Molecular Evolution and Expansion of the KUP Family in the Allopolyploid Cotton Species *Gossypium hirsutum* and *Gossypium barbadense*

Kai Fan [1,2,3], Zhijun Mao [1,2,3], Jiaxin Zheng [1,2,3], Yunrui Chen [1,2,3], Zhaowei Li [1,2,3], Weiwei Lin [1,2,3], Yongqiang Zhang [1,2,3], Jinwen Huang [1,2,3*] and Wenxiong Lin [1,2,3*]

[1] Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, College of Agriculture, Fujian Agriculture and Forestry University, Fuzhou, China, [2] Fujian Provincial Key Laboratory of Agroecological Processing and Safety Monitoring, College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, China, [3] Key Laboratory of Crop Ecology and Molecular Physiology (Fujian Agriculture and Forestry University), Fujian Province University, Fuzhou, China

The comprehensive analysis of gene family evolution will elucidate the origin and evolution of gene families. The K+ uptake (KUP) gene family plays important roles in K+ uptake and transport, plant growth and development, and abiotic stress responses. However, the current understanding of the KUP family in cotton is limited. In this study, 51 and 53 KUPs were identified in *Gossypium barbadense* and *Gossypium hirsutum*, respectively. These KUPs were divided into five KUP subfamilies, with subfamily 2 containing three groups. Different subfamilies had different member numbers, conserved motifs, gene structures, regulatory elements, and gene expansion and loss rates. A paleohexaploidization event caused the expansion of GhKUP and GbKUP in cotton, and duplication events in *G. hirsutum* and *G. barbadense* have happened in a common ancestor of *Gossypium*. Meanwhile, the KUP members of the two allopolyploid subgenomes of *G. hirsutum* and *G. barbadense* exhibited unequal gene proportions, gene structural diversity, uneven chromosomal distributions, asymmetric expansion rates, and biased gene loss rates. In addition, the KUP families of *G. hirsutum* and *G. barbadense* displayed evolutionary conservation and divergence. Taken together, these results illustrated the molecular evolution and expansion of the KUP family in allopolyploid cotton species.

Keywords: cotton, KUP family, polyploidization, molecular evolution, expansion

INTRODUCTION

Cotton (*Gossypium*) is a globally important natural source of fibers and is a good model for studying genome evolution and polyploidization in the plant kingdom. *Gossypium* has experienced two major whole-genome duplication events: a paleohexaploidization event and a cotton-specific decaploidy event (Paterson et al., 2012; Wang et al., 2016; Hu et al., 2019). The genus *Gossypium* consists of approximately 45 diploid (2n = 2x = 26) and five allopolyploid (2n = 4x = 52) species (Wendel and Albert, 1992). Diploid cotton species include eight genome groups (A–G and K genomes). Diploid cotton species, such as the A-genome species *Gossypium arboreum* and the

D-genome species *Gossypium raimondii*, diverged from a common eudicot ancestor belonging to the *Gossypium* genus approximately 5–10 million years ago (MYA). Approximately 1–2 MYA, allopolyploid species, including the two widely cultivated species *Gossypium hirsutum* and *Gossypium barbadense*, originated from the transoceanic hybridization of the A and D genome species. Then, allopolyploid species independently evolved in diverse environments (Wendel, 1989; Paterson et al., 2012; Wang et al., 2017). *G. hirsutum* and *G. barbadense* have different morphologies and economic traits (Hu et al., 2019; Wang M. et al., 2019). Physiological and phenotypic divergences are related to the metabolism of mineral nutrients in cotton. Many mineral elements, such as phosphorus and calcium, play critical roles in regulating numerous cotton biological processes (Tang et al., 2014; Santos et al., 2019).

In plants, K+ is an essential macronutrient and controls many key biological processes, such as root growth and development (Sustr et al., 2019), leaf senescence (Wang et al., 2012), and fruit quality (Hartz et al., 2005). K+ can influence cotton yield and quality (Pettigrew, 2008). For example, the cell expansion of cotton fibers can be regulated by turgor pressure under the mediation of K+ concentration (Dhindsa et al., 1975). Plants absorb and transport K+ through K+ channels or transporters (Véry et al., 2014). K+ transporters in plants can be divided into the KT/HAK/K+ uptake (KUP), Trk/HKT, CHX, and KEA families (Mäser et al., 2001; Gupta et al., 2008). The KUP family is the largest K+ transporter family (Kim et al., 1998; Alemán et al., 2014; Han et al., 2016). It can regulate plant growth and developmental processes, including root hair (Ahn et al., 2004) and shoot cell expansion (Elumalai et al., 2002), and is associated with salt stress and osmotic regulation (Osakabe et al., 2013; Chen et al., 2015).

Members of the KUP family have been found in fungi, bacteria, and plants (Rodríguez-Navarro, 2000). A KUP member in plants was first identified in *Arabidopsis* (Quintero and Blatt, 1997). Since then, many KUP members have been reported in various flowering plants, such as *Oryza sativa* (Gupta et al., 2008), *Zea mays* (Zhang et al., 2012), *Triticum aestivum* (Cheng et al., 2018), *Solanum lycopersicum* (Hyun et al., 2014), *Populus trichocarpa* (He et al., 2012), *Prunus persica* (Song et al., 2015), *Cicer arietinum* (Azeem et al., 2018), *Nicotiana tabacum* (Song et al., 2019), and *Manihot esculenta* (Ou et al., 2018). Some KUP genes have been reported in cotton (Wang et al., 2018; Wang Y. et al., 2019). *G. hirsutum* K+ transporter 2 (GhKT2) may be related to K+ acquisition, transport, and distribution (Wang et al., 2018), and *G. hirsutum* high-affinity K+ transporter 5a (GhHAK5a) is essential for the shoot regulation of root K+ uptake under potassium deficiency (Wang Y. et al., 2019). However, a genome-wide systematic analysis of the KUP family in cotton does not exist.

The complete genome sequences of *G. hirsutum*, *G. barbadense*, and their putative diploid donor species *G. arboreum* and *G. raimondii* provide ideal opportunities for investigating the evolutionary and functional genomics of cotton (Du et al., 2018; Hu et al., 2019; Udall et al., 2019). A comprehensive investigation of molecular evolution is urgently needed to elucidate the complex evolutionary history of KUP proteins after cotton polyploidization. The comprehensive evolutionary history of the KUP family can unravel selection and accelerate the molecular breeding of cotton. In the present study, the KUP members of *Arabidopsis thaliana*, *O. sativa*, *Vitis vinifera*, *Theobroma cacao*, *Bombax ceiba*, *Corchorus capsulari*, *Corchorus olitorius*, *Herrania umbratica*, *Durio zibethinus*, *Gossypioides kirkii*, *Gossypium austral*, *Gossypium turneri*, *G. raimondii*, *G. arboretum*, *G. hirsutum*, and *G. barbadense* were subjected to genome-wide and comparative genomic analyses to reveal the molecular evolutionary history of the gene family, including the asymmetric evolution of subfamilies, the unequal evolution of subgenomes, and the conserved and divergent evolution of the two cultivated allopolyploid species.

MATERIALS AND METHODS

Sequence Retrieval and Phylogenetic Analysis

The genome sequences of *A. thaliana*, *O. sativa*, *V. vinifera*, *T. cacao*, *B. ceiba*, *C. capsulari*, *C. olitorius*, *H.umbratica*, *D. zibethinus*, *G. kirkii*, *G. austral*, *G. turneri*, *G. raimondii*, *G. arboretum*, *G. hirsutum*, and *G. barbadense* were downloaded from publically available databases (**Table S1**). Hmmsearch in HMMER 3.0 program was conducted to identify the KUP members of the above-mentioned species by using the KUP domain (Pfam ID: PF02705) as the query (Finn et al., 2011). Then, the conserved KUP domains of the candidate KUP sequences were further confirmed by utilizing the CDD program with default settings (https://www.ncbi.nlm.nih.gov) (Marchler-Bauer et al., 2015). Finally, all of the KUP members with KUP domains (Pfam ID: PF02705) were retained for further analysis. The KUP members of *T. cacao*, *G. raimondii*, *G. arboretum*, *G. hirsutum*, and *G. barbadense* were designated as TcKUP, GrKUP, GaKUP, GhKUP, and GbKUP, respectively.

All of the KUP sequences were aligned by using MAFFT (Katoh and Standley, 2013). Parameters, except for globalpair and maxiterate, were set with the default settings. Maxiterate was set to 1,000. A phylogenetic tree was inferred *via* maximum likelihood (ML) by using IQ-tree (Kalyaanamoorthy et al., 2017). The best model was selected by applying the ModelFinder program, and the bootstrap value was set to 1,000 (Kalyaanamoorthy et al., 2017). All trees were visualized with MEGA software (Kumar et al., 2018).

Structural and Promoter Analysis

The conserved motifs of the KUP family in *G. hirsutum*, *G. barbadense*, *G. arboretum*, and *G. raimondii* were identified with MEME (http://meme-suite.org/tools/meme) (Bailey et al., 2009). Parameters were set on the basis of values that were previously used to analyze putative motifs (Fan et al., 2019; Fan et al., 2020). The identified motifs were further annotated by using CDD (Marchler-Bauer et al., 2015). The structural and chromosomal location information of all the *KUP* genes in cotton was downloaded from CottonGen (https://www.cottongen.org/). The gene structures and chromosomal distribution images of

*GhKUP*s, *GbKUP*s, *GaKUP*s, and *GrKUP*s were visualized with TBtools (Chen et al., 2020).

The 1000-bp upstream sequences of *GhKUP*s and *GbKUP*s were extracted from CottonGen. Potential *cis*-elements in the *GhKUP* and *GbKUP* promoters were identified with PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Stress-responsive regulatory elements, including the ABA-responsive element (ABRE), gibberellin-responsive element (GARE), defense and stress responsiveness (DSR), low-temperature-responsive element (LTR), and MYB-binding site (MBS), were visualized by using TBtools (Chen et al., 2020).

Identification of Syntenic Blocks and Orthologous Gene Pairs

Syntenic blocks in *G. hirsutum*, *G. barbadense*, *G. arboretum*, and *G. raimondii* were identified with MCScan (Python version) (https://github.com/tanghaibao/jcvi/wiki/MCscan-[Python-version]). Each block contained at least five genes. Circos was utilized to visualize syntenic relationships (Krzywinski et al., 2009). Moreover, the orthologous gene pairs of four cotton species were searched on the basis of syntenic block and phylogenetic tree results. In addition, the nonsynonymous distance (Ka), synonymous distance (Ks), and their ratio (Ka/Ks) were estimated by using TBtools software with the Nei–Gojobori model (Nei and Gojobori, 1986; Chen et al., 2020). The density plots of Ks were analyzed and visualized by using ggplot2 package in RStudio (Studio, 2012; Wickham et al., 2016). Statistical analyses were performed in RStudio (Studio, 2012). In accordance with the neutral substitution rate of cotton ($r = 2.6 \times 10^{-9}$), the estimated divergence times were calculated by using the formula t = Ks/2r (Hu et al., 2019).

Expression Analysis

G. hirsutum acc. TM-1 and *G. barbadense* acc. Hai7124 were used for expression analysis. Roots, stems, and leaves were harvested from 4-week-old seedlings grown in a greenhouse. For abiotic stress treatments, four-week-old seedlings were subjected to drought stress (20% PEG6000), salt stress (250 mM NaCl), and ABA treatment (100 mM ABA) in Hoagland's solution. The third true leaves were collected at 0, 6, 12, and 24 h after treatment. Three biological replications were assayed. All collected samples were quick-frozen in liquid nitrogen and then stored at −80°C.

Total RNA was extracted from frozen samples with the RNAprep pure Plant Kit (TIANDZ, China), and first-strand cDNA was synthesized from DNase-treated RNA with a PrimerScript 1st Strand cDNA synthesis kit (TaKaRa, Dalian, China). Gene-specific primers for quantitative real-time PCR (qRT-PCR) analysis were designed and synthesized (**Table S2**). BLAST searches against the *G. hirsutum* and *G. barbadense* genomes were performed to confirm the primers' specificity. qRT-PCR was run on a CFX96 Realtime System (BioRad) with SYBR premix Extaq (TaKaRa, Dalian, China). An annealing temperature of 60°C and 40 cycles were set for all qRT-PCR reactions. The cotton *EF1α* gene was selected as the endogenous control to normalize expression data. Relative expression levels

were calculated in accordance with the comparative cycle threshold method. All reactions were conducted with three biological replications. The expression levels of *GhKUP*s and *GbKUP*s were clustered by using ComplexHeatmap package in RStudio (Studio, 2012; Gu et al., 2016). Statistical analyses were performed in RStudio (Studio, 2012). The Pearson correlation coefficients (PCCs) of the expression data were calculated and visualized by using ggcorrplot package in RStudio (Studio, 2012; Kassambara, 2018).
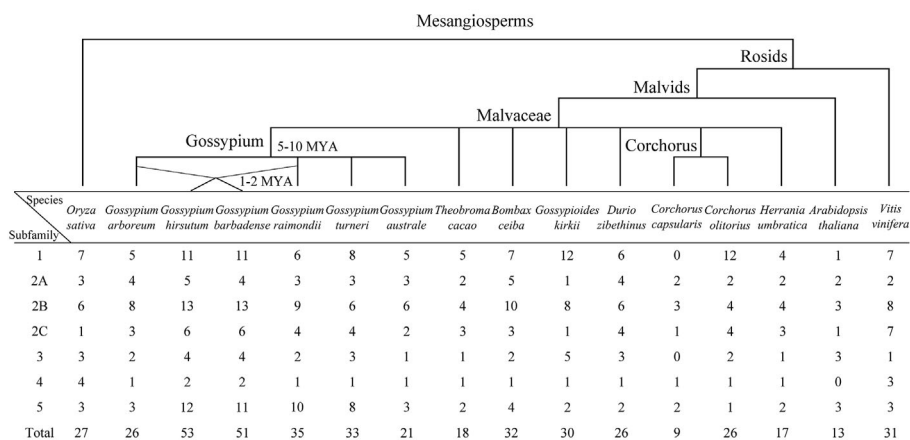
RESULTS

Identification and Phylogenetic Analysis of KUPs in Malvaceae, *Arabidopsis*, Grape, and Rice

A total of 448 KUP members were identified in 13 Malvaceae species, *Arabidopsis*, grape, and rice (**Figures 1**, **2**; **Tables S3**, **S4**). The lowest number of KUPs was found in *C. capsularis* (nine members). *Arabidopsis thaliana*, *Herrania umbratica*, and *Theobroma cacao* had less than 20 members. The highest number of members was identified in *G. barbadense* (51) and *G. hirsutum* (53), whereas the diploid donor species *G. arboretum* and *G. raimondii* had 26 and 35 members, respectively. In addition, 24 GbKUPs and 25 GhKUPs were found in the *G. barbadense* At (GbAt) and *G. hirsutum* At (GhAt) subgenomes, respectively. The *G. barbadense* Dt (GbDt) and *G. hirsutum* Dt (GhDt) subgenomes contained 27 GbKUPs and 28 GbKUPs, respectively.

On the basis of previously reported similar classifications (Cheng et al., 2018; Ou et al., 2018; Song et al., 2019), the KUP family can be divided into five subfamilies, which were numbered from 1 to 5. Subfamily 2 could be further classified into three groups, namely, 2A, 2B, and 2C (**Figures 1**, **2**). The KUP members were unevenly distributed throughout each subfamily. In most species, subfamily 2B contained the highest number of KUPs (more than 22%), followed by subfamilies 1 and 5, and subfamily 4 usually had the fewest KUPs (less than 10%). Nearly 70% of the KUP members of *G. barbadense* and *G. hirsutum* belonged to subfamilies 1, 2B, or 5, whereas subfamily 4 of these species contained only two KUPs (**Figure S1**). The number of GbKUPs in subfamilies 1, 3, and 4 was equal to the sum of GaKUPs and GrKUPs. The other subfamilies had low numbers of GbKUPs. GhKUPs also had the same distribution. Most of KUPs that had been lost from *G. barbadense* and *G. hirsutum* were located in the GbDt and GhDt subgenomes. For example, GbKUPs and GhKUPs in subfamily 2B primarily lost three KUPs from the GbDt and GhDt subgenomes.

Conserved Motifs and Gene Structures of the KUP Family in Cotton

All of the KUP members had the conserved KUP domain (**Table S5**). The MEME tool identified 20 putative conserved motifs in GbKUPs, GhKUPs, GaKUPs, and GrKUPs (**Figure 3**; **Table S6**).

Species / Subfamily	Oryza sativa	Gossypium arboreum	Gossypium hirsutum	Gossypium barbadense	Gossypium raimondii	Gossypium turneri	Gossypium australe	Theobroma cacao	Bombax ceiba	Gossypioides kirkii	Durio zibethinus	Corchorus capsularis	Corchorus olitorius	Herrania umbratica	Arabidopsis thaliana	Vitis vinifera
1	7	5	11	11	6	8	5	5	7	12	6	0	12	4	1	7
2A	3	4	5	4	3	3	3	2	5	1	4	2	2	2	2	2
2B	6	8	13	13	9	6	6	4	10	8	6	3	4	4	3	8
2C	1	3	6	6	4	4	2	3	3	1	4	1	4	3	1	7
3	3	2	4	4	2	3	1	1	2	5	3	0	2	1	3	1
4	4	1	2	2	1	1	1	1	1	1	1	1	1	1	0	3
5	3	3	12	11	10	8	3	2	4	2	2	2	1	2	3	3
Total	27	26	53	51	35	33	21	18	32	30	26	9	26	17	13	31

FIGURE 1 | Distribution of KUP members in 13 Malvaceae species, *Arabidopsis*, grape, and rice. The upper branch displays the simple phylogenetic relationships of the 16 species. The bottom table represents the number of KUP genes in each KUP subfamily.

Consistent with phylogenetic analysis, KUP members could be further classified into five subfamilies, with subfamily 2 containing three groups, on the basis of the distribution of the identified motifs. After CDD search, 18 motifs were annotated as KUP domains. Motifs 8 and 17 lacked any annotations. More than 100 cotton KUP members had 15 motifs. Motifs 4, 8, 18, 19, and 20 existed in some specific subfamilies. Over 75% of the exon numbers of the cotton KUP members ranged from 6 to 10 (**Figure 4**; **Table S7**). The highest exon numbers of GaKUPs, GrKUPs, GbKUPs, and GhKUPs were 16, 15, 13, and 11, respectively. Moreover, the At subgenome of *G. barbadense* and *G. hirsutum* had more highly fragmented KUPs than the Dt subgenome of the two species. The highest exon numbers of the GbAt, GbDt, GhAt, and GhDt subgenomes were 13, 10, 11, and 10, respectively. The exon numbers of the At subgenome were significantly higher than those of the Dt subgenome. Furthermore, gene structure was highly conserved within each subfamily. The KUP members of subfamily 4 had eight exons, and many KUP members with highly fragmented gene structures could be found in subfamily 5.

Expression Patterns of *GhKUP*s and *GbKUP*s in Different Tissues

*GhKUP*s and *GbKUP*s showed tissue-specific expression patterns in roots, stems, and leaves (**Figure 5**). In *G. hirsutum*, six *GhKUP*s (*GhKUP02*, *GhKUP03*, *GhKUP13*, *GhKUP29*, *GhKUP36*, and *GhKUP46*) had significantly higher expression levels in stems and leaves than in other tissues. *GhKUP04*, *GhKUP05*, and *GhKUP09* showed the significantly highest expression levels in roots among all the tested tissues. Eleven *GhKUP*s (*GhKUP10*, *GhKUP17*, *GhKUP18*, *GhKUP23*, *GhKUP33*, *GhKUP37*, *GhKUP39*, *GhKUP40*, *GhKUP41*, *GhKUP45*, and *GhKUP52*) exhibited significantly higher expression patterns in stems than in other tissues, whereas GhKUP07 showed opposite expression patterns. Compared with those in other tissues, the expression levels of eight *GhKUP*s (*GhKUP11*, *GhKUP12*, *GhKUP14*, *GhKUP16*, *GhKUP26*, *GhKUP35*, *GhKUP44*, and *GhKUP49*) were significantly higher and those of *GhKUP08*, *GhKUP24*, *GhKUP27*, and *GhKUP51* were significantly lower in leaves. In *G. barbadense*, *GbKUP02*, *GbKUP04*, *GbKUP12*, and *GbKUP43* significantly peaked in roots. *GbKUP15*, *GbKUP23*, *GbKUP31*, *GbKUP42*, and *GbKUP50* showed significantly higher expression patterns in leaves than in other tissues, whereas *GbKUP08*, *GbKUP17*, *GbKUP18*, and *GbKUP44* had the opposite expression patterns. Eight *GbKUP*s (*GbKUP03*, *GbKUP10*, *GbKUP13*, *GbKUP22*, *GbKUP35*, *GbKUP40*, *GbKUP47*, and *GbKUP49*) showed significantly higher expression levels in stems than in other tissues.

Regulatory Elements of the GhKUP and GbKUP Promoters

The regulatory elements of the promoters are crucial. Many stress-responsive regulatory elements, including ABRE, GARE, DSR, LTR, and MBS, were identified in the promoter regions of *GhKUP*s and *GbKUP*s through PlantCARE analysis (**Figure 6**; **Table S8**). Over 80% of the *KUP* promoters in *G. hirsutum* and *G. barbadense* had the above-mentioned elements. The most common element was ABRE (19 *GbKUP*s and 24 *GhKUP*s), followed by GARE (16 *GbKUP*s and 16 *GhKUP*s), DSR (17 *GbKUP*s and 13 *GhKUP*s), and MBS (13 *GbKUP*s and 15 *GhKUP*s). The promoter regions of 44 *KUP*s contained at least two stress-responsive regulatory elements. For example, *GhKUP27* contained DSR and MBS in its promoter region, and ABRE and MBS were found in the promoter region of *GbKUP47*. Moreover, *GbKUP*s and *GhKUP*s from the same subfamily had similar elements in their promoter regions. To illustrate, ABRE, DSR, and MBS were abundant in the promoter regions of *GbKUP*s and *GhKUP*s from subfamily 2B. ABRE, GARE, and MBS existed in most of the promoter regions of *GbKUP*s and *GhKUP*s from subfamily 5.
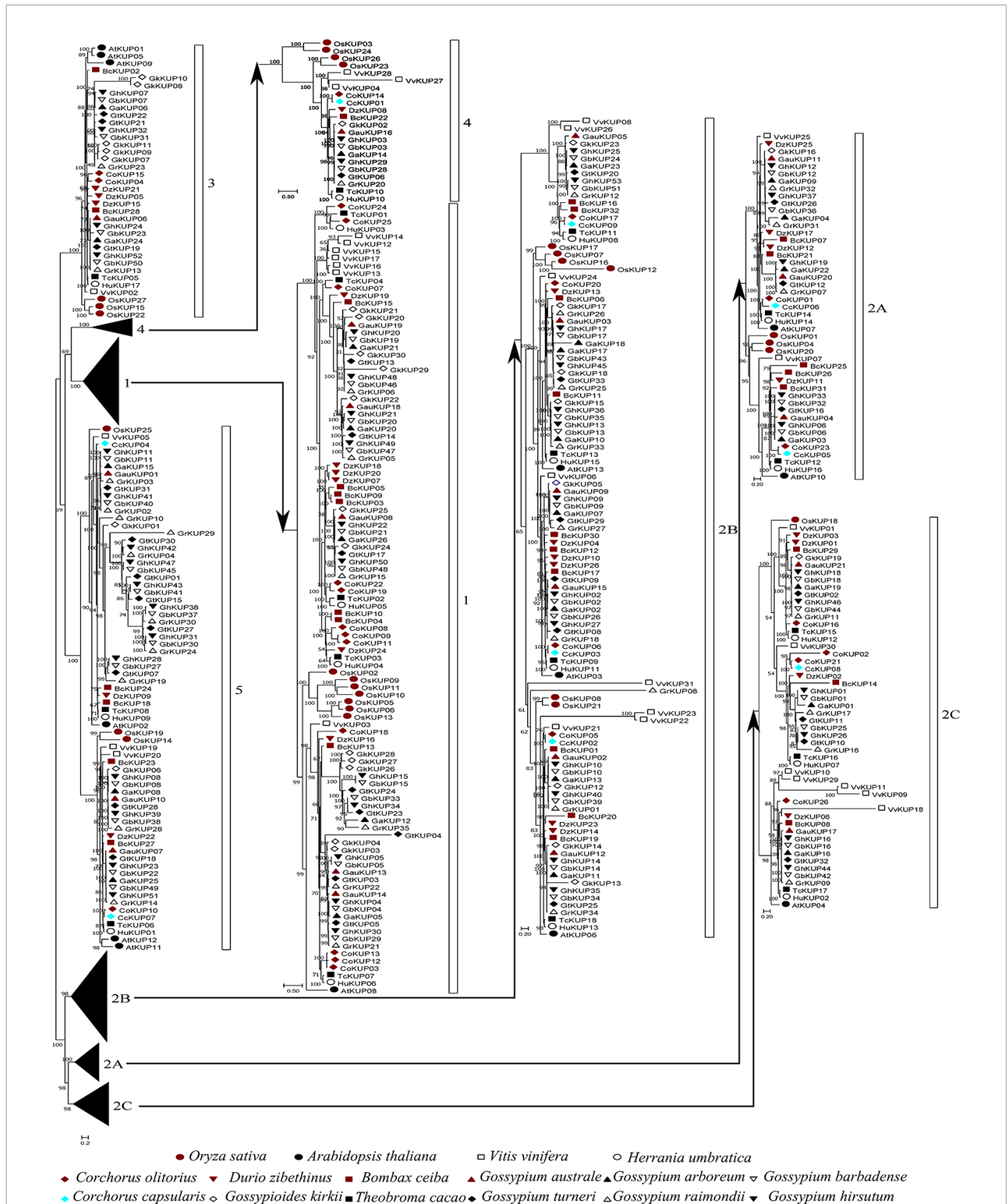
FIGURE 2 | Phylogenetic analysis of the KUP family in 13 Malvaceae species, *Arabidopsis*, grape, and rice. The phylogenetic tree was constructed *via* the ML method. Numbers in the clades represent bootstrap values, and KUP subfamilies are indicated by different letters. KUP members from the investigated species are marked by differently colored shapes.
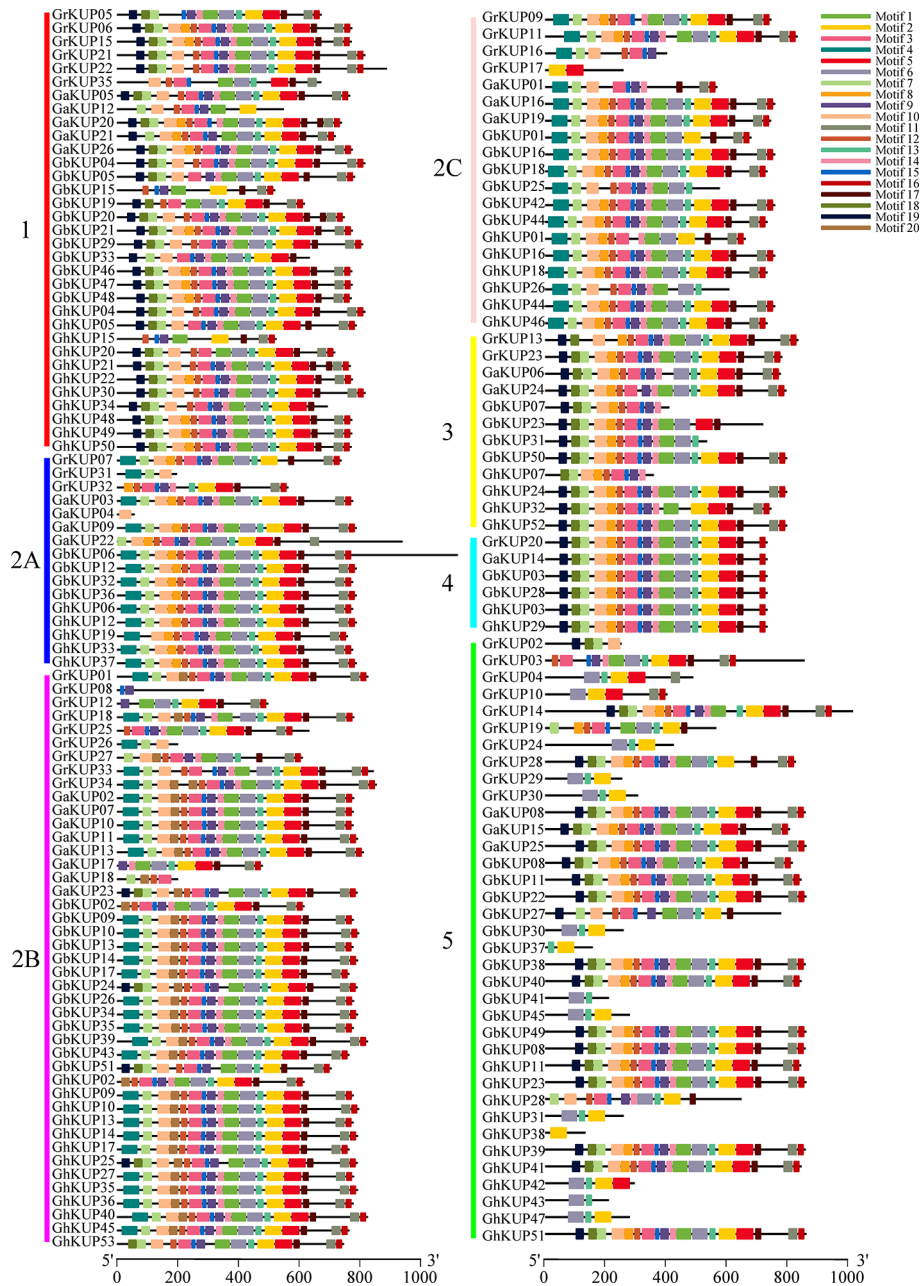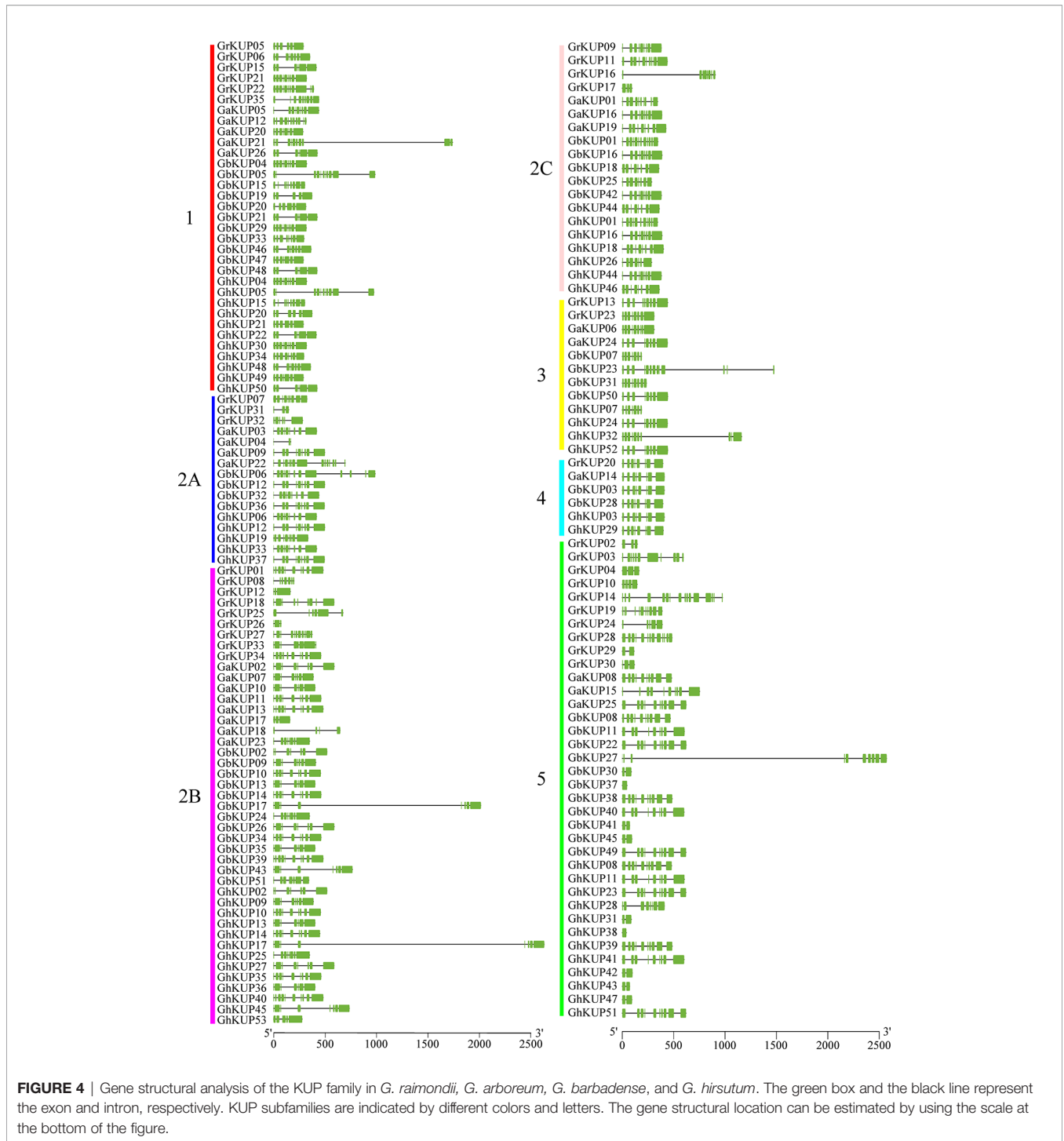
FIGURE 3 | Putative motif distributions of the KUP family in *G. raimondii*, *G. arboreum*, *G. barbadense*, and *G. hirsutum*. Different motifs are highlighted by colored boxes. KUP subfamilies are indicated by different colors and letters. Motif location can be estimated by using the scale at the bottom of the figure.

## Expression Patterns of *GhKUP*s and *GbKUP*s Under Abiotic Stresses

On the basis of the results for regulatory elements, the expression levels of *GbKUP*s and *GhKUP*s under three abiotic stress conditions (drought, salt, and ABA) were further investigated. Most of the KUPs in *G. hirsutum* and *G. barbadense* exhibited different expression levels after exposure to various stresses (**Figures 7**, **8**). Under drought stress, *GhKUP02*, *GhKUP08*,

GhKUP10, *GhKUP11*, *GhKUP12*, *GhKUP14*, *GhKUP16*, *GhKUP18*, *GhKUP33*, *GhKUP35*, *GhKUP41*, *GhKUP46*, *GhKUP51*, *GhKUP52*, and *GbKUP15* were significantly downregulated, and others were significantly upregulated (**Figure 7**). *GhKUP13*, *GhKUP17*, *GhKUP24*, *GhKUP39*, *GhKUP40*, *GhKUP44*, *GhKUP45*, *GbKUP23*, *GbKUP31*, and *GbKUP35* were rapidly upregulated after 6 h or 12 h of drought treatment, and others were upregulated after 24 h of

FIGURE 4 | Gene structural analysis of the KUP family in *G. raimondii*, *G. arboreum*, *G. barbadense*, and *G. hirsutum*. The green box and the black line represent the exon and intron, respectively. KUP subfamilies are indicated by different colors and letters. The gene structural location can be estimated by using the scale at the bottom of the figure.

drought treatment. In particular, *GhKUP03*, *GhKUP27*, *GbKUP08*, and *GbKUP47* were highly upregulated after 24 h of drought treatment. Under salt stress, all *GbKUPs* and *GhKUPs* showed significantly upregulated expression levels (**Figure 7**). *GhKUP27*, GbKUP10, GbKUP17, and GbKUP40 were upregulated after 24 h of salt treatment, and others were rapidly upregulated after 6 or 12 h of salt treatment. *GhKUP04*,

GhKUP07, *GhKUP27*, *GbKUP15*, and *GbKUP47* were highly upregulated under salt stress. Under ABA treatment, all *GbKUPs*, particularly *GbKUP03*, *GbKUP17*, *GbKUP35*, and *GbKUP47*, were rapidly upregulated after 6 h, whereas 12 *GhKUPs* (*GhKUP03*, *GhKUP04*, *GhKUP07*, *GhKUP09*, *GhKUP17*, *GhKUP24*, *GhKUP26*, *GhKUP27*, *GhKUP37*, *GhKUP44*, *GhKUP45*, and *GhKUP49*) were significantly
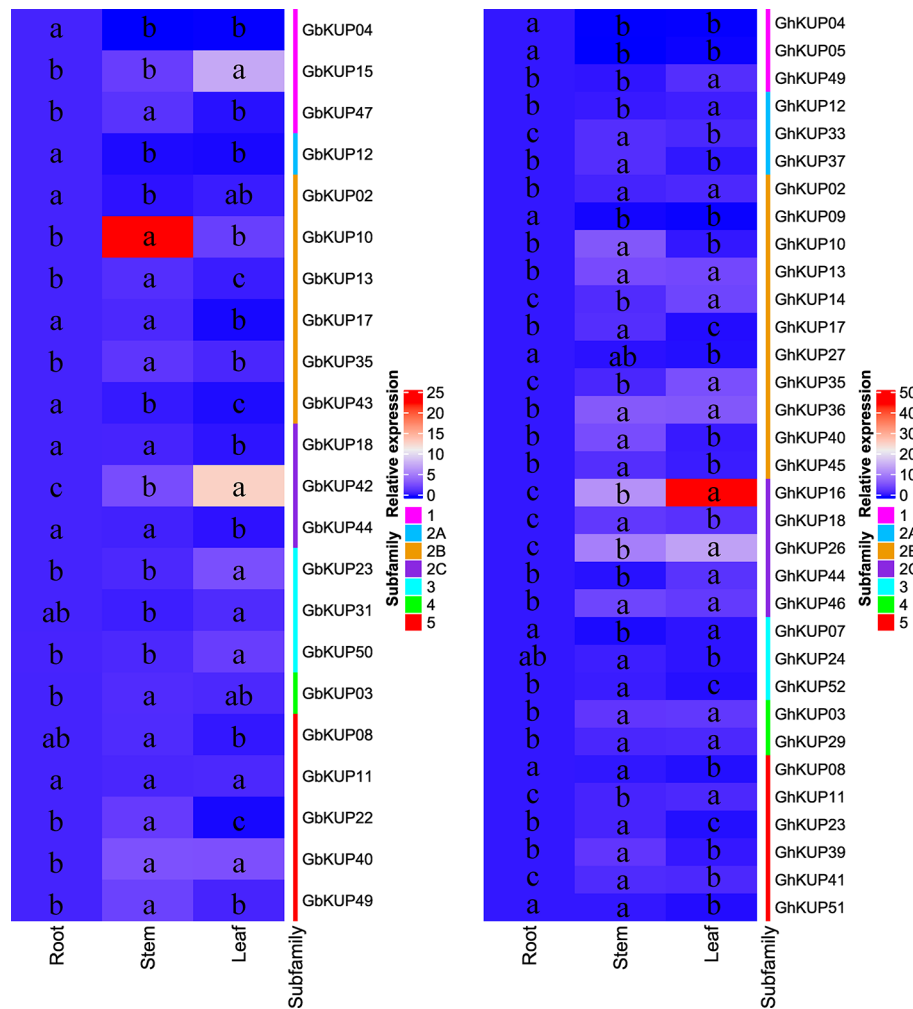
FIGURE 5 | Heat map representation of *GhKUP* and *GbKUP* expression levels in roots, stems, and leaves. Two color bars represent relative expression values and KUP subfamilies. Expression levels with the same letter in each cell do not significantly differ at *P* < 0.05 as revealed by Duncan's multiple test.

upregulated after 12 or 24 h of stress treatment (**Figure 8**). *GhKUP04*, *GhKUP27*, and *GhKUP45* were highly upregulated under ABA treatment. Other *GhKUPs* were significantly downregulated under ABA treatment.H

Genomic Locations and Gene Duplication of GhKUPs and GbKUPs

GbKUPs and *GhKUPs* were located unevenly across chromosomes (**Figures S1-S3**; **Table S9**). Chromosomes 6, 7, and 13 in each subgenome lacked *KUP* genes. At05 and Dt04 in the At and Dt subgenomes had the highest number of *KUPs*. Furthermore, *GbKUPs* and *GhKUPs*, except for At11 and Dt5, had the same distribution pattern on each chromosome. In At11 and Dt5, *G. hirsutum* had one more *KUP* member than *G. barbadense*.

G. hirsutum and *G. barbadense* had each undergone nine KUP family gene duplication events, whereas *G. arboreum* had experienced three and *G. raimondii* underwent five

(**Figures 9A, B**; **Table S10**). All duplicated genes were located in different chromosomes, indicating that the gene duplication events were segmental duplication events. The KUP family of cacao, which shares a common ancestor with cotton, underwent one segmental gene duplication event (*TcKUP* 09/13). Duplicated *GhKUPs* and *GbKUPs* belonged to subfamilies 2B, 3, and 5 (**Figure 9C**). Subfamilies 2B and 3 of *G. hirsutum* and *G. barbadense* had twice as many duplicated gene pairs as those of their diploid donor species. However, two additional duplicated gene pairs in subfamily 5 existed only in *G. hirsutum* and *G. barbadense*. Moreover, four and five duplicated gene pairs were distributed in the At and Dt subgenomes, respectively. The orthologous genes of duplicated *GhKUPs* and *GbKUPs* were identified in their corresponding diploid donor species (**Figures 9A, B**; **Table S11**). The orthologous genes of six duplicated *KUPs* (*GaKUPs* and *GrKUPs*) from subfamilies 2B and 3 were duplicated in each allopolyploid species. In subfamily 5, the
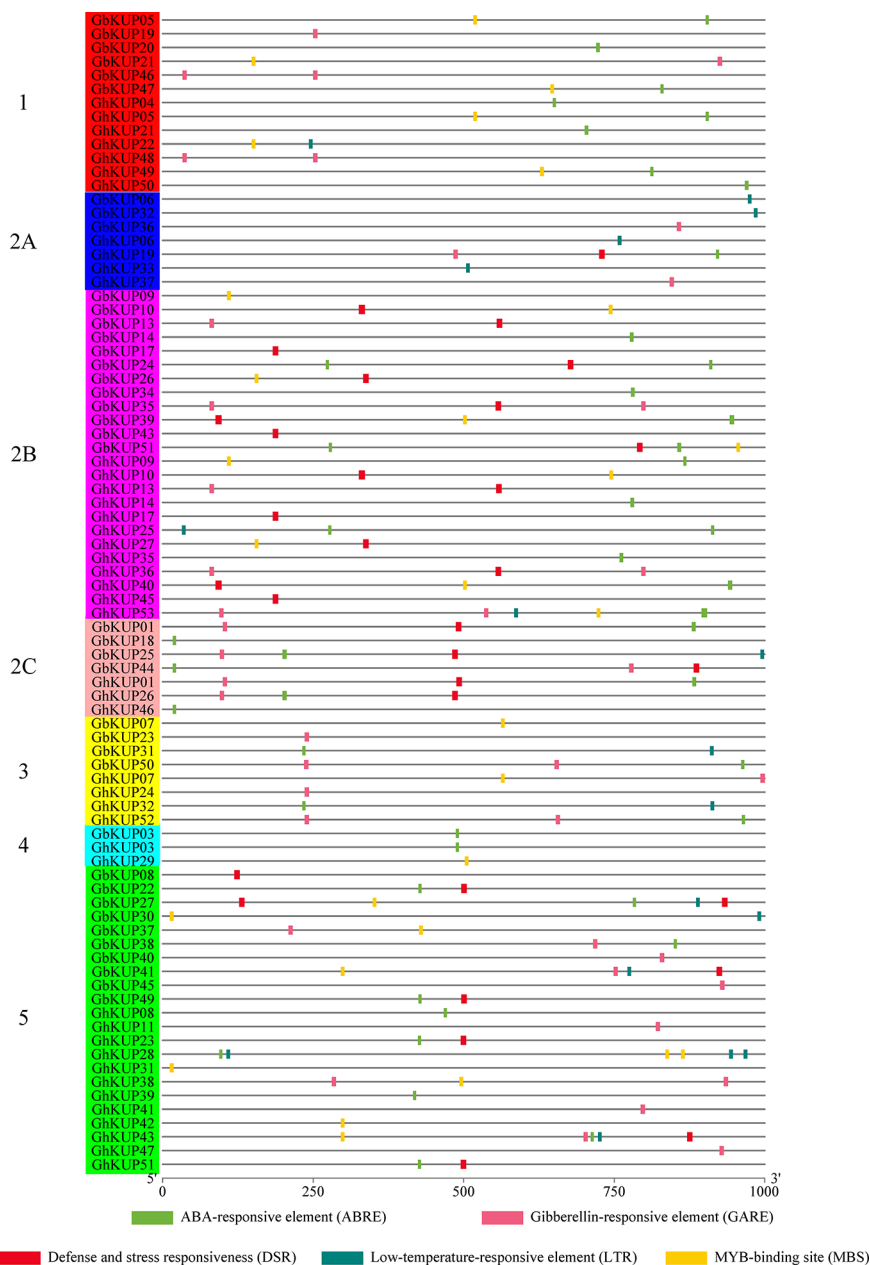
FIGURE 6 | Stress-responsive regulatory elements in the promoter regions of *GhKUP*s and *GbKUP*s. The stress-responsive regulatory elements were predicted in the promoter region of *GhKUP*s and *GbKUP*s by applying the PlantCARE database. The KUP subfamilies are marked by using different colors and letters. The colored boxes represent stress-responsive elements. The location of the stress-responsive elements can be estimated by using the scale at the bottom of the figure.

orthologous genes of two duplicated KUPs (*GhKUP 39/51* and *GbKUP 38/49*) were duplicated in *G. raimondii* (*GrKUP 14/18*), whereas the orthologous genes of four duplicated KUPs (*GhKUP 08/23*, *GhKUP 28/41*, *GbKUP 08/22*, and *GbKUP 27/40*) were not duplicated in *G. arboreum* and *G. raimondii*. The phylogenetic tree of the duplicated *KUP*s in cotton and their orthologous genes in cacao was used to predict the relative time of gene duplication (**Figure 10**). The duplication events of the KUP

family were found only in cotton, and orthologous genes in cacao were not duplicated.

The evolutionary distances of nine duplicated *KUP* pairs in *G. hirsutum*, nine duplicated *KUP* pairs in *G. barbadense*, three duplicated *KUP* pairs in *G. arboreum*, and five duplicated *KUP* pairs in *G. raimondii* were calculated (**Figure 9D**; **Table S10**). The Ks values peaked at approximately 0.489, and the divergence time of the duplicated gene pairs corresponded to
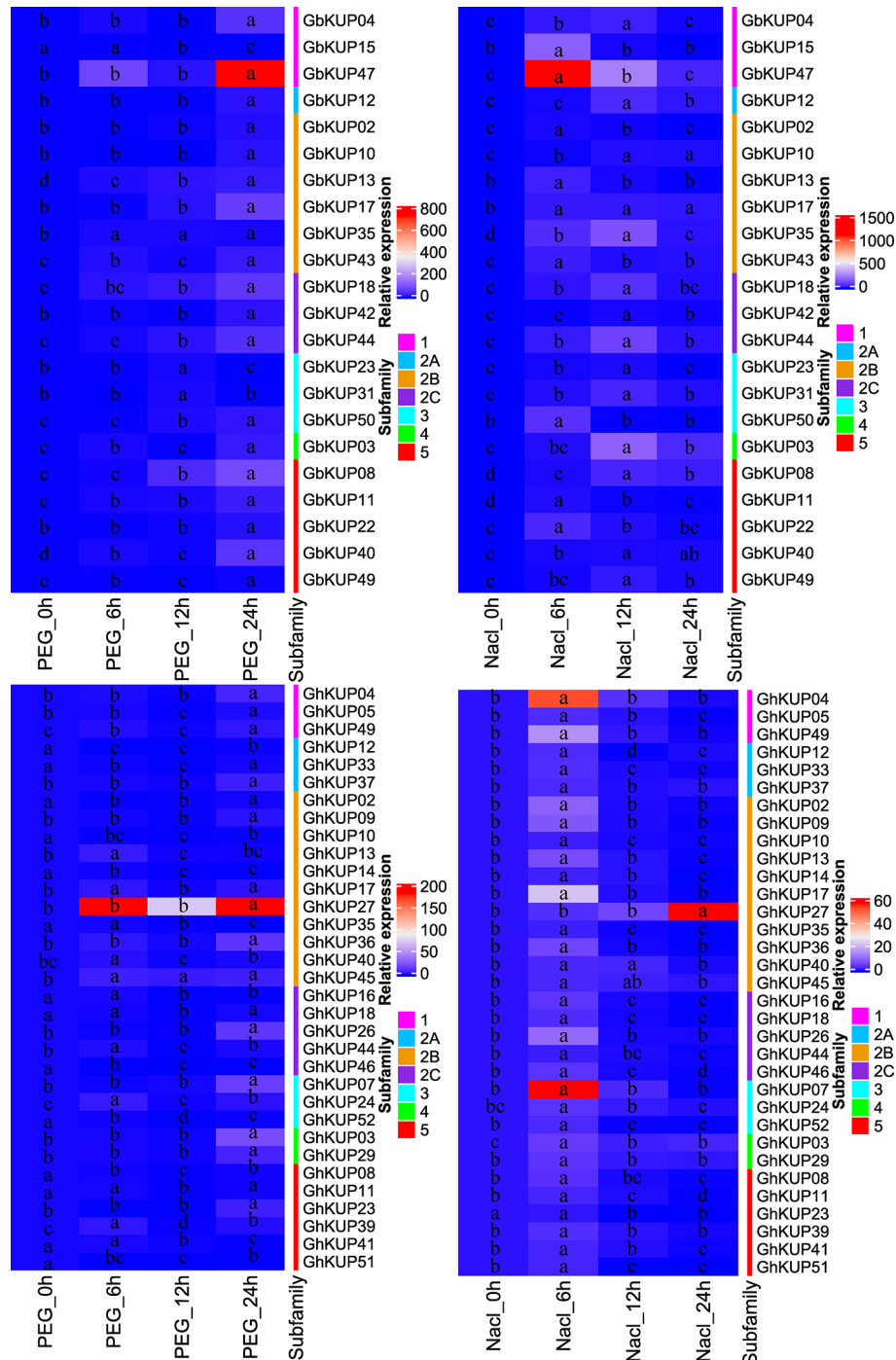
FIGURE 7 | Heat map representation of *GhKUP* and *GbKUP* expression levels under drought and salt stresses. Two color bars represent the relative expression value and KUP subfamily. Expression levels with the same letter in each cell do not significantly differ at *P* < 0.05 as revealed by Duncan's multiple test.

approximately 94 MYA. Furthermore, the Ka/Ks ratios of all duplicated *KUP* members were less than 1. PCC values were calculated on the basis of the expression levels of *GhKUP*s and *GbKUP*s in different tissues and stresses (**Figures 9E, F**). All of the expression levels of the duplicated gene pairs

were positively correlated (PCC > 0), and over 50% of the duplicated gene pairs had highly positively correlated expression levels (PCC > 0.4). For example, the PCC values of the expression levels of the duplicated *GhKUP 13/17* and *GbKUP 35/43* were 0.66 and 0.64, respectively.
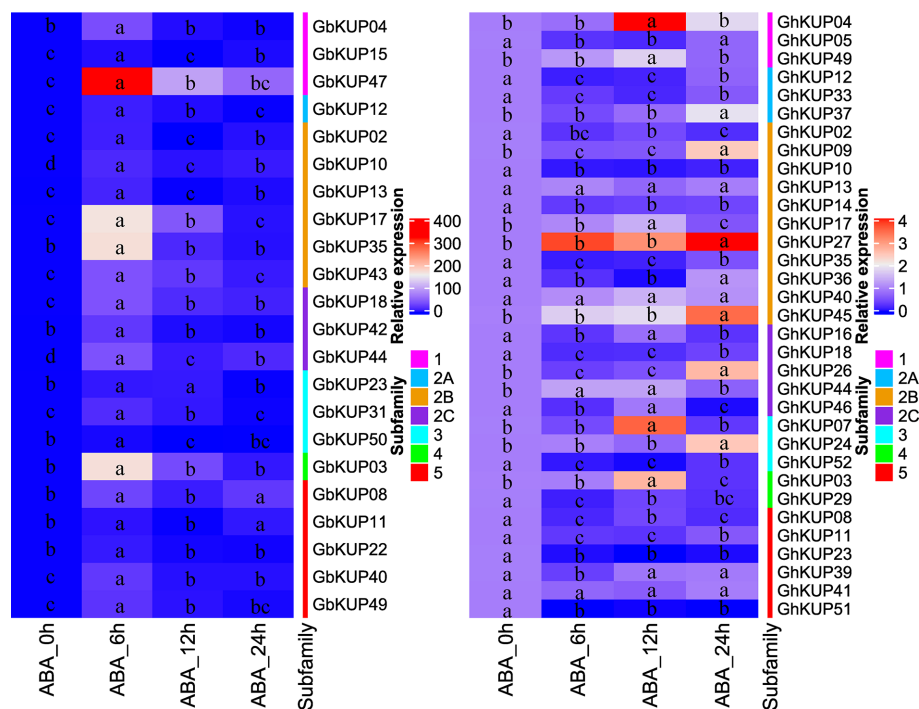
FIGURE 8 | Heat map representation of *GhKUP* and *GbKUP* expression levels under ABA stress. Two color bars represent relative expression values and KUP subfamilies. Expression levels with the same letter in each cell do not significantly differ at *P* < 0.05 as revealed by Duncan's multiple test.

Gene Loss During Cotton KUP Evolution

In this study, 31 orthologous gene groups of the KUP family were identified in four cotton species (**Figures 11A, B**; **Table S11**). A total of 20 orthologous gene groups in *G. hirsutum* and *G. barbadense* were also highly conserved in *G. arboreum* and *G. raimondii*, and two orthologous gene groups lost a *GaKUP* or *GrKUP*. In the allopolyploid species, two orthologous genes were found only in the GhAt and GbAt subgenomes, whereas the GhDt and GbDt subgenomes contained five unique orthologous genes. Meanwhile, only two orthologous gene groups were identified in *G. hirsutum* and *G. arboreum* or *G. raimondii*. Furthermore, six and seven *KUP*s were absent from the GhAt and GbAt subgenomes, respectively, whereas three *KUP*s were lost from the GhDt subgenome and four *KUP*s were lost from the GbDt subgenome. A total of eight GaKUPs, five GrKUPs, nine GhKUPs, and 11 GbKUPs were lost from the KUP family. Moreover, the lost *GhKUP*s or *GbKUP*s were distributed unevenly in the KUP subfamily (**Figure 11B**). A total of 13 genes were absent from subfamily 5. Two, three, and two *KUP*s were lost from subfamilies 1, 2A, and 2B, respectively. However, no gene loss occurred in subfamilies 2C, 3, and 4.

Evolutionary Analysis of the KUP Family in the Cotton Subgenomes

The evolutionary distances (Ka, Ks, and Ka/Ks ratio) of the 159 KUP gene pairs that were orthologous between the allopolyploid species and their diploid donor genomes were calculated (**Table S12**). The Ks values of the KUP members between *G. arboreum* and *G. raimondii* peaked at 0.0549, those between GbAt and GbDt peaked at 0.0391, and those between GhAt and GhDt peaked at 0.0334 (**Figure 11C**). The corresponding divergence times between *G. arboreum* and *G. raimondii* were 10.6 MYA, that between GbAt and GbDt were 7.5 MYA and that between GhAt and GhDt were 6.4 MYA. Ks values of the KUP members between GhAt and GhDt were significantly lower than those between *G. arboreum* and *G. raimondii*. The divergence time between the At subgenome and *G. arboreum* genome was approximately 1.38–2.27 MYA (Ks peaked at 0.0072 between the GhAt subgenome and *G. arboreum* genome and 0.0118 between the GbAt subgenome and *G. arboreum* genome). The divergence time between the Dt subgenome and *G. raimondii* genome was predicted to be 3.02–3.40 MYA (Ks peaked at 0.0157 between the GbDt subgenome and *G. raimondii* genome and 0.0177 between the GhDt subgenome and *G. raimondii* genome). In addition, the Ka, Ks, and Ka/Ks ratios between the At or Dt subgenome and their corresponding diploid donor genomes were similar in *G. hirsutum* and *G. barbadense* (**Figures 11D–F**).

DISCUSSION

In this study, we identified 51 GbKUPs and 53 GhKUPs in *G. hirsutum* and *G. barbadense*, respectively. These numbers were higher than those reported for any other investigated species
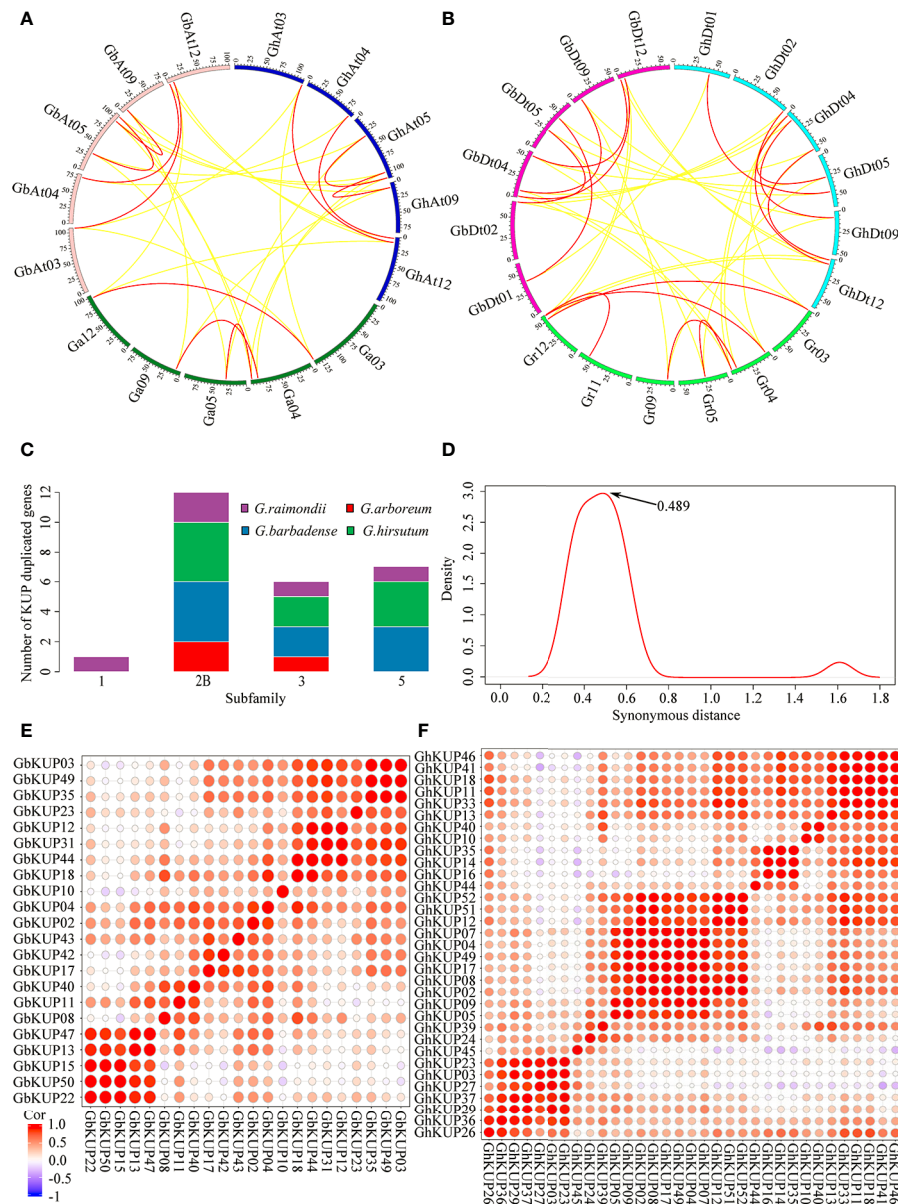
FIGURE 9 | Syntenic analysis, age distribution, and expression relationship of the duplicated *KUP* members in cotton. **(A, B)** Syntenic relationship of the duplicated *GhKUP*s and *GbKUP*s in the A-related genome **(A)** and D-related genome **(B)**. Red lines connect duplicated *KUP*s, and yellow lines connect orthologous genes in other species. The different colored sections of the circles indicate different genomes or subgenomes. **(C)** Subfamily distribution of the duplicated *KUP* members in cotton. **(D)** Age distribution of the duplicated *KUP*s in cotton based on Ks values. The peak value of the duplicated *KUP*s is marked with an arrow. **(E, F)** Heat map of the PCC of the expression profiles of the GbKUP **(E)** and GhKUP **(F)** families.

(**Figure 1**; **Tables S3**, **S4**). The increase in the KUP members of *G. hirsutum* and *G. barbadense* was associated with the allopolyploidization of the A and D diploid donor species (Wendel, 1989; Paterson et al., 2012; Wang et al., 2017). Although the A-related genome was larger than the D-related genome, the number of KUP members in the A-related genome was lower than that in the D-related genome likely because of the presence of additional LTR-type retrotransposons in the A-related genome (Li et al., 2014; Li et al., 2015; Du et al., 2018).

Additional LTR-type retrotransposons in the A-related genome might have resulted in the reduced protein-coding capacities of the A-related genome. Thus, the number of KUP members in the A-related genome was not larger than that in the D-related genome. The low distribution of members in the A-related genome is also exhibited by cotton CDK and PP2C families (Magwanga et al., 2018; Shazadee et al., 2019). Moreover, in contrast to the classification provided by previous studies, we found evidence for the existence of five KUP subfamilies with
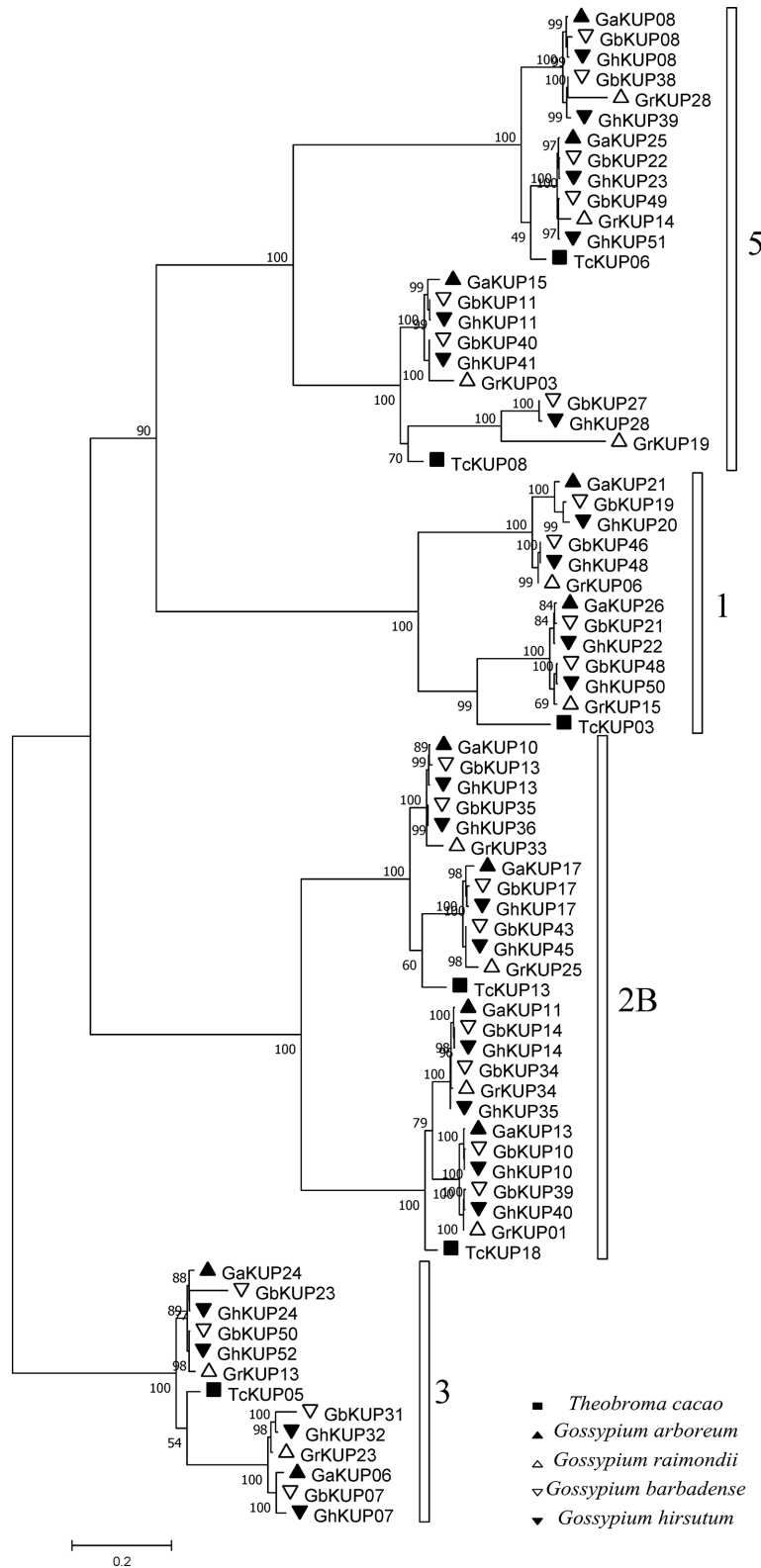
FIGURE 10 | Phylogenetic analysis of duplicated *GaKUPs*, *GrKUPs*, *GhKUPs*, and *GbKUPs* and their orthologous genes in cacao. The phylogenetic tree was constructed *via* the ML method. Numbers in clades represent bootstrap values, and KUP subfamilies are indicated by different letters. KUP members from the investigated species are marked by different shapes.
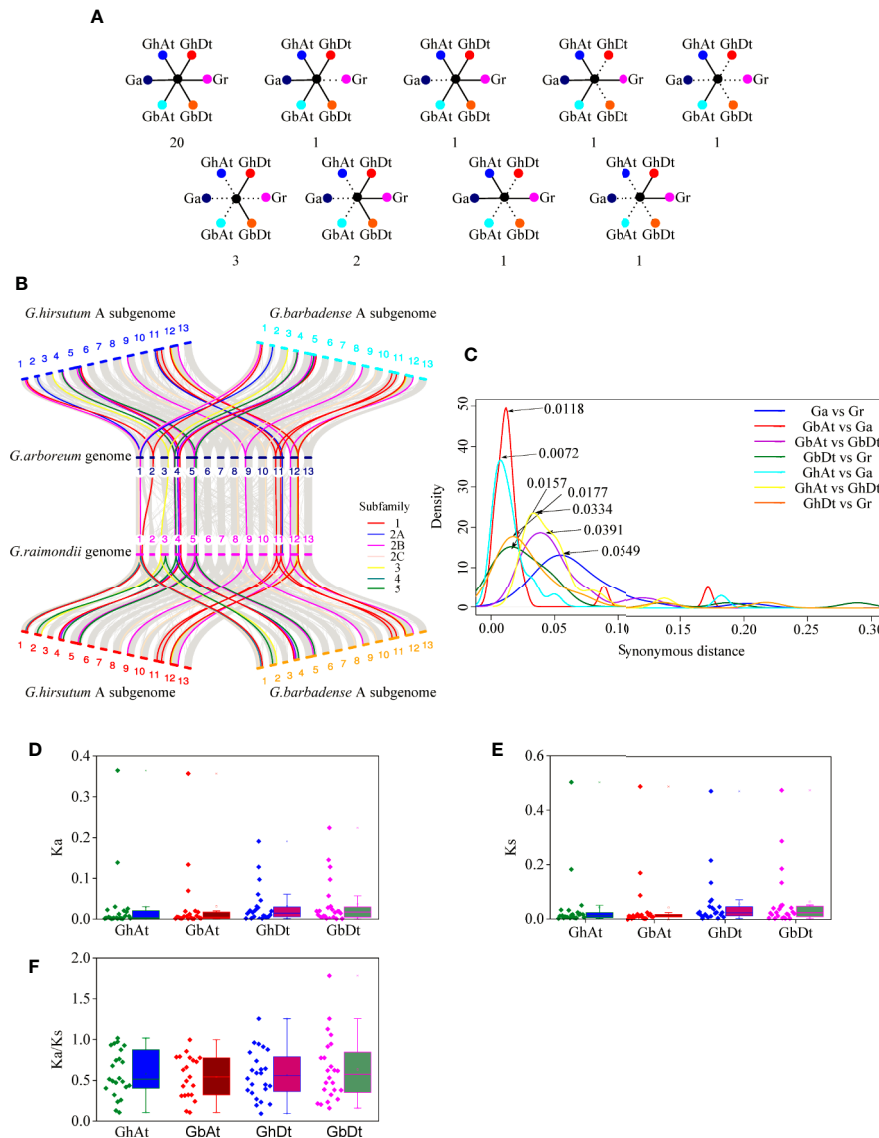
FIGURE 11 | Syntenic and evolutionary analyses of the KUP members of *G. hirsutum* (GhAt and GhDt), *G. barbadense* (GbAt and GbDt), *G. arboretum* (Ga), and *G. raimondii* (Gr). **(A)** Scenarios and numbers of the conserved KUPs in cotton. Solid lines represent extant genes, and dotted lines represent lost genes. The numbers beneath each drawing show the number of gene pairs found in four cotton genomes that fit the corresponding model. From left to right in the first line: KUPs present in all four genomes; KUPs not observed in Gr; KUPs not observed in Ga; KUPs not observed in GhDt and GbDt; and KUPs not observed in GhDt, GbDt, Ga and Gr. From left to right in the second line: KUPs not observed in GhAt, GbAt, Ga, and Gr; KUPs not observed in GhAt, GbAt and Ga; KUPs not observed in GbAt, GbDt, and GhDt; and KUPs not observed in GbAt, GbDt, GhAt, and Ga. **(B)** Microcollinearity analysis of genomic regions from four cotton species. Gray lines link orthologous gene pairs, and orthologous KUP members from different subfamilies are highlighted with different colored lines. **(C)** Ks distribution for the orthologous KUP gene pairs of four cotton species. Peak values for each comparison are marked with arrows. Blue, red, purple, dark green, cyan, yellow, and orange lines represent the Ks distribution of the orthologous KUP gene pairs of Ga and Gr, GbAt and Ga, GbAt and GbDt, GbDt and Gr, GhAt and Ga, GhAt and GhDt, and GhDt and Gr, respectively. **(D–F)** Distribution of Ka values **(D)**, Ks values **(E)**, and Ka/Ks ratio **(F)** between allopolyploid subgenomes and their corresponding ancestral species.

three groups in subfamily 2 (**Figures 1**, **2**). The subfamily that was regarded as group 3B in previous reports was identified as subfamily 5 in the present study due to its independent phylogenetic relationship with subfamily 3 (Cheng et al., 2018; Ou et al., 2018; Song et al., 2019). Meanwhile, each subfamily of *G. hirsutum* and *G. barbadense* had similar numbers of KUP members. Subfamilies 1, 2B, and 5 had the highest number of

members, whereas subfamily 4 contained the lowest number of KUPs. Similar distribution patterns of KUP members have been observed in other species (Cheng et al., 2018; Ou et al., 2018; Song et al., 2019). The distribution of conserved motifs, gene structures, and regulatory elements within subfamilies largely supported the results of phylogenetic analysis (**Figures 3**, **4**, **6**). Moreover, *GbKUP*s and *GhKUP*s were unevenly distributed

across cotton chromosomes, with the largest number of *KUPs* in At05 and Dt04 (**Figures S1-S3**; **Table S9**). Uneven chromosomal distributions have been also reported for the NAC, PP2C, and CDK gene families in cotton (Fan et al., 2018; Magwanga et al., 2018; Shazadee et al., 2019).

Expression patterns can reflect the possible functions of *GhKUPs* and *GbKUPs*. The high expression levels of most *GhKUPs* and *GbKUPs* in leaves and stems were suggestive of their roles in K^+ transport (**Figure 5**). The expression levels of the KUP family in cotton were similar to those in other species (Ahn et al., 2004; Gupta et al., 2008; Ou et al., 2018). Three *GhKUPs* and four *GbKUPs* showed significantly high expression levels only in roots. These genes may participate in K+ absorption. An orthologous gene of *GhKUP04*, *GhKUP05*, and *GbKUP04* in rice (*OsKUP02*) plays a major role in K+ absorption by roots (Yang et al., 2014), and another orthologous gene of *GbKUP12* in *Arabidopsis* (*AtKUP10*) may be related to the absorption of K+ and the establishment of root tip growth (Rigas et al., 2001). Moreover, the promoter regions of *GhKUPs* and *GbKUPs* contained numerous stress-responsive regulatory elements (**Figure 6**; **Table S8**). The significantly different expression levels of *GbKUPs* and *GhKUPs* under drought, salt, and ABA stresses indicated their possible functions in abiotic stress response (**Figures 7, 8**). Many KUP members in *Arabidopsis* (e.g., *AtKUP03*, *AtKUP06*, *AtKUP08*, and *AtKUP13*) and rice (e.g., *OsKUP02*, *OsKUP04*, *OsKUP10*, *OsKUP11*, *OsKUP13*, and *OsKUP17*) can participate in abiotic stress response (Ahn et al., 2004; Osakabe et al., 2013; Yang et al., 2014; Chen et al., 2015; Shen et al., 2015). These results indicated that the KUP family might be related to K+ uptake and efflux through ABA regulation during osmotic adjustment (Very and Sentenac, 2003; Fujita et al., 2011). In addition, *GhKUP27* and *GbKUP47* were highly upregulated under drought, salt, and ABA treatments. Therefore, *GhKUP27* and *GbKUP47* might play a vital role in abiotic stresses. Their corresponding orthologous genes in *Arabidopsis* and rice can respond to multiple abiotic stresses (Osakabe et al., 2013; Yang et al., 2014; Shen et al., 2015).

The expansion of the KUP family in *G. hirsutum* and *G. barbadense* was analyzed (**Figures 9A, B; Table S10**). Duplication events occurred more frequently in cotton than in cacao. The greater number of duplication events in cotton than in cacao partially explained the increase in the KUPs of four cotton species (**Figure 1**). Meanwhile, the KUP members of the two subgenomes expanded asymmetrically. Compared with the At subgenome, the Dt subgenome had more duplicated gene pairs, which resulted in additional *KUP* members. Furthermore, duplication events were not random across the KUP family, and expansion was preferred in subfamilies 2B, 3, and 5 (**Figures 9C, 10**). The orthologous genes of *GhKUPs* and *GbKUPs* that were duplicated in subfamilies 2B and 3 were also duplicated in their diploid donor species (**Figures 9A, B; Table S11**). In subfamily 5, the four orthologous genes of duplicated *GhKUPs* and *GbKUPs* were absent from the syntenic blocks of the diploid donor species. These results may be associated with genomic fractionation during cotton evolution (Wang et al., 2016). Preferential subfamily expansion was found in other families (Fan et al., 2018; Liu et al., 2019; Shazadee et al., 2019). Moreover, segmental duplication dominated the expansion of *KUP* genes in cotton. In this study, the

Ks peak was approximately 0.489, suggesting that the divergence times of all of the duplicated *KUPs*, except *GrKUP06/15*, corresponded to a paleohexaploidization event (**Figure 9D**; **Table S10**). Most of the *KUPs* that were duplicated in the allopolyploid species could be used to identify their orthologous duplicated genes in their diploid donor species (**Figures 9A, B**; **Table S11**). However, their orthologous genes were not duplicated in cacao. Phylogenetic analysis indicated that the duplication events of the KUP family might have occurred in the cotton ancestor (**Figure 10**). Thus, the duplicated *KUPs* in *G. hirsutum* and *G. barbadense* have happened in a common ancestor of *Gossypium*, and all of the duplication events occurred during the cotton paleohexaploidization event. In addition, the Ka/Ks ratios were less than 1 for all duplication events, indicating that the retained KUP members mainly experienced purifying selection (**Table S10**). The positive correlation among the expression levels of the duplicated *GhKUPs* and *GbKUPs* suggested that the duplicated genes maintained their original or similar functions during sequential evolution (**Figures 9E, F**) (Adams et al., 2003). Thus, subfunctionalization was the evolutionary fate of the duplicated *KUPs* in the cotton genome. The NAC and HMGS families underwent a similar evolutionary fate (Fan et al., 2018; Liu et al., 2019).

Postpolyploidization genomes may be reshuffled extensively partly due to an increase in homologous chromosomes (Gordon et al., 2009; Wang et al., 2009). Genome restructuring leads to large-scale chromosome loss, fusion, and fission (Buggs et al., 2012). After allopolyploid formation, *G. hirsutum* and *G. barbadense* tended to regain diploid heredity (Wang et al., 2016). More KUP members were lost from *G. hirsutum* and *G. barbadense* than from the diploid species (**Figure 11A**; **Table S11**). Gene loss accounted for the low numbers of the KUP members of *G. barbadense*. At11 and Dt5 in *G. barbadense* lost one KUP more than those in *G. hirsutum*. Meanwhile, gene loss was asymmetric in the At and Dt subgenomes. The high rate of gene losses in the At subgenome might have resulted from intensive human selection after domestication (Zhang et al., 2015; Hu et al., 2019). The high retention rates in the Dt subgenome might be related to abiotic stress tolerance (Zhang et al., 2015). The expression profile results also suggested that *GhKUPs* and *GbKUPs* could respond to abiotic stresses (**Figures 7, 8**). Furthermore, different KUP subfamilies had different loss rates (**Figure 11B**). Subfamily 5 showed the highest gene loss rate, and subfamilies 2C, 3, and 4 did not lose any genes. The different loss rates shown by KUP families resulted in unequal subfamily distributions.

The *Gossypium* genus underwent two major events: diploid species divergence (5–10 MYA) and interspecific hybridization (1–2 MYA; Zhang et al., 2015; Hu et al., 2019). The Ks distribution values provided evidence for the occurrence of these two major events in the KUP family (**Figure 11C**; **Table S12**). The Ks values between the GhAt and GhDt subgenomes were significantly lower than those between the *G. arboreum* and *G. raimondii* genomes. The significantly reduced Ks values might result from widespread homologous chromosome recombination and crop domestication in *G. hirsutum* (Wang M. et al., 2019). Moreover, the divergence time between the At subgenomes and *G. arboretum* genome was similar to that between the Dt subgenomes and *G. raimondii*

genome (**Figures 11D–F**). The At and Dt subgenomes contributed to fiber improvement and stress tolerance traits, respectively (Zhang et al., 2015). Thus, the KUP members of *G. hirsutum* and *G. barbadense* evolved fiber-improvement and stress-tolerant physiologies or phenotypes after domestication at the same time.

In summary, 51 GbKUPs and 53 GhKUPs were identified in *G. hirsutum* and *G. barbadense*, respectively. The KUP family could be divided into five subfamilies with subfamily preference. The duplication events of the KUP family in the cultured allopolyploid species originated from their diploid donor species and occurred during the cotton paleohexaploidization event. Subfunctionalization was the evolutionary fate of duplicated *GhKUP*s and *GbKUP*s. The KUP members in *G. hirsutum* and *G. barbadense* showed tissue-specific expression patterns and could respond to various stresses. Moreover, the KUP family in the At and Dt subgenomes of the allopolyploid species underwent asymmetric evolution. Meanwhile, *G. hirsutum* and *G. barbadense* exhibited conserved and divergent KUP evolution. The present study provided a comprehensive understanding of the KUP family in allopolyploid cotton species.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

KF, JH, and WXL designed the research. KF, ZM, JZ, YC, and ZL performed the experiments. KF, ZM, WWL, and YZ analyzed the data. KF, JH, and WXL wrote the paper with contributions from all the authors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.545042/full#supplementary-material

REFERENCES

Adams, K., Cronn, R., Percifield, R., and Wendel, J. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4649–4654. doi: 10.1073/pnas.0630618100

Ahn, S., Shin, R., and Schachtman, D. (2004). Expression of KT/KUP genes in Arabidopsis and the role of root hairs in K+ uptake. *Plant Physiol.* 134, 1135–1145. doi: 10.1104/pp.103.034660

Alemán, F., Caballero, F., Ródenas, R., Rivero, R., Martínez, V., Rubio, F., et al (2014). The F130S point mutation in the Arabidopsis high-affinity K+ transporter AtHAK5 increases K+ over Na+ and Cs+ selectivity and confers Na+ and Cs+ tolerance to yeast under heterologous expression. *Front. Plant Sci.* 5, 430. doi: 10.3389/fpls.2014.00430

Azeem, F., Ahmad, B., Atif, R., Ali, M., Nadeem, H., Hussain, S., et al. (2018). Genome-wide analysis of potassium transport-related genes in chickpea (Cicer arietinum L.) and their role in abiotic stress responses. *Plant Mol. Biol. Rep.* 36, 451–468. doi: 10.1007/s11105-018-1090-2

Bailey, T., Boden, M., Buske, F., Frith, M., Grant, C., Clementi, L., et al. (2009). MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Buggs, R., Chamala, S., Wu, W., Tate, J., Schnable, P., Soltis, D., et al. (2012). Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin. *Curr. Biol.* 22, 248–252. doi: 10.1016/j.cub.2011.12.027

Chen, G., Hu, Q., Luo, L., Yang, T., Zhang, S., Hu, Y., et al. (2015). Rice potassium transporter OsHAK 1 is essential for maintaining potassium-mediated growth and functions in salt tolerance over low and high potassium concentration ranges. *Plant Cell Environ.* 38, 2747–2765. doi: 10.1111/pce.12585

Chen, C., Chen, H., Zhang, Y., Thomas, H., Frank, M., He, Y., et al (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 8, 1193–1202. doi: 10.1016/j.molp.2020.06.009

Cheng, X., Liu, X., Mao, W., Zhang, X., Chen, S., Zhan, K., et al. (2018). Genome-wide identification and analysis of HAK/KUP/KT potassium transporters gene family in wheat (*Triticum aestivum* L.). *Int. J. Mol. Sci.* 19, 3969. doi: 10.3390/ijms19123969

Dhindsa, R., Beasley, C., and Ting, I. (1975). Osmoregulation in cotton fiber: accumulation of potassium and malate during growth. *Plant Physiol.* 56, 394–398. doi: 10.1104/pp.56.3.394

Du, X., Huang, G., He, S., Yang, Z., Sun, G., Ma, X., et al. (2018). Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat. Genet.* 50, 796–802. doi: 10.1038/s41588-018-0116-x

Elumalai, R., Nagpal, P., and Reed, J. (2002). A mutation in the Arabidopsis KT2/KUP2 potassium transporter gene affects shoot cell expansion. *Plant Cell* 14, 119–131. doi: 10.1105/tpc.010322

Fan, K., Li, F., Chen, J., Li, Z., Lin, W., Cai, S., et al. (2018). Asymmetric evolution and expansion of the NAC transcription factor in polyploidized cotton. *Front. Plant Sci.* 9:47:47. doi: 10.3389/fpls.2018.00047

Fan, K., Yuan, S., Chen, J., Chen, Y., Li, Z., Lin, W., et al. (2019). Molecular evolution and lineage-specific expansion of the PP2C family in *Zea mays*. *Planta* 250, 1521–1538. doi: 10.1007/s00425-019-03243-x

Fan, K., Chen, Y., Mao, Z., Fang, Y., Li, Z., Lin, W., et al. (2020). Pervasive duplication, biased molecular evolution and comprehensive functional analysis of the PP2C family in *Glycine max*. *BMC Genomics* 21, 465. doi: 10.1186/s12864-020-06877-4

Finn, R., Clements, J., and Eddy, S. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Fujita, Y., Fujita, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J. Plant Res.* 124, 509–525. doi: 10.1007/s10265-011-0412-3

Gordon, J., Byrne, K., and Wolfe, K. (2009). Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern

Saccharomyces cerevisiae genome. *PloS Genet.* 5, e1000485. doi: 10.1371/journal.pgen.1000485

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

Gupta, M., Qiu, X., Wang, L., Xie, W., Zhang, C., Xiong, L., et al. (2008). KT/HAK/KUP potassium transporters gene family and their whole-life cycle expression profile in rice (*Oryza sativa*). *Mol. Genet. Genomics* 280, 437. doi: 10.1007/s00438-008-0377-7

Han, M., Wu, W., Wu, W., and Wang, Y. (2016). Potassium transporter KUP7 is involved in K+ acquisition and translocation in Arabidopsis root under K+-limited conditions. *Mol. Plant* 9, 437–446. doi: 10.1016/j.molp.2016.01.012

Hartz, T., Johnstone, P., Francis, D., and Miyao, E. (2005). Processing tomato yield and fruit quality improved with potassium fertigation. *Hortscience* 40, 1862–1867. doi: 10.21273/HORTSCI.40.6.1862

He, C., Cui, K., Duan, A., Zeng, Y., and Zhang, J. (2012). Genome-wide and molecular evolution analysis of the Poplar KT/HAK/KUP potassium transporter gene family. *Ecol. Evol.* 2, 1996–2004. doi: 10.1002/ece3.299

Hu, Y., Chen, J., Fang, L., Zhang, Z., Ma, W., Niu, Y., et al. (2019). *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat. Genet.* 51, 739–748. doi: 10.1038/s41588-019-0371-5

Hyun, T., Rim, Y., Kim, E., and Kim, J. (2014). Genome-wide and molecular evolution analyses of the KT/HAK/KUP family in tomato (*Solanum lycopersicum* L.). *Genes Genom.* 36, 365–374. doi: 10.1007/s13258-014-0174-0

Kalyaanamoorthy, S., Minh, B., Wong, T., von Haeseler, A., and Jermiin, L. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587. doi: 10.1038/nmeth.4285

Kassambara, A. (2018). *ggcorrplot: Visualization of a Correlation Matrix using "ggplot2"*. R Package Version 0.1.2. Available at: https://CRAN.R-project.org/package=ggcorrplot.

Katoh, K., and Standley, D. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kim, E., Kwak, J., Uozumi, N., and Schroeder, J. (1998). AtKUP1: an Arabidopsis gene encoding high-affinity potassium transport activity. *Plant Cell* 10, 51–62. doi: 10.2307/3870628

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., Song, G., et al. (2014). Genome sequence of the cultivated cotton *Gossypium arboreum. Nat. Genet.* 46, 567–572. doi: 10.1038/ng.2987

Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R., et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33, 524–530. doi: 10.1038/nbt.3208

Liu, W., Zhang, Z., Zhu, W., Ren, Z., Jia, L., Li, W., et al. (2019). Evolutionary conservation and divergence of genes encoding 3-hydroxy-3-methylglutaryl coenzyme A synthase in the allotetraploid cotton species *Gossypium hirsutum*. *Cells* 8, 412. doi: 10.3390/cells8050412

Magwanga, R., Lu, P., Kirungu, J., Cai, X., Zhou, Z. , Wang, X., et al. (2018). Whole genome analysis of cyclin dependent kinase (CDK) gene family in cotton and functional evaluation of the role of CDKF4 gene in drought and salt stress tolerance in plants. *Int. J. Mol. Sci.* 19, 2625. doi: 10.3390/ijms19092625

Marchler-Bauer, A., Derbyshire, M., Gonzales, N., Lu, S., Chitsaz, F., Geer, L., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221

Mäser, P., Thomine, S., Schroeder, J., Ward, J., Hirschi, K., Sze, H., et al. (2001). Phylogenetic relationships within cation transporter families of Arabidopsis. *Plant Physiol.* 126, 1646–1667. doi: 10.2307/4280038

Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 418–426. doi: 10.1093/oxfordjournals.molbev.a040410

Osakabe, Y., Arinaga, N., Umezawa, T., Katsura, S., Nagamachi, K., Tanaka, H., et al. (2013). Osmotic stress responses and plant growth controlled by potassium transporters in Arabidopsis. *Plant Cell* 25, 609–624. doi: 10.2307/41812291

Ou, W., Mao, X., Huang, C., Tie, W., Yan, Y., Ding, Z., et al. (2018). Genome-wide identification and expression analysis of the KUP family under abiotic stress in cassava (*Manihot esculenta* Crantz). *Front. Physiol.* 9, 17. doi: 10.3389/fphys.2018.00017

Paterson, A., Wendel, J., Gundlach, H., Guo, H., Jenkins, J., Jin, D., et al. (2012). Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492, 423–427. doi: 10.1038/nature11798

Pettigrew, W. (2008). Potassium influences on yield and quality production for maize, wheat, soybean and cotton. *Physiol. Plant.* 133, 670–681. doi: 10.1111/j.1399-3054.2008.01073.x

Quintero, F., and Blatt, M. (1997). A new family of K+ transporters from Arabidopsis that are conserved across phyla. *FEBS Lett.* 415, 206–211. doi: 10.1016/s0014-5793(97)01125-3

Rigas, S., Debrosses, G., Haralampidis, K., Vicente-Agullo, F., Feldmann, K. A., Grabov, A., et al. (2001). TRH1 encodes a potassium transporter required for tip growth in Arabidopsis root hairs. *Plant Cell* 13, 139–151. doi: 10.1105/tpc.13.1.139

Rodríguez-Navarro, A. (2000). Potassium transport in fungi and plants. *BBA-Rev. Biomembr.* 1469, 1–30. doi: 10.1016/S0304-4157(99)00013-1

Santos, E., Pongrac, P., Reis, A., White, P., and Lavres, J. (2019). Phosphorus-zinc interactions in cotton: consequences for biomass production and nutrient-use efficiency in photosynthesis. *Physiol. Plant.* 166, 996–1007. doi: 10.1111/ppl.12867

Shazadee, H., Khan, N., Wang, J., Wang, C., Zeng, J., Huang, Z., et al. (2019). Identification and expression profiling of protein phosphatases (PP2C) gene family in *Gossypium hirsutum* L. *Int. J. Mol. Sci.* 20:1395. doi: 10.3390/ijms20061395

Shen, Y., Shen, L., Shen, Z., Jing, W., Ge, H., Zhao, J., et al. (2015). The potassium transporter OsHAK21 functions in the maintenance of ion homeostasis and tolerance to salt stress in rice. *Plant Cell Environ.* 38, 2766–2779. doi: 10.1139/gen-2018-0187

Song, Y., Ma, R., and Yu, M. (2015). Genome-wide analysis and identification of KT/HAK/KUP potassium transporter gene family in peach (*Prunus persica*). *Genet. Mol. Res.* 14, 774–787. doi: 10.4238/2015.January.30.21

Song, Z., Wu, X., Gao, Y., Cui, X., Jiao, F., Chen, X., et al. (2019). Genome-wide analysis of the HAK potassium transporter gene family reveals asymmetrical evolution in tobacco (*Nicotiana tabacum*). *Genome* 62, 267–278. doi: 10.1139/gen-2018-0187

Studio, R. (2012). *RStudio: integrated development environment for R* (Boston, Massachusetts: RStudio Inc). Available at: https://rstudio.com/.

Sustr, M., Soukup, A., and Tylova, E. (2019). Potassium in Root Growth and Development. *Plants* 8:435. doi: 10.3390/plants8100435

Tang, W., Tu, L., Yang, X., Tan, J., Deng, F., Hao, J., et al. (2014). The calcium sensor GhCaM7 promotes cotton fiber elongation by modulating reactive oxygen species (ROS) production. *New Phytol.* 202, 509–520. doi: 10.1111/nph.12676

Udall, J., Long, E., Hanson, C., Yuan, D., Ramaraj, T., Conover, J., et al. (2019). *De novo* genome sequence assemblies of *Gossypium raimondii* and *Gossypium turneri. G3: Genes Genom. Genet.* 9, 3079–3085. doi: 10.1534/g3.119.400392

Very, A., and Sentenac, H. (2003). Molecular mechanisms and regulation of K+ transport in higher plants. *Annu. Rev. Plant Biol.* 54, 575–603. doi: 10.1146/annurev.arplant.54.031902.134831

Véry, A., Nieves-Cordones, M., Daly, M., Khan, I., Fizames, C., and Sentenac, H. (2014). Molecular biology of K+ transport across the plant cell membrane: what do we learn from comparison between plant species? *J. Plant Physiol.* 171, 748–769. doi: 10.1016/j.jplph.2014.01.011

Wang, X., Tang, H., Bowers, J., and Paterson, A. (2009). Comparative inference of illegitimate recombination between rice and sorghum duplicated genes produced by polyploidization. *Genome Res.* 19, 1026–1032. doi: 10.1101/gr.087288.108

Wang, Y., Li, B., Du, M., Eneji, A., Wang, B., Duan, L., et al. (2012). Mechanism of phytohormone involvement in feedback regulation of cotton leaf senescence induced by potassium deficiency. *J. Exp. Bot.* 63, 5887–5901. doi: 10.1093/jxb/ers238

Wang, X., Guo, H., Wang, J., Lei, T., Liu, T., Wang, Z., et al. (2016). Comparative genomic de-convolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation. *New Phytol.* 209, 1252–1263. doi: 10.1111/nph.13689

Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q., et al. (2017). Asymmetric subgenome selection and *cis*-regulatory divergence during cotton domestication. *Nat. Genet.* 49, 579–587. doi: 10.1038/ng.3807

Wang, Y., Xu, J., Zhang, M., Tian, X., and Li, Z. (2018). GhKT2: a novel K$^+$ transporter gene in cotton (*Gossypium hirsutum*). *Front. Agr. Sci. Eng.* 5, 226–235. doi: 10.15302/J-FASE-2017170

Wang, M., Tu, L., Yuan, D., Zhu, D., Shen, C., Li, J., et al. (2019). Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nat. Genet.* 51, 224–229. doi: 10.1038/s41588-018-0282-x

Wang, M., Wang, Y., Li, B., Xiong, C., Eneji, A., Zhang, M., et al. (2019). The cotton high-affinity K$^+$ transporter, GhHAK5a, is essential for shoot regulation of K$^+$ uptake in root under potassium deficiency. *Plant Cell Physiol.* 60, 888–899. doi: 10.1093/pcp/pcz003

Wendel, J., and Albert, V. (1992). Phylogenetics of the cotton genus (*Gossypium*): character-state weighted parsimony analysis of chloroplast-DNA restriction site data and its systematic and biogeographic implications. *Syst. Bot.* 60, 115–143. doi: 10.2307/2419069

Wendel, J. (1989). New World tetraploid cottons contain Old World cytoplasm. *Proc. Natl. Acad. Sci. U.S.A.* 86, 4132–4136. doi: 10.1073/pnas.86.11.4132

Wickham, H., Winston, C.RStudio (2016). *ggplot2: create elegant data visualisations using the grammar of graphics.* Available at: https://CRAN.R-project.org/package=ggplot2.

Yang, T., Zhang, S., Hu, Y., Wu, F., Hu, Q., Chen, G., et al. (2014). The role of a potassium transporter OsHAK5 in potassium acquisition and transport from roots to shoots in rice at low potassium supply levels. *Plant Physiol.* 166, 945–959. doi: 10.1104/pp.114.246520

Zhang, Z., Zhang, J., Chen, Y., Li, R., Wang, H., and Wei, J. (2012). Genome-wide analysis and identification of HAK potassium transporter gene family in maize (*Zea mays* L.). *Mol. Biol. Rep.* 39, 8465–8473. doi: 10.1007/s11033-012-1700-2

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207