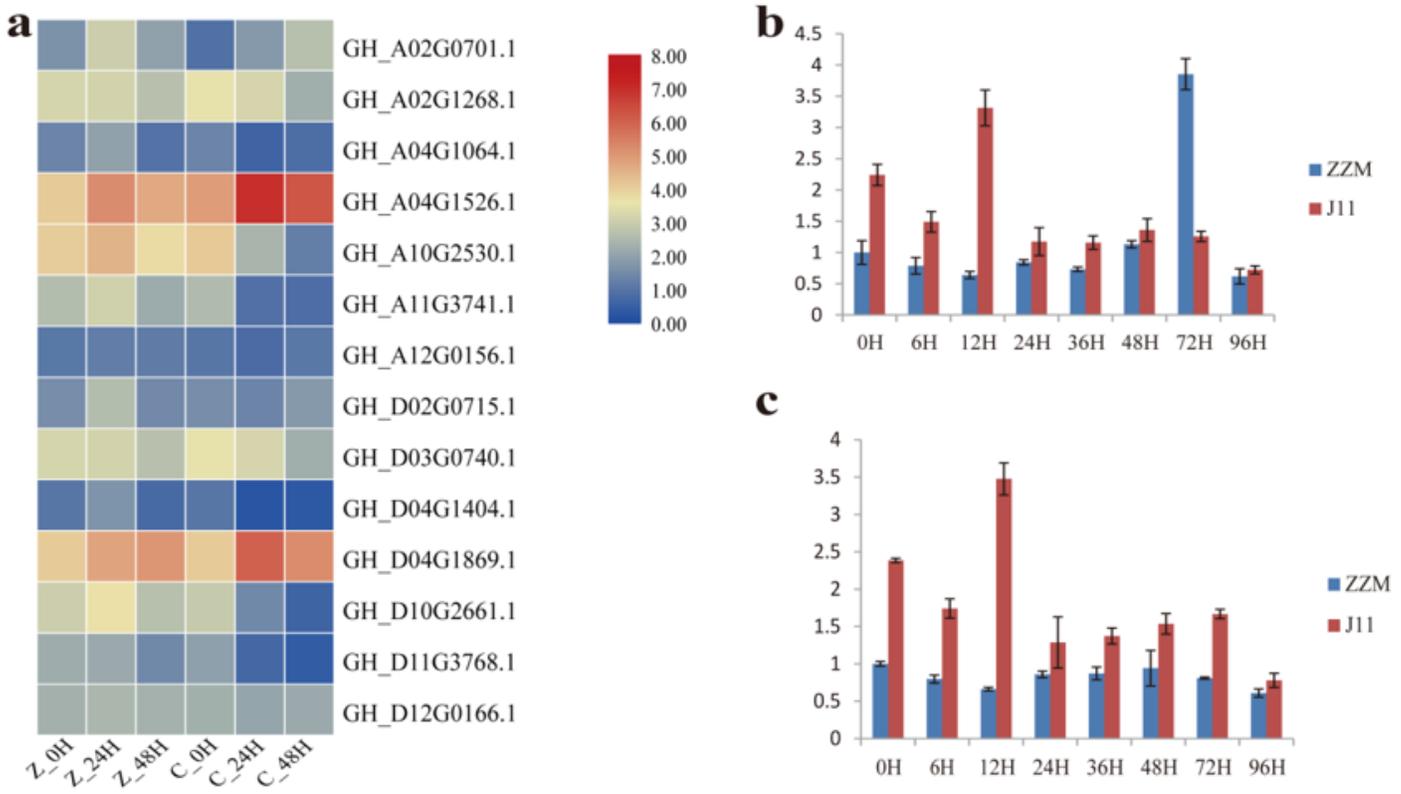
**Figure 5**

Multiple comparison of Ka/Ks ratio of genes pairs in four Gossypium species



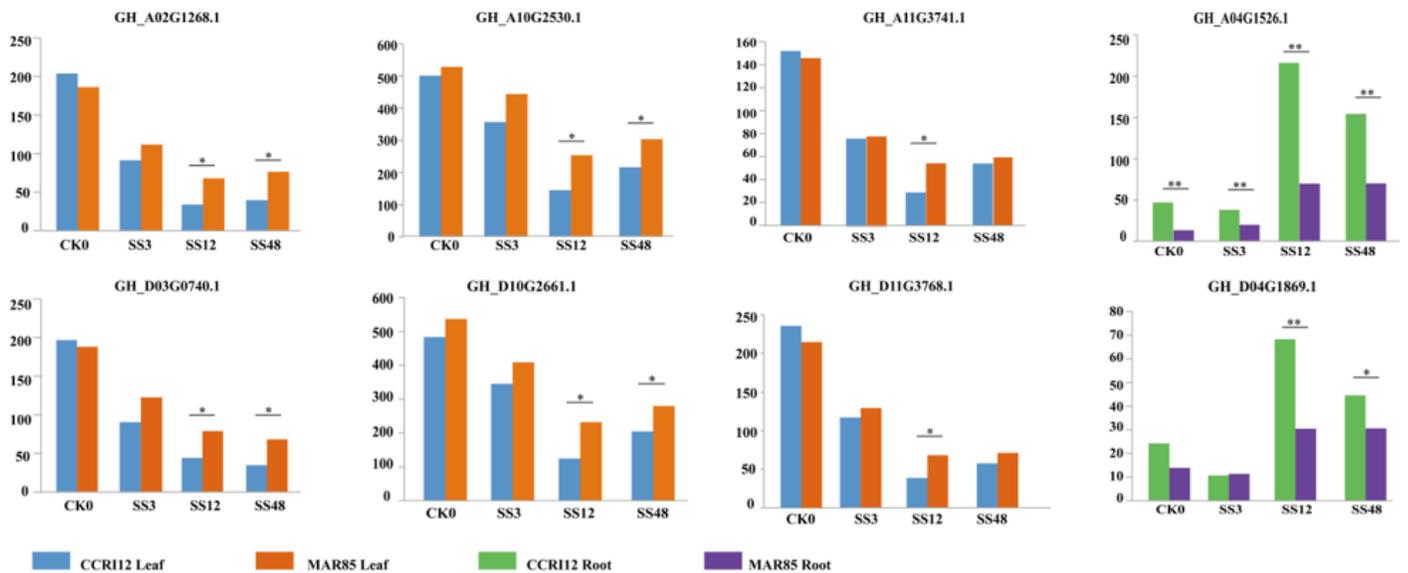
**Figure 6**

The expression levels of FBP genes in different stages of cotton fiber development. (a) Transcriptome analysis of FBP genes expression levels during the fiber development in *G. hirsutum*; (b) Transcriptome analysis of FBP genes expression levels during the fiber development in *G. barbadense*; (c) qRT-PCR verification of GH\_D02G0715.1; (d) qRT-PCR verification of GH\_A02G0701.1.



**Figure 7**

Expression profiling of FBP genes in response to *Verticillium* wilt stresses. a. Result of transcriptome analysis, Z represents the susceptible line CCRI36, C represents the resistant line MBI8255; 0, 24 and 48 represent hours after inoculation of *Verticillium dahliae*; b. qRT-PCR result of GH\_A04G1526.1; c. qRT-PCR result of GH\_D04G1869.1. The red bars represent Jimian11(J11) and the blue ones represent ZYM.



**Figure 8**

Expression profiling of upland cotton FBP genes evaluated using RNA transcriptome data analysis under salt stress.

## Supplementary Files

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

- [FigureS1.tif](#)
- [supplementaryfile1.fasta](#)
- [supplementaryfile2.xls](#)
- [supplementaryfile4.xls](#)
- [supplementaryfile3.xls](#)