*Article*

# *Magnesium transporter* Gene Family: Genome-Wide Identification and Characterization in *Theobroma cacao*, *Corchorus capsularis*, and *Gossypium hirsutum* of Family Malvaceae

**Parviz Heidari[1,*], Abdullah[2], Sahar Faraji[3], and Peter Poczai [4,5,6,*]**

[1] Faculty of Agriculture, Shahrood University of Technology, Shahrood 3619995161, Iran
[2] Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan; abd.ullah@bs.qau.edu.pk
[3] Department of Plant Breeding, Faculty of Crop Sciences, Sari Agricultural Sciences and Natural Resources University (SANRU), Sari 4818166996, Iran; sahar.faraji@rocketmail.com
[4] Finnish Museum of Natural History, P.O. Box 7, FI-00014 University of Helsinki, Finland
[5] Faculty of Biological and Environmental Sciences, PO Box 65 FI-00065 University of Helsinki, Finland
[6] Department of B
* Correspondence: heidarip@shahroodut.ac.ir; Tel.: +98-912-0734-034.(P.H.); Péter Poczai (peter.poczai@helsinki.fi (P.P.)

**Abstract:** Magnesium (Mg) is a vital element, which involves in various key cellular processes in plants. Mg transporter (MGT) genes play an important role in magnesium distribution and ionic balance maintenance. Here, MGT family members were identified and characterized in three species of the plant family Malvaceae, including *Theobroma cacao*, *Corchorus capsularis*, and *Gossypium hirsutum* to improve our understanding of their structure, regulatory systems, functions, and possible interactions. We identified 18, 41, and 16 putative non-redundant *MGT* genes from the genome of *T. cacao*, *G. hirsutum*, and *C. capsularis*, respectively, which were clustered into five groups within maximum likelihood tree. Several segmental/tandem duplication events were determined between *MGTs* genes. It seems that *MGTs* were slowly evolving and have been evolved under a purifying selection. Analysis of gene promoter regions showed that *MGTs* have a high potential to respond to biotic/abiotic stresses as well as hormones. The expression patterns of *MGT* genes revealed their possible *T. cacao*role in response to *P. megakarya* fungi in *T. cacao*, whereas *MGT* genes showed differential expression in various tissues and respond to several abiotic stresses, including cold, salt, drought, and heat stress in *G. hirsutum*. Besides, the co-expressions network of MGTs indicated that genes involved in auxin responsive, lipid metabolism, cell wall organization, and photoprotection can interact with MGTs.

**Keywords:** Magnesium transporter; Comparative analysis; Malvaceae; *Theobroma, Gossypium; Corchorus*; Expression analysis; Gene structure; Phylogenetic analysis.

## 1. Introduction

Magnesium (Mg) is one of the critical bimetals which regulate the biochemical processes and provide stability to membrane in plants [1,2]. Magnesium act is a cofactor for polymerase, kinase, and H+-ATPase which are necessary to synthesize proteins, nucleic acid, and generate energy [3,4]. Moreover, it is also required to maintain the homeostasis of cation-anion in the cell [5]. Various types of adverse effects have been reported on the plant during deficiency of Mg, including a reduction in photosynthesis, macromolecules synthesis, and plant growth and development [6–8]. Therefore, plants adapted an efficient transport system for absorption, storage, and Mg translocation [2]. The *Mg transporter* (*MGT*) gene family, also known as MRS2 or CorA, significantly plays the aforementioned

essential functions [9,10]. Members of the MTG family are defined by two transmembrane domains in which a tripeptide motif GMN (Glycine-Methionine-Asparagine) exists at the C-terminal terminal domain of the first transmembrane [11,12]. *MGTs* are expressed in root tissues of plants, which are more involved in up taking Mg such as *MGT1* in rice and *MGT6* in Arabidopsis, transferring of Mg from root to shoot (such as *MGT9* in Arabidopsis), homeostasis by maintaining ionic balance (such as *MGT10* in Arabidopsis), accumulation and translocation of Mg within the parts of the cell such as vacuole, *MGT2* and *MGT3* in Arabidopsis [11,13–16]. These genes are also crucial for pollen mitosis and pollen intine formation [15,17,18].

*MGT* genes also respond to changes in elemental concentration in soil, i.e., *MGT* genes showed high expression due to Aluminum (Al) toxicity in acidic soil. It was observed in Arabidopsis and maize that Al-tolerant genotypes have a high capability of Mg uptake and accumulation [9,19]. The transformation and expression of Arabidopsis *MGT1* in *Nicotiana benthamiana* increase uptake of Mg and reduced toxicity of Al in transgenic lines [20]. In contrast, the knockout of *MGT1* in rice reduced tolerance to salt and was found linked to a high content of sodium in shoot tissues [21]. The *MGT* genes are also important for plant adaptation to changing Mg status in soil [13]. The genes of *MGT* family has been identified and characterized in several plant species, including Arabidopsis [22], rice [11], maize [9], pear [23], citrus [24], rapeseed[25], wild sugarcane[26], and tomato [27]. The plant family Malvaceae is one of the largest dicot families comprises 244 genera and 4,225 species [28]. The family include significant economic plant species such as cotton (Gossypium) and jute (Corchorus) which are important for fiber, whereas cacao (*Theobroma cacao* L.) is important for chocolate production [29,30]. To date, none of the studies, to the best of our knowledge, focus on the characterization of the *MGT* gene family despite the availability of nuclear genomes of the species of family Malvaceae [31,32] with the advancement of sequencing technologies. Here, we aim to: (i) identify and characterize *MGT* genes within three species of Malvaceae including *T. cacao*, *C. capsularis*, and *G. hirsutum*, (ii) study evolutionary pattern and phylogenetic relationship, (iii) determining roles of MGTs in growth and development of Malvaceae species.

## 2. Materials and Methods

### *2.1. Identification and characterization of MGT genes in T. cacao, C. capsularis, and G. hirsutum*

The MGT proteins of Arabidopsis were BLAST in Ensembl Plants [33] for *T. cacao*, and *Corchorus capsularis*, and cotton genome database [34] for *Gossypium hirsutum* to identify *MGT* genes and retrieve proteins, coding sequences, genomic sequences, and promoter regions for various analyses following the previous approach [35]. The non-redundant protein sequences were selected based on CDD search (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). The ProtParam program [36] was used to determining molecular weight (MW) and isoelectric points (*pI*), TMHMM Server v. 2.0 [37] to predict transmembrane domains, and CELLO2GO [38] to determine the location of MGT proteins in the cell. The PlantCARE [39] was used to analyze promoter regions.

### 2.2. Phylogenetic inference, conserve protein motifs, and gene structure

The protein sequences of *T. cacao*, *C. capsularis*, and *G. hirsutum* was aligned in Geneious R8.1[40] using Clustal W [41] and was analyzed in IQ-tree for construction of maximum likelihood phylogenetic tree under default parameters and 1000 bootstrap replications [42–44]. The integrative tree of life version.4 [45] was used for the improvement of tree representation. The MEME (Multiple Em for Motif Elicitation) server [46] was used to identify the conserved protein motifs in MGT. The gene structure of each *MGT* gene was constructed using the Gene Structure Display Server [47].

*2.3. Gene duplications and synteny analysis*

The *MGT* genes with more than 85% identity in each species were selected as duplicated genes. Then, the location diagram of duplicated genes in cacao and *G. hirsutum* were constructed using TBtools [48]. In addition, the synonymous (Ks) and non-synonymous (Ka) of each duplicated gene pair were calculated by DnaSP v6 software [49]. Finally, the time of divergence of duplicated genes was determined using the following equation, $T = (Ks/2\lambda) \times 10^{-6}$ [50]. In equation, $\lambda$ is substitutions per synonymous site per year and $\lambda = 6.5 \times 10^{-9}$. Moreover, the synteny relationships diagrams of *MGTs* genes among the orthologous pairs of *T. cacao-G. hirsutum*, and *T. cacao- C. capsularis* were created using Circos software [51].

*2.4. Structure analyses of MGT proteins*

The transmembrane and three-dimensional structures of the candidates of MGT subgroups, MRS2, NIPA, and CorA proteins, in *T. cacao*, *G. hirsutum*, and *C. capsularis* were predicted using the Phyre2 server [52], whereas docking analysis was performed to predict the ligand-binding regions (pocket sites) using DeepSite [53] and CASTp [54] tools and finally constructed in PyMOL [55].

*2.5. Expression analysis of tcMGTs and ghMGTs using RNA-seq data*

In the present study, the publicly available RNA-seq data of cacao transcriptome with accession number GSE11604 were retrieved from NCBI to determine the response and expression of cacao *MGTs* in response to a fungi disease caused by *Phytophthora megakarya* after inoculation of different time courses, 0h (hour), 6h, 24h, 48h, and 72h, in two contrasting cultivars, Nanay (NA-32), as a susceptible cultivar and Scavina (SCA-6) as fungal resistant cultivar. Finally, the expression patterns of *tcMGTs* were illustrated in heatmaps based on log2 transformed using TBtools [48]. Furthermore, expression profile of *ghMGTs* in various tissues (ovule, fiber, anther, bract, filament, leaf, pental, root, sepal, stem, and torus) and in response to various abiotic stresses (cold, heat, salt, and drought) was retrieved from available RNA-seq data of cotton genome database (https://cottonfgd.org/) under project PRJNA490626 using the gene ID of each gene of newly assembled genome as query [56]. The expression of *ghMGTs* in various tissues were analyzed and represented as heatmap based on percentage expression of each gene using TBtools [48] while the data of abiotic stresses were analyzed and represented through heatmap after log2 transformation through TBtools [48].

## 3. Results

*3.1. Sequence and structure of MGT genes*

In the present study, 18, 41, and 16 putative non-redundant *MGT* genes were identified from the genome of *T. cacao*, *G. hirsutum*, and *C. capsularis*, respectively. All sequences (genomic, amino acids, coding sequences) of identified *MGT* genes were shown in Table S1. MGTs were characterized based on their sequences structure (Table S2) and three MGT sub-groups, including MRS2, NIPA, and CorA were recognized according to the specific domain distribution, (Table S2). Our findings revealed that MGTs in three studied plant species are diverse in terms of sequence length, molecular weight (MW), the isoelectric point (pI), and exon number (Table S2). For instance, protein length varied from 321 amino acids to 632 amino acids (aa) in *T. cacao*, from 210 aa to 474 aa in *G. hirsutum*, and 262 to 2417 aa in *C. capsularis* (Table 1). Besides, the predicted MW was ranging from 32.75 kDa to 70.91 kDa in *T. cacao*, 32.66 kDa to 53.95 kDa in *G. hirsutum*, and 29.82 kDa to 268.42 kDa in *C. capsularis* (Table 1). Moreover, the pI of MGTs ranged between 4.48 and 8.57 in *T. cacao*, between 4.76 and 9.57 in *G. hirsutum*, and between 4.79 and 8.60 in *C. capsularis* (Table 1). Based on pI value, 75% of MGTs in *C. capsularis*, 56% in *T. cacao* and 49% in *G. hirsutum* were predicted as acidophilic protein (Table S2). In addition, the prediction of subcellular localization illustrated that most MGTs are located in endomembrane and

plasma membrane (Table S2). The exon number of MGT genes was varied between 4 and 15 in *T. cacao* and *G. hirsutum*, while the exon number was varied from 4 to 21 in *C. capsularis* (Table 1).

**Table 1.** Summary of MGTs properties in three studied plant species, including *Theobroma cacao, Gossypium hirsutum*, and *Corchorus capsularis.*

| Organism | Gene number | Gene length (bp) | Protein length (aa) | MW (KDa) | pI | Exon number |
|---|---|---|---|---|---|---|
| *T. cacao* | 18 | 1212-2632 | 321-632 | 32.75-70.91 | 4.48-8.57 | 4-15 |
| *G. hirsutum* | 41 | 633-1425 | 210-474 | 32.66-53.95 | 4.76-9.57 | 4-15 |
| *C. capsularis* | 16 | 789-7254 | 262-2417 | 29.82-268.42 | 4.79-8.60 | 4-21 |

*3.2. Phylogenetic analysis and classification of the MGT gene family*

In this study, a phylogenetic tree of MGT proteins was constructed, comprising 18 tcMGT proteins from *T. cacao*, 41 ghMGTs from *G. hirsutum*, and 16 ccMGTs from *C. capsularis*. The MGT proteins were clustered into five groups (Groups I, II, III, IV, and V; Figure 1). Six MGTs, including a CorA of Jute, ccMGT08, a MRS2 protein from cacao, tcMGT03, along four MRS2 proteins of cotton, ghMGT11, ghMGT12, ghMGT32, and ghMGT33 with similar structure contained 11 exons were located in group I (Figure 1a, b). Besides, five MGT proteins were located in group II and four CorA proteins, tcMGT01, tcMGT16, ccMGT02, and ccMGT16 were clustered in group III. In addition, 16 MRS2 proteins along with two CorA proteins were located in group IV. Most MGTs were clustered in group V and all NIPA type proteins were located in this group (Figure 1a, b). In addition, MGTs were analyzed based on distribution of conserved motifs in their protein sequence. In the present study, 35 conserved motifs were identified and MGTs from group V showed more diversity than other groups (Figure 2). Besides, motifs 12, 5, 18, 7, 10, and 2 frequently observed in MGTs and NIPA proteins illustrated different patterns of conserved-motifs distribution. Moreover, CorA proteins (group III) also showed various conserved motifs (Figure 2).

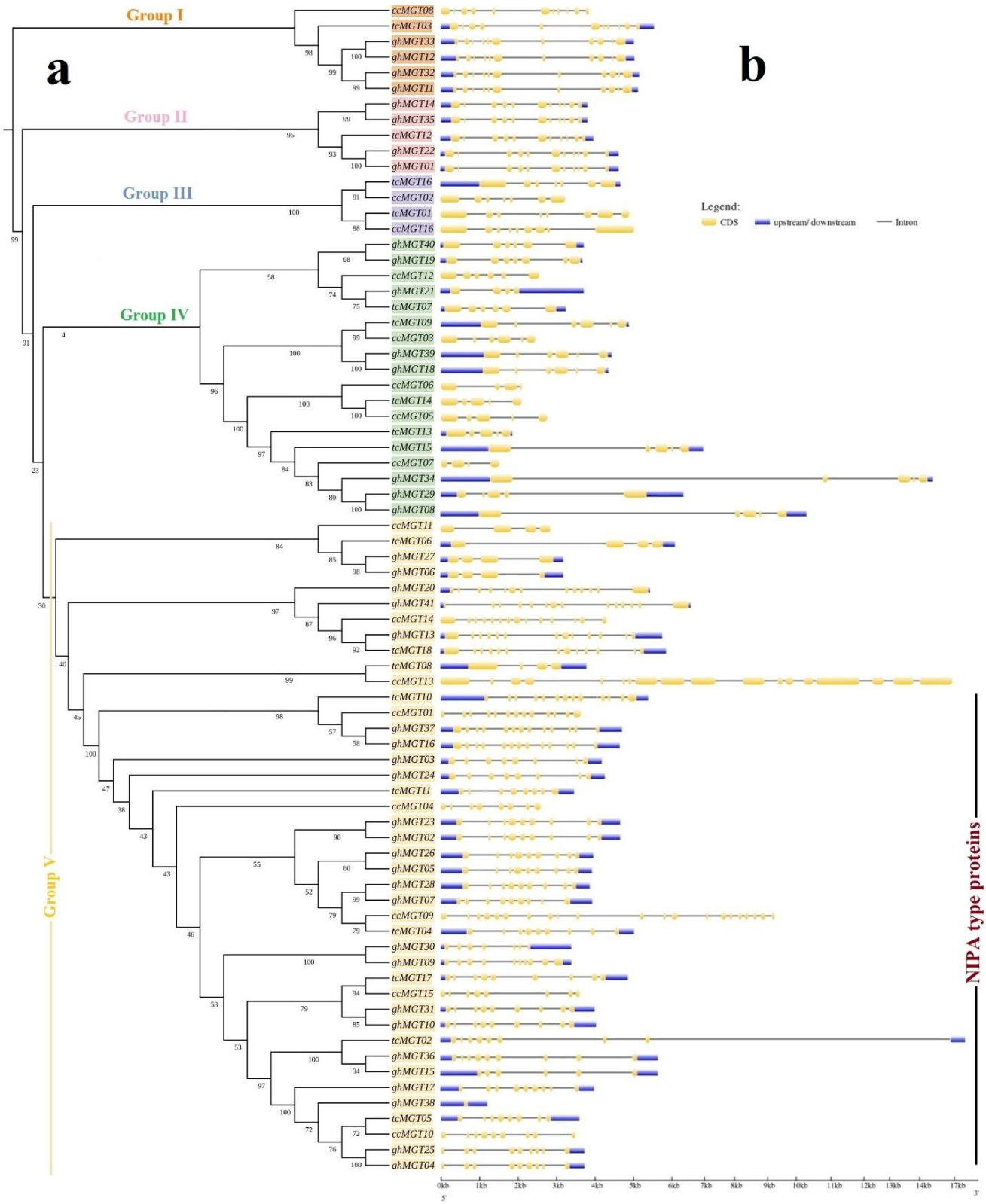


**Figure 1.** The phylogenetic tree of MGT proteins (a) and gene structure of *MGT* genes (b) of *Theobroma cacao*, *Gossypium hirsutum*, and *Corchorus capsularis*. The start of each gene name for the species as: tc: *Theobroma cacao*; gh: *Gossypium hirsutum*; cc; *Corchorus capsularis*.

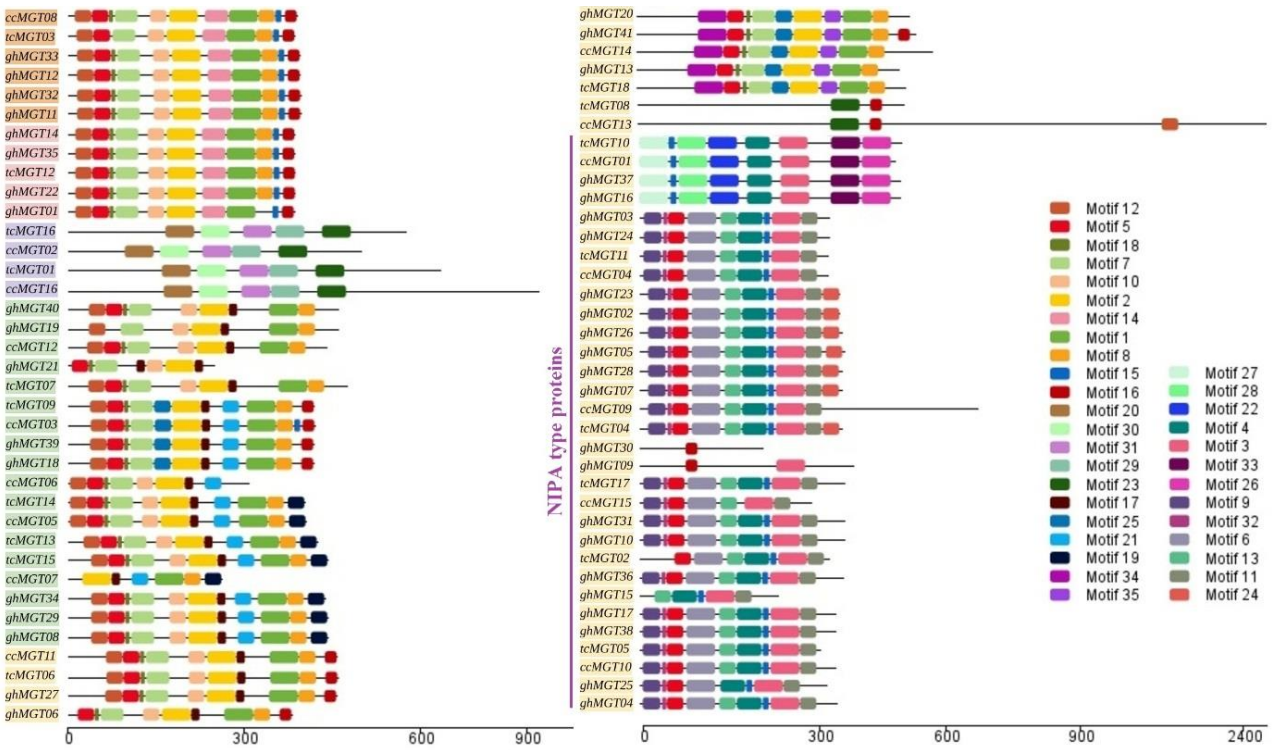


**Figure 2.** Conserved motifs distribution in MGT proteins of *Theobroma cacao*, *Gossypium hirsutum*, and *Corchorus capsularis*.

*3.3. Duplication events and synteny analysis*

The *tcMGT* genes were mapped onto the 5 out of 10 chromosomes in the cacao genome (Figure 3a), while *ghMGTs* were distributed in 22 of 26 chromosomes of *G. hirsutum* genome (Figure 3b). Due to the incomplete physical map of the *C. capsularis* plant, the mapping of *ccMGT* genes is not provided. In cacao genome, chromosome 2 encompassed the most significant number of *tcMGTs* with five genes, while in *G. hirsutum* genome, the most significant number of *ghMGTs* were located on chromosomes A04 and D01 (Figure 3). In addition, the duplication events of *MGT* genes in selected plant species were investigated. Eight segmental-duplicated gene pairs were identified between 12 *tcMGT* genes of cacao (Figure 3a and Table S3). It seems that tandem duplication events occurred on chromosome 6 between three *tcMGT* genes of cacao including, *tcMGT013*, *tcMGT14*, and *tcMGT15* (Figure 3a). Besides, four segmental-duplicated events were predicted for *tcMGT03* gene (Table S3). Notably, a duplication event occurred around 10 MYA between two *CorA like* genes in cacao, *tcMGT01* and *tcMGT16*. In addition, five segmental-duplicated gene pairs were recognized between *ccMGT* genes, and a triplication event was predicted between *ccMGT05* and *ccMGT06*, *ccMGT07*, and *ccMGT04* (Table S3). According to the *Ka*/*Ks* ratio, the first duplication was approximately 103 million years ago (MYA) between *ccMGT05* and *ccMGT04*. The most duplication events were observed between *ghMGT* genes in *G. hirsutum* genome with 22 segmental-duplicated gene pairs. Moreover, four *ghMGT* genes, *ghMGT02*, *ghMGT05*, *ghMGT23*, and *ghMGT26*, had a common ancestor and probably the first duplication event occurred approximately 68 MYA between *ghMGT23* and *ghMGT26* (Table S3). The intraspecies synteny of *MGT* genes was constructed between *T. cacao* and *G. hirsutum*, and *T. cacao* and *C. capsularis* (Figure 4). The 18 *tcMGT* genes in *T. cacao* illustrated 15 and 16 syntenic blocks relationships with *MGT* genes in the *G. hirsutum* and *C. capsularis*, respectively (Figure 4a, b). Interestingly, tcMGT genes of cacao showed more syntenic relationships with *ghMGTs* of D-genome than the A-genome of *G. hirsutum*.

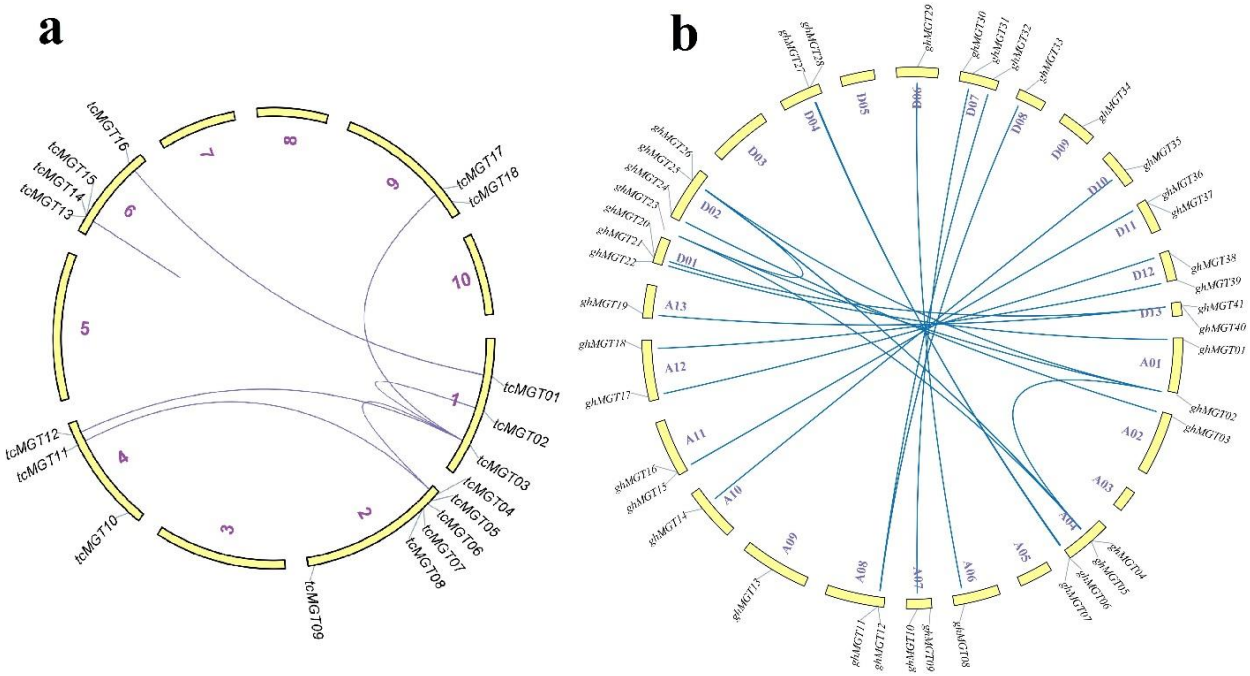


**Figure 3.** Location of *MGT* genes on the chromosome in *Theobroma cacao* (a), and *Gossypium hirsutum* (b). The duplicated genes are connected using blue lines.

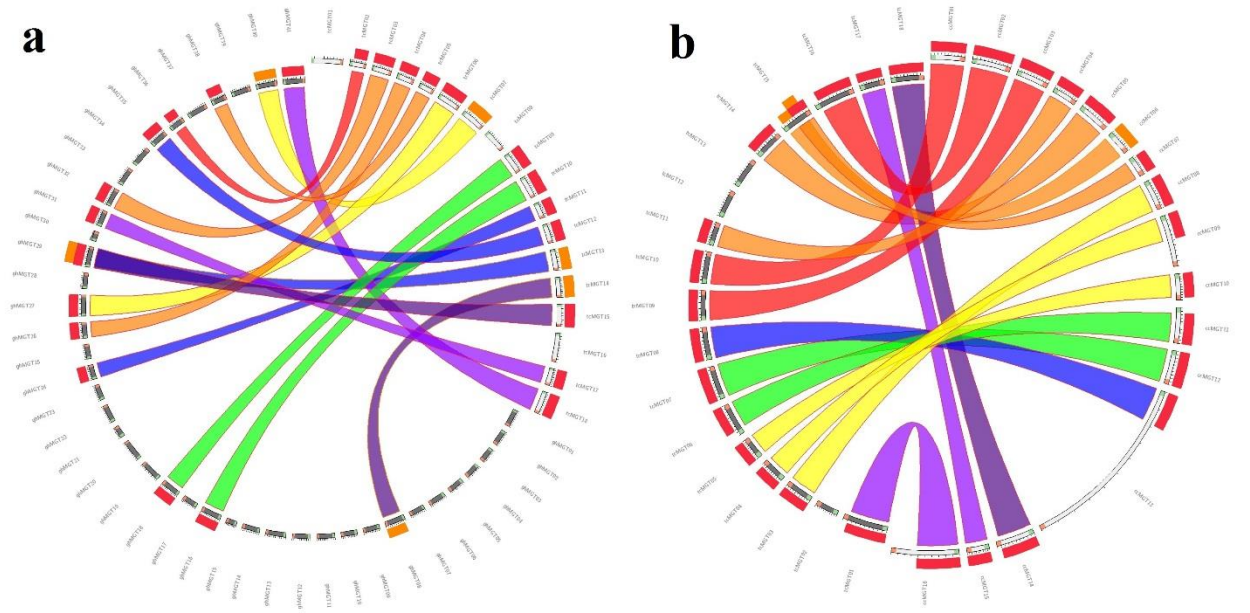


**Figure 4.** Synteny analysis of *MGT* genes. The syntenic blocks of cacao *MGTs* are constructed with *Gossypium hirsutum* (a), and *Corchorus capsularis* (b).

*3.4. Protein structure and docking analysis*

The 3D structures of all the candidates of three types of MGTs, NIPA, MRS2, and CorA, were predicted in *T. cacao, G. hirsutum* and *C. capsularis* (Figure S1). Nine $\alpha$-helixes were observed in the predicted 3D structure of NIPAs in three studied plants, while fewer $\alpha$-helixes were predicted in the structure of MRS2 and CorA proteins (Figure S1). Besides, nine transmembrane helices with eight pores were predicted in the structure of NIPAs in three plants while in candidate MRS2 proteins, two transmembrane helices were observed in all studied plant species. However, both N-terminal and C-terminal of candidate MRS2 proteins from *T. cacao* and *C. capsularis* were

predicted in extracellular part while in *G. hirsutum*, both N- and C-terminal were observed cytoplasmic part. In the candidate CorA protein, three transmembrane helices were predicted in *G. hirsutum,* and two transmembrane helices were predicted in *T. cacao* and *C. capsularis* (Figure 5).

Moreover, pocket sites of MGT proteins related to the active binding site were predicted into structures of candidate proteins. The results illustrated that sub-groups of MGT proteins are different based on the residues present in predicted pocket sites (Figure 6). Phenylalanine (PHE) amino acid was frequently observed in binding sites of NIPA proteins from *T. cacao* and *G. hirsutum*, while in *C. capsularis*, isoleucine (ILE) and lysine (LYS) were more observed in pocket sites. In candidate MRS2 proteins, proline (PRO), PHE, glutamine (GLN), asparagine (ASN), glutamic acid (GLU), and glycine (GLY) were frequently predicted in *T. cacao* as binding sites, while in *G. hirsutum*, tyrosine (TYR), leucine (LEU), GLU, GLY, and GLN were more repeated in pocket sites. Besides, LEU residue was highly observed as key binding sites in the candidate ccMGT of *C. capsularis*. In the candidate CorA proteins, GLU amino acid was more predicted in pocket sites of *T. cacao* and *G. hirsutum*, while PHE was frequently observed in pocket sites of candidate CorA protein in *C. capsularis*.

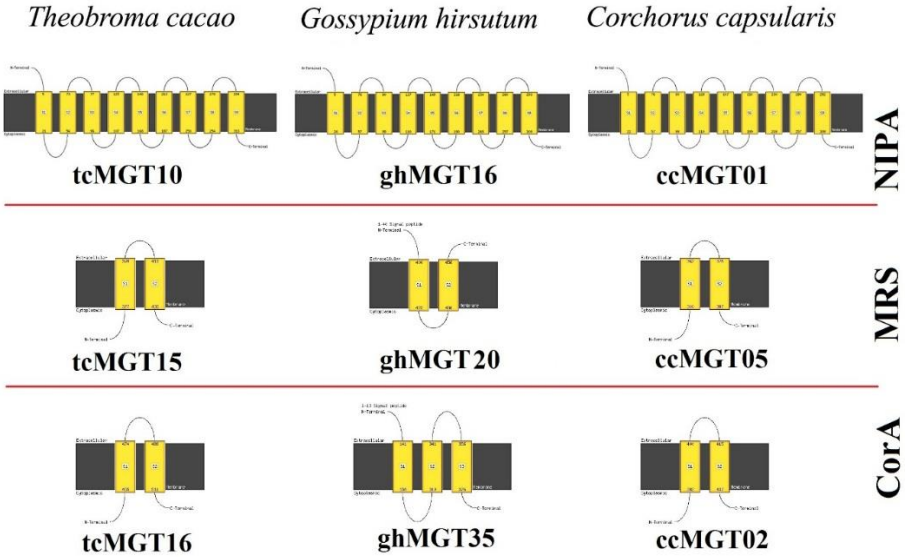


**Figure 5.** The predicted transmembrane helices in sub-groups of MGTs, NIPA, MSR2, and CorA, in *Theobroma cacao*, *Gossypium hirsutum*, and *Corchorus capsularis*

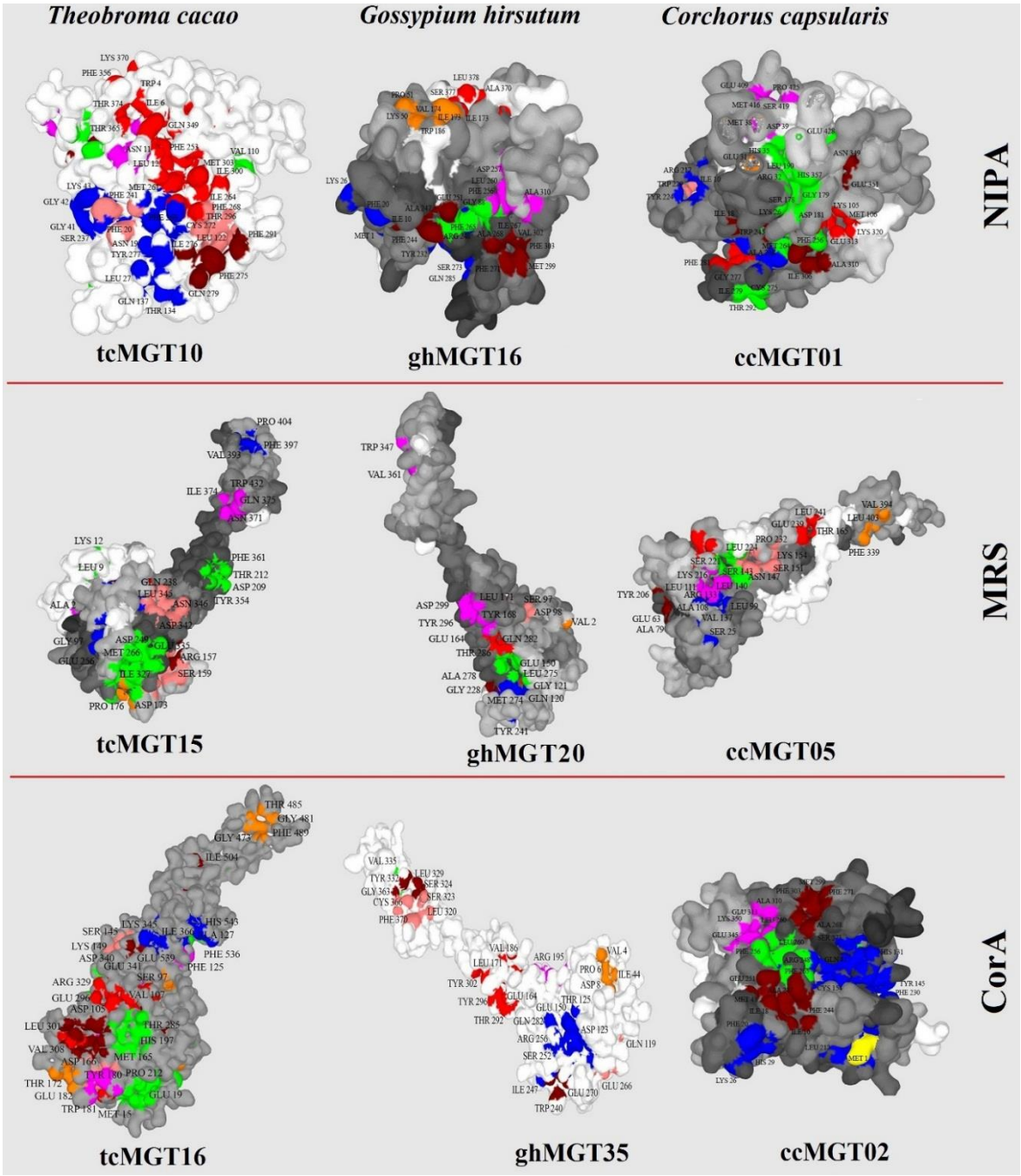


**Figure 6.** Docking analysis of candidates of sub-groups of MGTs, NIPA, MSR2, and CorA, in *Theobroma cacao*, *Gossypium hirsutum*, and *Corchorus capsularis*

*3.5. Distribution of cis-regulatory elements in promoter region of MGT genes*

In the present study, the promoter region of *MGTs* in three plant species, including *T. cacao, G. hirsutum* and *C. capsularis*, were analyzed and compared based on type and frequency of cis-regulatory elements. All recognized elements were classified in five groups, including hormone-responsive elements, stress responsive, light-responsive elements, growth -responsive elements, and binding sites of transcription factors. Our results revealed that *MGT*' promoters contain more cis-regulatory elements related to stress response (Figure 7 and Table S4). In addition, all stress-responsive elements were grouped in six classes related to drought, wounding, anaerobic, low temperature, biotic stress, and general stresses. However, percentage of cis-regulatory elements responsive to low temperature, anaerobic, and biotic stresses was more observed in promoter region of *tcMGT* genes (Figure 7b). Besides, regulatory elements related to response to abscisic acid (ABA),

salicylic acid (SA), auxin, gibberellin (GA), and methyl jasmonate (MeJA) were observed among hormone-responsive elements (Figure 7b, f, and i). We found that *MGT* genes might be more induced in response to ABA hormone.

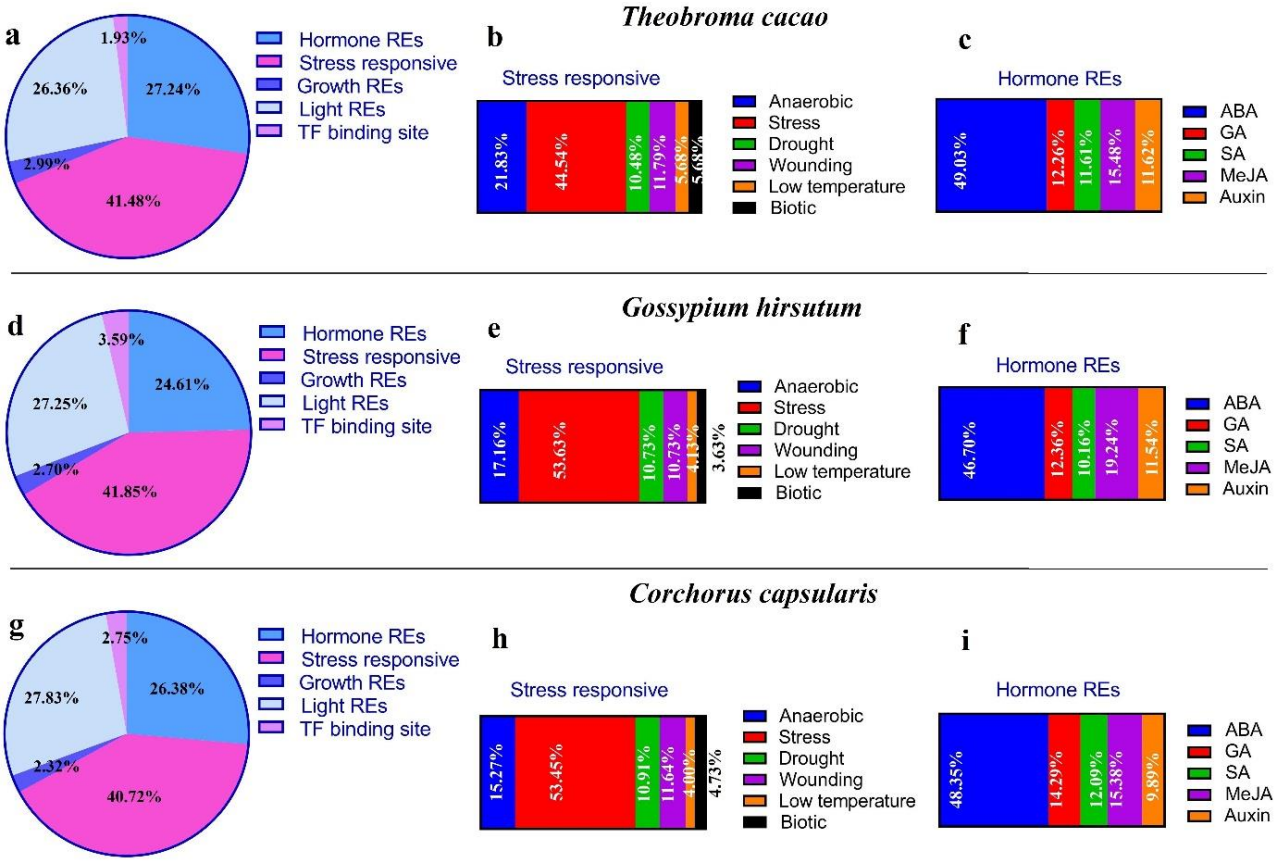


**Figure 7.** Percentage of cis-regulatory elements in promoter regions of *MGT* genes. Classification of identified cis-regulatory elements in term of hormone responsive elements (REs), stress responsive, growth REs, light Res, and transcription factor (TF) binding site in *Theobroma cacao* (a), *Gossypium hirsutum* (d), and *Corchorus capsularis* (g). Percentage of cis-regulatory elements related to stress responsiveness in *Theobroma cacao* (b), *Gossypium hirsutum* (e), and *Corchorus capsularis* (h). Percentage of different groups of hormone-related cis-regulatory elements in *Theobroma cacao* (c), *Gossypium hirsutum* (f), and *Corchorus capsularis* (i).

### *3.6. Expression profile of tcMGT genes*

In the current study, the expression levels of *tcMGTs* were also provided in response to *P. megakarya* after 0h, 6h, 24h, and 72h after infection using available RNA-seq data of two contrasting genotypes *T. cacao* Nanay (fungal susceptible cultivar) and Scavina (fungal resistant cultivar) (Figure 8 a, b). According to expression heatmaps, most *tcMGT* genes were less induced by a fungal infection, *P. megakarya*. A *NIPA* gene, *tcMGT02*, showed an upregulation after 48h of fungal infection in both cultivars (Figure 8). Besides, two *MRS2* genes, *tcMGT18* and *tcMGT12*, were more expressed after 72h in fungal resistant cultivar (Figure 8b).

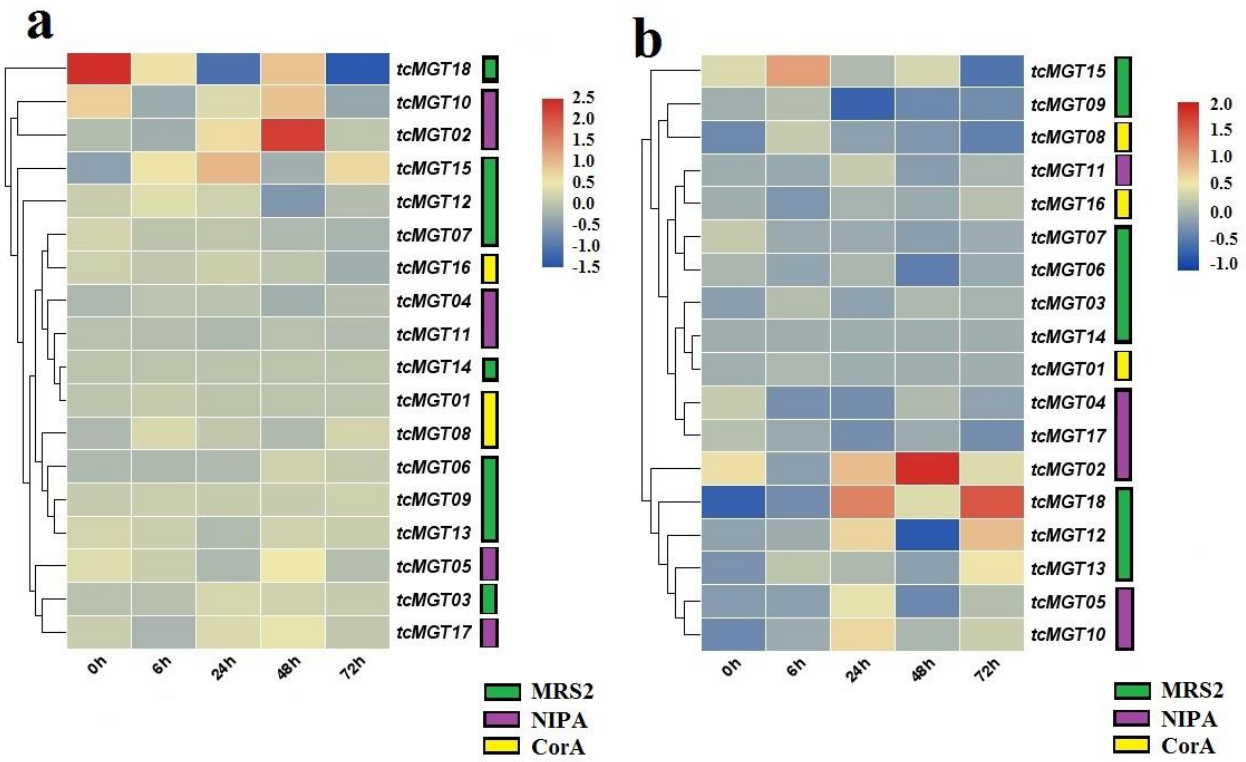


**Figure 8.** Expression profile of *tcMGT* genes in response *to P. megakarya* inoculation after 0h, 6h, 24h, and 72h. In this figure, Nanay (NA-32) is susceptible cultivar (a), and Scavina (SCA6) is tolerant cultivar (b). The type of tcMGTs proteins sub-groups is highlighted using different color.

*3.7. Expression profile of ghMGT genes*

The expression profile of *ghMGT* genes was investigated in different tissues of *G. hirsutum* and under abiotic stresses including cold, heat, drought, and salinity (Figure 9). Results illustrated that *ghMGTs* are involved in early and late response to abiotic stresses (Figure 9a). For instance, a *MRS2* gene, *ghMGT*32, showed an upregulation in response to temperature stresses, cold and heat, after one hour (Figure 9a). Besides, *ghMGT*24, as a *NIPA* gene, was more upregulated in response to all studied abiotic stresses after 6 and 12h. Moreover, expression profile of *ghMGT* genes showed that six *MRS2* genes including *ghMGT12*, *ghMGT33*, *ghMGT06*, *ghMGT41*, *ghMGT13*, and *ghMGT20* are more expressed after 24h of drought and salt stress. In the first hours of heat stress, *NIPA* genes are more expressed than *MRS2* genes in cotton. Besides, *CorA* like genes, *ghMGT*35 and *ghMGT14*, showed more upregulation in response to drought and salt stress. In the present study, expression levels of *ghMGTs* were also evaluated in different tissues and organs of *G. hirsutum* (Figure 9b). Results show that *ghMGTs* were expressed in different organs for the proper distribution of magnesium throughout the cotton plant. In root tissues, *NIPA* genes are more expressed, while in leaf and torus tissues, *MRS2* genes are more expressed (Figure 9b). Besides, two *CorA* like genes showed high expression in filament tissues. In addition, most *ghMGTs* are expressed in ovule tissues in 10 days post anthesis (DPA) (Figure 9b).

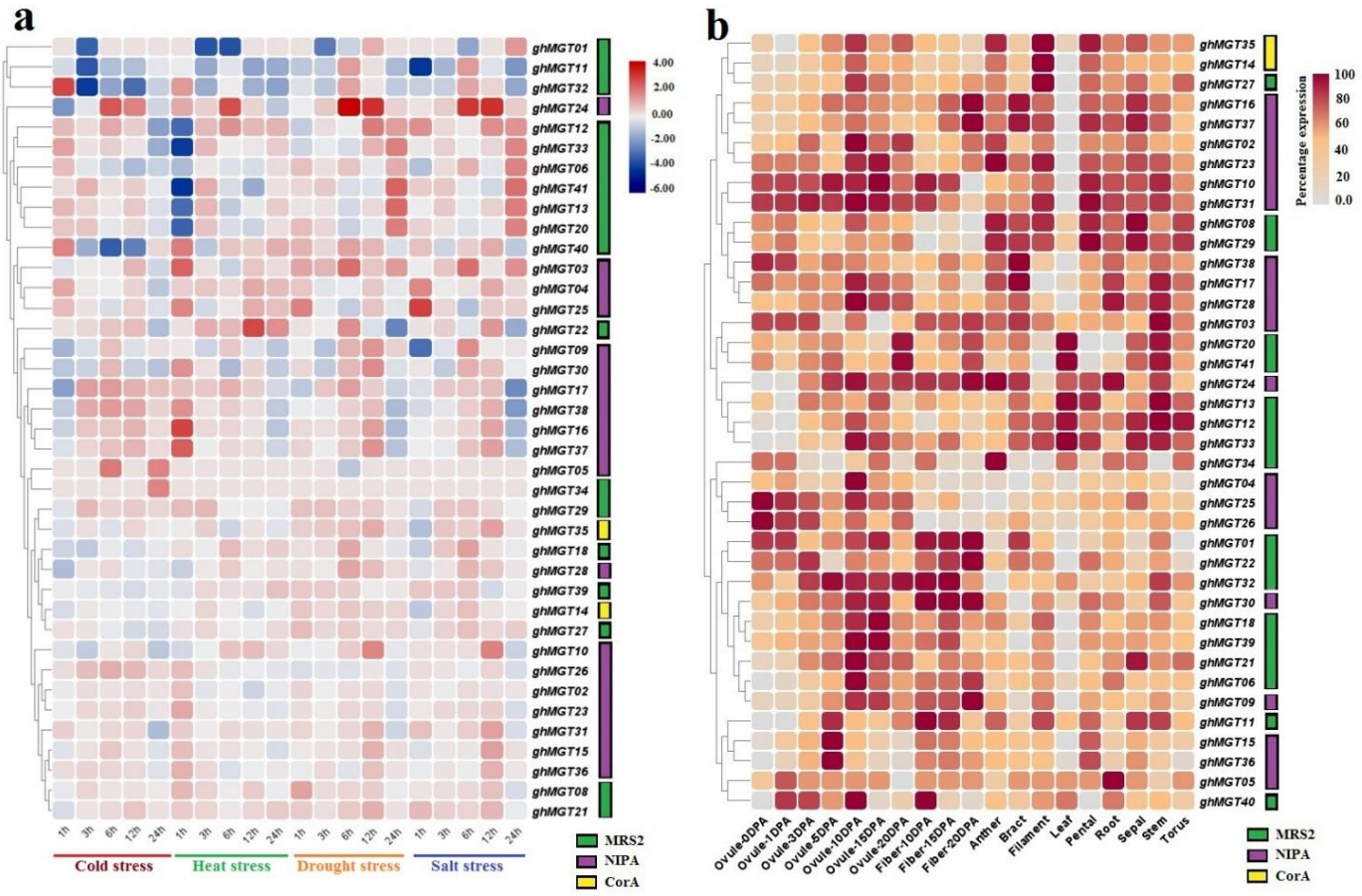


**Figure 9.** Expression profile of *ghMGT* genes in response to abiotic stresses, including cold, heat, drought, and salt stress (a), and different tissues of cotton. The type of ghMGT proteins sub-groups is highlighted using different color.

*3.8. Co-expression network of MGT genes*

In the current study, a co-expression network of *tcMGTs* was formed using their orthologues in the diploid model plant, Arabidopsis (Figure 10). All nodes in the co-expression network of *MGTs* were classified into four groups (I to IV) (Figure 10). In group I, three auxin-responsive genes, *SAUR4*, *7*, and *12*, two genes affecting lipid metabolism, *PLC* genes, two genes involving in disease resistance-responsive, *AT4G38700* and *AT2G21110*, and *CYP21-1* affecting in protein folding showed high co-expression score with *MGT* genes. In group II, a protein phosphatase 2C, *AT2G20050*, was found with high co-expression with three *MGT* genes. In group III, *MGT* genes showed high co-expression with *AT5G42070*, a hypothetical protein, *RPL15* (ribosomal protein), *AT2G23390* (acyl-CoA protein) and *CH1*. In group IV, high co-expression connections were observed between *MGT* genes and genes involving in cell wall organization/modification processes such as *S2LB*, *AGM1*, *AGM2*, *EXPA13*, *GXM2*, *GXM1*, and *QRT1*. In addition, *LrgB* involving in response to water deprivation, and *MHX* encoding a mg/proton exchanger showed string co-expression with *MGTs*. Gene ontology (GO) analysis illustrated that the biological processes including mg ion transporter, photoprotection, cell wall pectin biosynthetic, xylan biosynthetic, and chloroplast organization were significantly enriched based on all nodes of co-expression networks of *MGTs*.

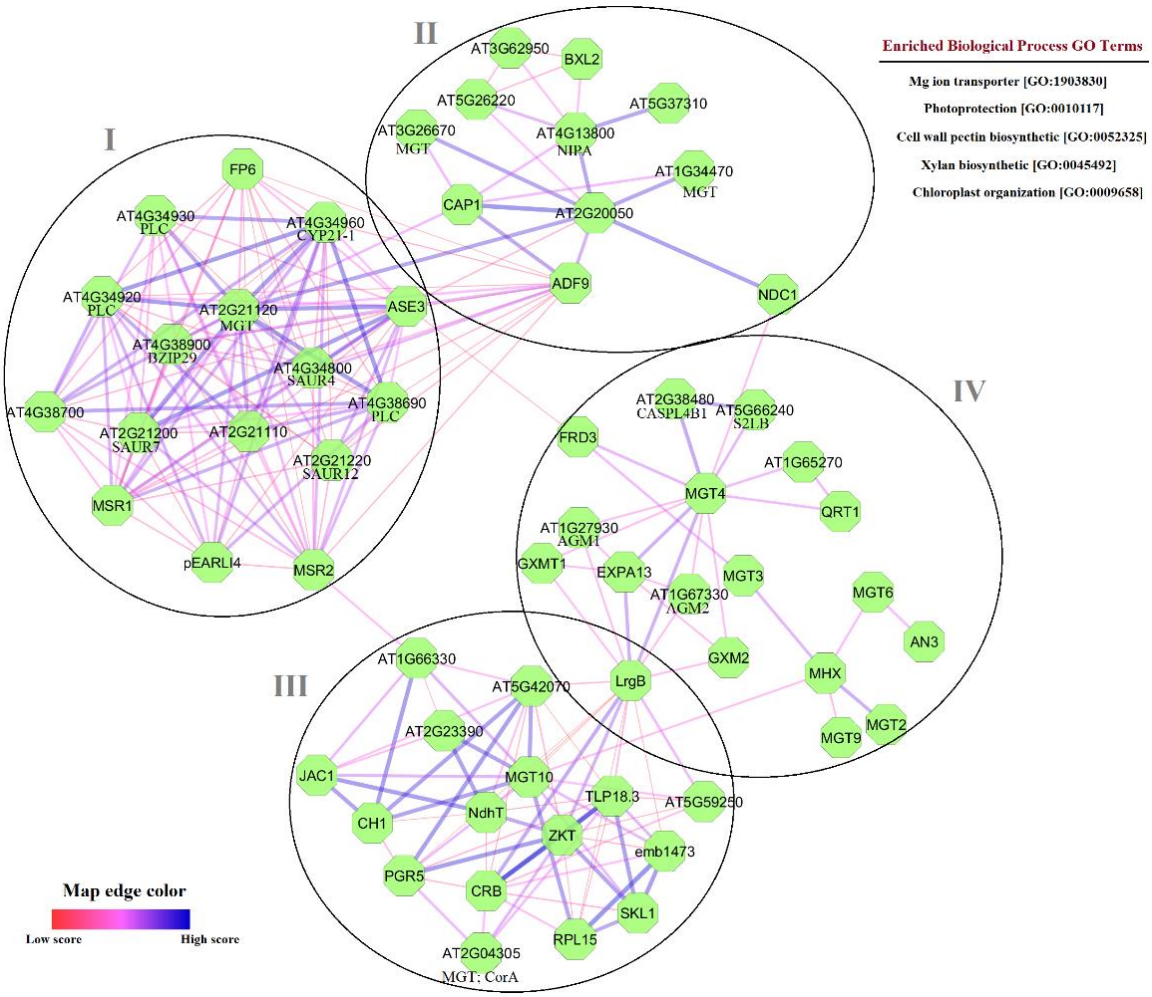


**Figure 10.** The co-expression networks of *MGT* genes based on model plant, *Arabidopsis thaliana*. Thickness and color intensity of each edge indicate the value of co-expression between two genes.

## 4. Discussion

Cocoa (*Theobroma cacao* L.) is an economical plant due to its wide used to produce chocolate, the most popular commodity worldwide. Scientific research is underway to better understand the genome and metabolomes of this plant. Besides, several gene families including *GASA* [35], *sucrose transporter* [57], *WRKY* [58], *NAC* [59], *desaturase* [60], and *sucrose synthase* [61] are investigated and characterized in cacao. Due to the important role of magnesium ions in regulating plant growth and development, magnesium transporters (MGTs) have been studied in various plant species. In the present study, *MGT* family members in cacao were widely characterized and compared with their orthologous in upland cotton, *Gossypium hirsutum*, and white jute (*Corchorus capsularis*). As the first report, we characterized 18 putative non-redundant *MGT* genes in cacao (*tcMGTs*) along with 41*MGTs* in *G. hirsutum* (*ghMGTs*), and 16 *MGTs* from the genome of *C. capsularis* (*ccMGTs*). According the previous studies, the 62 *MGTs* in *Camelina sativa* [12], the 41 *MGTs* in *Triticum turgidum* [12], the 12 *MGTs* in *Zea mays* [62], the 16 *MGTs* in *Pyrus bretschneideri* [23], the 36 *MGTs* in *Brassica napus* [25], the 12 *MGTs* in *Fagaria vesca* [27], and the 8 *MGTs* s in *Poncirus trifoliata* [24] were characterized. It seems that the number of *MGTs* is probably correlated with polyploidy events and genome size [12,63]. The prediction of the pI value of MGT proteins illustrated that ccMGTs are more acidophilic proteins than tcMGTs and ghMGTs, indicating that ccMGTs are mostly active under acidic conditions, pI < 6.50. It can be related to the optimal growing environment of white jute.

According to the phylogenetic analysis, MGT family members from *T. cacao*, *G. hirsutum*, and *C. capsularis* can be classified into five groups and ccMGTs showed close relationships to tcMGTs. In addition, the MGT family proteins of Arabidopsis and rice were divided in five clusters based on phylogenetic analysis [11,64]. In the current study, MRS2 sub-group proteins showed more diversity than the other two sub-groups, indicating that *NIPA* and *CorA* genes may be derived from *MRS2* genes during the process of evolution. In addition, more *MGT* genes were identified in cotton, which is due to polyploidy in *G. hirsutum*. The analysis of gene structure showed that *tcMGTs* and *ghMGTs* contained 4 to 15 exons, while *ccMGTs* contained 4 to 21 exons. It suggests that under evolution events, more insertion and deletion of introns were occurred in *MGT* genes, especially in *ccMGTs*. In addition, the first duplication event estimated approximately 103 MYA between *ccMGT05*, as a *MRS2* gene, and *ccMGT04*, as a *NIPA* gene.

Several segmental/tandem duplication events were estimated between *tcMGTs*. Notably, the gene cluster of three MRS2/*tcMGT* genes including *tcMGT013*, *tcMGT14*, and *tcMGT15* was observed in chromosome 6, which are great potential to further molecular investigation in cacao. Furthermore, the $K_a/K_s$ ratios in most duplicated genes were less than one, suggesting that *MGTs* have been evolved under a purifying selection and they were slowly evolving [63,65]. The comparisons between structures of MGT proteins revealed that NIPA proteins include the conserved structures with more transmembrane regions. More transmembrane regions could indicate a more important role for NIPA proteins group in the transport of magnesium within the plant cell [12]. However, transmembrane structure of MSR2 and CorA proteins in *T.* cacao and *C. capsularis* is highly conserved, while these proteins possess different transmembrane structure in *G. hirsutum*, which may affect the obligations in the magnesium transmembrane transport process [66,67]. Besides, sub-groups of MGTs in three studied plants showed diversity in the predicted pocket sites into 3D structure. Overall, our findings suggest that PHE, GLU, LEU, GLY, and ILE as the key binding sites are associated with the function and interaction of MGTs in response to environmental stimuli as well as changes in ion/mg concentration [65,68,69].

Magnesium transporters as well as other ion transporters, are not only involved in responding to Mg concentrations, but also their activity can be affected by changes in environmental conditions [10,12]. Analysis of gene promoter regions is one of the strategies that can predict the response of target genes to various environmental factors [70,71]. *MGT* genes have a high potential to respond to stresses, biotic and abiotic, as well as ABA based on the distribution of corresponding cis-regulatory elements in the promoter region. However, investigations that are more molecular need to confirm their functions. The black rod disease caused by the genus *Phytophthora* is one of production limiting factors, reducing cacao production around 20-25% [72]. On the other hand, expression studies of genes in beans of bulk cultivars, like disease resistance and fine flavor cocoa (disease susceptibility) may be of interest [73,74]. In the current study, the expression level of *tcMGT* genes using RNA sequencing data responses to *P. megakarya* in two contrasting cultivars of *T. cacao*, susceptible and fungal resistant cultivar. Results reveal that two tcMGTs including *tcMGT12*, and *tcMGT18* were identified as *P. megakarya* responsiveness genes by their upregulation specifically in cacao tolerant cultivar. These may be a good target for further molecular studies related to introducing the new cocao-resistant cultivars with high quality, delicious chocolate. Furthermore, the expression profile of *ghMGTs* suggested that *MGT* genes are involved in response to abiotic stresses such as temperature stresses (cold and heat stress), drought, and salinity stress. Nevertheless, *ghMGTs* can express in different plant tissues to regulate Mg homeostasis. Besides, previous studies indicated that *MGT* genes are associated with controlling ions homeostasis in plant tissues against adverse conditions [20,21].

*Mg transporter* genes by affecting Na+ transporters and $K^+$ transporters (HKTs) can improve salinity tolerance in plant species [75]. Besides, it was stated that MGTs could regulate the downstream pathways related to response to abiotic stresses by interacting with $Ca^{2+}$ sensors [76]. Our findings revealed that the *MGT* duplicated gene pair could have diverse expression patterns, suggesting that these genes probably under some modifications or insertion/deletion in their sequence, CDS or promoter regions, received novel functions [63,77]. The modifications such as gain and loses of cis elements in promoters between duplicated gene pair, parent and daughter genes, could occur after duplication events, affecting the expression levels [78,79]. By constructing a network of co-expression genes, it is possible to identify other molecular pathways in which target genes are involved to gain a better understanding of the function of genes [65,80]. In the present study, *MGT* genes showed the diverse co-expressions with genes involved in auxin responsive, lipid metabolism, cell wall organization, photoprotection, and chloroplast organization. Magnesium is a critical element of chlorophyll affecting photosynthesis rate and biomass production [81,82]. Overall, it seems that *MGTs* involve in various pathways to control the plant growth and development as well as response to adverse conditions.

## 5. Conclusions

In the current study, a genome-wide analysis of the *MGT* family genes was performed in genome of *Theobroma cacao, Corchorus capsularis, and Gossypium hirsutum*. Our findings provide insight into certain aspects of the sequence structure, evolutionary events, regulatory systems and function of *MGT* genes in three species of Malvaceae. Furthermore, our results show that MGTs are involved in diverse cellular pathways, and they can interact with proteins associated with growth and development as well as response to environmental stimuli. These findings can be improved by further functional-molecular analyses to improve our understanding of MGTs role in cacao resistance to stress and quality delicious chocolate.

**Supplementary Materials:**

**Table S1.** The complete detail of each sequence of MGTs in *Theobroma cacao, Corchorus capsularis*, and *Gossypium hirsutum* analyzed in current study including protein sequences, coding sequences, genomic sequences, and promoter regions.

**Table S2.** List and properties of studied *MGT* genes in *Theobroma cacao, Corchorus capsularis*, and *Gossypium hirsutum*

**Table S3.** The predicted duplicated gene pairs in the MGT protein family in *Theobroma cacao, Corchorus capsularis*, and *Gossypium hirsutum*

**Table S4.** Cis-regulatory elements in promoter regions of *MGT* gene family in *Theobroma cacao, Corchorus capsularis*, and *Gossypium hirsutum*

**Figure S1.** The 3D structures of MGT proteins from various clades, NIPA, MRS2, and CorA, in *Theobroma cacao, Corchorus capsularis*, and *Gossypium hirsutum*

**Author Contributions:** Conceptualization, P.H. and A.; methodology, P.H., A. and S.F.; formal analysis, P.H., A. and S.F.; investigation, P.H., A. and P.P.; writing—original draft preparation, P.H. and A.; writing—review and editing, P.H., A. and P.P.; funding acquisition, P.P. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The public data set is analyzed in the manuscript. All the analyzed data are available in the main manuscript or as Supplementary Materials.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


1. Ceylan, Y.; Kutman, U.B.; Mengutay, M.; Cakmak, I. Magnesium applications to growth medium and foliage affect the starch distribution, increase the grain size and improve the seed germination in wheat. *Plant Soil* **2016**, *406*, 145–156.

2. Chen, Z.C.; Peng, W.T.; Li, J.; Liao, H. Functional dissection and transport mechanism of magnesium in plants. In Proceedings of the Seminars in cell & developmental biology; Elsevier, 2018; Vol. 74, pp. 142–152.

3. Hermans, C.; Conn, S.J.; Chen, J.; Xiao, Q.; Verbruggen, N. An update on magnesium homeostasis mechanisms in plants. *Metallomics* **2013**, *5*, 1170–1183.

4. Guo, W.; Nazim, H.; Liang, Z.; Yang, D. Magnesium deficiency in plants: An urgent problem. *Crop J.* **2016**, *4*, 83–91.

5. Marschner, H. *Marschner's mineral nutrition of higher plants*; Academic press, 2011; ISBN 0123849063.

6. Hermans, C.; Bourgis, F.; Faucher, M.; Strasser, R.J.; Delrot, S.; Verbruggen, N. Magnesium deficiency in sugar beets alters sugar partitioning and phloem loading in young mature leaves. *Planta* **2005**, *220*, 541–549.

7. Tang, N.; Li, Y.; Chen, L. Magnesium deficiency–induced impairment of photosynthesis in leaves of fruiting Citrus reticulata trees accompanied by up-regulation of antioxidant metabolism to avoid photo-oxidative damage. *J. plant Nutr. soil Sci.* **2012**, *175*, 784–793.

8. Peng, H.-Y.; Qi, Y.-P.; Lee, J.; Yang, L.-T.; Guo, P.; Jiang, H.-X.; Chen, L.-S. Proteomic analysis of Citrus sinensis roots and leaves in response to long-term magnesium-deficiency. *BMC Genomics* **2015**, *16*, 253.

9. Li, H.; Du, H.; Huang, K.; Chen, X.; Liu, T.; Gao, S.; Liu, H.; Tang, Q.; Rong, T.; Zhang, S. Identification, and functional and expression analyses of the CorA/MRS2/MGT-type magnesium transporter family in maize. *Plant Cell Physiol.* **2016**, *57*, 1153–1168.

10. Li, H.; Wang, N.; Ding, J.; Liu, C.; Du, H.; Huang, K.; Cao, M.; Lu, Y.; Gao, S.; Zhang, S. The maize CorA/MRS2/MGT-type mg transporter, ZmMGT10, responses to magnesium deficiency and confers low magnesium tolerance in transgenic Arabidopsis. *Plant Mol. Biol.* **2017**, *95*, 269–278.

11. Gebert, M.; Meschenmoser, K.; Svidová, S.; Weghuber, J.; Schweyen, R.; Eifler, K.; Lenz, H.; Weyand, K.; Knoop, V. A root-expressed magnesium transporter of the MRS2/MGT gene family in Arabidopsis thaliana allows for growth in low-Mg2+ environments. *Plant Cell* **2009**, *21*, 4018–4030.

12. Faraji, S.; Ahmadizadeh, M.; Heidari, P. Genome-wide comparative analysis of Mg transporter gene family between Triticum turgidum and Camelina sativa. *BioMetals* **2021**, *4*.

13. Mao, D.; Chen, J.; Tian, L.; Liu, Z.; Yang, L.; Tang, R.; Li, J.; Lu, C.; Yang, Y.; Shi, J. Arabidopsis transporter MGT6 mediates magnesium uptake and is required for growth under magnesium limitation. *Plant Cell* **2014**, *26*, 2234–2248.

14. Chen, Z.C.; Yamaji, N.; Motoyama, R.; Nagamura, Y.; Ma, J.F. Up-regulation of a magnesium transporter gene

OsMGT1 is required for conferring aluminum tolerance in rice. *Plant Physiol.* **2012**, *159*, 1624–1633.

15. Drummond, R.S.M.; Tutone, A.; Li, Y.-C.; Gardner, R.C. A putative magnesium transporter AtMRS2-11 is localized to the plant chloroplast envelope membrane system. *Plant Sci.* **2006**, *170*, 78–89.

16. Lenz, H.; Dombinov, V.; Dreistein, J.; Reinhard, M.R.; Gebert, M.; Knoop, V. Magnesium deficiency phenotypes upon multiple knockout of Arabidopsis thaliana MRS2 clade B genes can be ameliorated by concomitantly reduced calcium supply. *Plant cell Physiol.* **2013**, *54*, 1118–1131.

17. Chen, J.; Li, L.; Liu, Z.; Yuan, Y.; Guo, L.; Mao, D.; Tian, L.; Chen, L.; Luan, S.; Li, D. Magnesium transporter AtMGT9 is essential for pollen development in Arabidopsis. *Cell Res.* **2009**, *19*, 887–898.

18. Li, J.; Huang, Y.; Tan, H.; Yang, X.; Tian, L.; Luan, S.; Chen, L.; Li, D. An endoplasmic reticulum magnesium transporter is essential for pollen development in Arabidopsis. *Plant Sci.* **2015**, *231*, 212–220.

19. Bose, J.; Babourina, O.; Shabala, S.; Rengel, Z. Low-pH and aluminum resistance in Arabidopsis correlates with high cytosolic magnesium content and increased magnesium uptake by plant roots. *Plant cell Physiol.* **2013**, *54*, 1093–1104.

20. Deng, W.; Luo, K.; Li, D.; Zheng, X.; Wei, X.; Smith, W.; Thammina, C.; Lu, L.; Li, Y.; Pei, Y. Overexpression of an Arabidopsis magnesium transport gene, AtMGT1, in Nicotiana benthamiana confers Al tolerance. *J. Exp. Bot.* **2006**, *57*, 4235–4243.

21. Chen, Z.C.; Yamaji, N.; Horie, T.; Che, J.; Li, J.; An, G.; Ma, J.F. A magnesium transporter OsMGT1 plays a critical role in salt tolerance in rice. *Plant Physiol.* **2017**, *174*, 1837–1849.

22. Schock, I.; Gregan, J.; Steinhauser, S.; Schweyen, R.; Brennicke, A.; Knoop, V. A member of a novel Arabidopsis thaliana gene family of candidate Mg2+ ion transporters complements a yeast mitochondrial group II intron-splicing mutant. *Plant J.* **2000**, *24*, 489–501.

23. Zhao, Z.; Wang, P.; Jiao, H.; Tang, C.; Liu, X.; Jing, Y.; Zhang, S.; Wu, J. Phylogenetic and expression analysis of the magnesium transporter family in pear, and functional verification of PbrMGT7 in pear pollen. *J. Hortic. Sci. Biotechnol.* **2018**, *93*, 51–63.

24. Liu, X.; Guo, L.-X.; Luo, L.-J.; Liu, Y.-Z.; Peng, S.-A. Identification of the magnesium transport (MGT) family in Poncirus trifoliata and functional characterization of PtrMGT5 in magnesium deficiency stress. *Plant Mol. Biol.* **2019**, *101*, 551–560.

25. Zhang, L.; Wen, A.; Wu, X.; Pan, X.; Wu, N.; Chen, X.; Chen, Y.; Mao, D.; Chen, L.; Luan, S. Molecular identification of the magnesium transport gene family in Brassica napus. *Plant Physiol. Biochem.* **2019**, *136*, 204–214.

26. Wang, Y.; Hua, X.; Xu, J.; Chen, Z.; Fan, T.; Zeng, Z.; Wang, H.; Hour, A.-L.; Yu, Q.; Ming, R. Comparative genomics revealed the gene evolution and functional divergence of magnesium transporter families in Saccharum. *BMC Genomics* **2019**, *20*, 83.

27. Regon, P.; Chowra, U.; Awasthi, J.P.; Borgohain, P.; Panda, S.K. Genome-wide analysis of magnesium transporter genes in Solanum lycopersicum. *Comput. Biol. Chem.* **2019**, *80*, 498–511.

28. Christenhusz, M.J.M.; Byng, J.W. The number of known plants species in the world and its annual increase. *Phytotaxa* **2016**, *261*, 201–217.

29. Xu, Z.; Deng, M. Malvaceae. In *Identification and control of common weeds*; Springer Netherlands, 2017; pp. 717–735 ISBN 9789402411577.

30. Bayer, C.; Kubitzki, K. Malvaceae. In *Flowering Plants · Dicotyledons*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2003; pp. 225–311.

31. Argout, X.; Martin, G.; Droc, G.; Fouet, O.; Labadie, K.; Rivals, E.; Aury, J.M.; Lanaud, C. The cacao Criollo genome v2.0: an improved version of the genome for genetic and functional genomic studies. *BMC Genomics 2017 181* **2017**, *18*, 1–9.

32. Islam, M.S.; Saito, J.A.; Emdad, E.M.; Ahmed, B.; Islam, M.M.; Halim, A.; Hossen, Q.M.M.; Hossain, M.Z.; Ahmed, R.; Hossain, M.S.; et al. Comparative genomics of two jute species and insight into fibre biogenesis. *Nat. Plants* **2017**, *3*, 16223.

33. Bolser, D.M.; Staines, D.M.; Perry, E.; Kersey, P.J. Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data. In *Plant Genomics Databases*; Springer, 2017; pp. 1–31.

34. Yu, J.; Jung, S.; Cheng, C.-H.; Ficklin, S.P.; Lee, T.; Zheng, P.; Jones, D.; Percy, R.G.; Main, D. CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Res.* **2014**, *42*, D1229–D1236.

35. Abdullah; Faraji, S.; Mehmood, F.; Malik, H.M.T.; Ahmed, I.; Heidari, P.; Poczai, P. The GASA Gene Family in Cacao (Theobroma cacao, Malvaceae): Genome Wide Identification and Expression Analysis. *Agronomy* **2021**, *11*, 1425.

36. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, 2005; pp. 571–607.

37. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580.

38. Yu, C.; Chen, Y.; Lu, C.; Hwang, J. Prediction of protein subcellular localization. *Proteins Struct. Funct. Bioinforma.* **2006**, *64*, 643–651.

39. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327.

40. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization

and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649.

41. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R. Clustal W and Clustal X version 2.0. *bioinformatics* **2007**, *23*, 2947–2948.

42. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274.

43. Hoang, D.T.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522.

44. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589.

45. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259.

46. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208.

47. Hu, B.; Jin, J.; Guo, A.-Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296–1297.

48. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202.

49. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302.

50. Yang, S.; Zhang, X.; Yue, J.-X.; Tian, D.; Chen, J.-Q. Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genomics* **2008**, *280*, 187–198.

51. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: an information aesthetic for comparative genomics. *Genome Res.* **2009**, *19*, 1639–1645.

52. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858.

53. Jiménez, J.; Doerr, S.; Martínez-Rosell, G.; Rose, A.S.; De Fabritiis, G. DeepSite: protein-binding site predictor using 3D-convolutional neural networks. *Bioinformatics* **2017**, *33*, 3036–3042.

54. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363–W367.

55. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. protein Crystallogr.* **2002**, *40*, 82–92.

56. Zhu, T.; Liang, C.; Meng, Z.; Sun, G.; Meng, Z.; Guo, S.; Zhang, R. CottonFGD: an integrated functional genomics

database for cotton. *BMC Plant Biol. 2017 171* **2017**, *17*, 1–9.

57. Li, F.; Wu, B.; Qin, X.; Yan, L.; Hao, C.; Tan, L.; Lai, J. Molecular cloning and expression analysis of the sucrose transporter gene family from Theobroma cacao L. *Gene* **2014**, *546*, 336–341.

58. Dayanne, S.M.D.A.; Oliveira Jordão Do Amaral, D.; Del-Bem, L.E.; Bronze Dos Santos, E.; Santana Silva Raner, J.; Peres Gramacho, K.; Vincentz, M.; Micheli, F. Genome-wide identification and characterization of cacao WRKY transcription factors and analysis of their expression in response to witches' broom disease. *PLoS One* **2017**, *12*.

59. Shen, S.; Zhang, Q.; Shi, Y.; Sun, Z.; Zhang, Q.; Hou, S.; Wu, R.; Jiang, L.; Zhao, X.; Guo, Y. Genome-wide analysis of the NAC domain transcription factor gene family in Theobroma cacao. *Genes (Basel).* **2020**, *11*.

60. Zhang, Y.; Maximova, S.N.; Guiltinan, M.J. Characterization of a stearoyl-acyl carrier protein desaturase gene family from chocolate tree, Theobroma cacao L. *Front. Plant Sci.* **2015**, *6*, 239.

61. Li, F.; Hao, C.; Yan, L.; Wu, B.; Qin, X.; Lai, J.; Song, Y. Gene structure, phylogeny and expression profile of the sucrose synthase gene family in cacao (Theobroma cacao L.). *J. Genet.* **2015**, *94*, 461–472.

62. Li, H.; Liu, C.; Zhou, L.; Zhao, Z.; Li, Y.; Qu, M.; Huang, K.; Zhang, L.; Lu, Y.; Cao, M. Molecular and functional characterization of the magnesium transporter gene ZmMGT12 in maize. *Gene* **2018**, *665*, 167–173.

63. Heidari, P.; Faraji, S.; Ahmadizadeh, M.; Ahmar, S.; Mora-Poblete, F. New insights into structure and function of TIFY genes in Zea mays and Solanum lycopersicum: a genome-wide comprehensive analysis. *Front. Genet.* **2021**, *12*, 534.

64. Saito, T.; Kobayashi, N.I.; Tanoi, K.; Iwata, N.; Suzuki, H.; Iwata, R.; Nakanishi, T.M. Expression and functional analysis of the CorA-MRS2-ALR-type magnesium transporter family in rice. *Plant Cell Physiol.* **2013**, *54*, 1673–1683.

65. Musavizadeh, Z.; Najafi-Zarrini, H.; Kazemitabar, S.K.; Hashemi, S.H.; Faraji, S.; Barcaccia, G.; Heidari, P. Genome-Wide Analysis of Potassium Channel Genes in Rice: Expression of the OsAKT and OsKAT Genes under Salt Stress. *Genes (Basel).* **2021**, *12*, 784.

66. Fukao, Y. Protein-protein interactions in plants. *Plant Cell Physiol.* **2012**, *53*, 617–625.

67. Braun, P.; Aubourg, S.; Van Leene, J.; De Jaeger, G.; Lurin, C. Plant protein interactomes. *Annu. Rev. Plant Biol.* **2013**, *64*, 161–187.

68. Galili, G.; Höfgen, R. Metabolic engineering of amino acids and storage proteins in plants. *Metab. Eng.* **2002**, *4*, 3–11.

69. Beauregard, M.; Hefford, M.A. Enhancement of essential amino acid contents in crops by genetic engineering and protein design. *Plant Biotechnol. J.* **2006**, *4*, 561–574.

70. Ahmadizadeh, M.; Chen, J.-T.; Hasanzadeh, S.; Ahmar, S.; Heidari, P. Insights into the genes involved in the ethylene biosynthesis pathway in Arabidopsis thaliana and Oryza sativa. *J. Genet. Eng. Biotechnol.* **2020**, *18*, 1–20.

71. Heidari, P.; Mazloomi, F.; Nussbaumer, T.; Barcaccia, G. Insights into the SAM synthetase gene family and its roles in tomato seedlings under abiotic stresses and hormone treatments. *Plants* **2020**, *9*.

72. Adeniyi, D. Diversity of Cacao Pathogens and Impact on Yield and Global Production. In *Theobroma Cacao - Deploying Science for Sustainability of Global Cocoa Economy*; IntechOpen, 2019.

73. De Wever, J.; Tulkens, D.; Verwaeren, J.; Everaert, H.; Rottiers, H.; Dewettinck, K.; Lefever, S.; Messens, K. A combined RNA preservation and extraction protocol for gene expression studies in cacao beans. *Front. Plant Sci.* **2020**, *11*, 992.

74. De Wever, J.; De Coninck, T.; Everaert, H.; Afoakwa, E.O.; Coppieters, F.; Rottiers, H.; Opoku, S.Y.; Lowor, S.; Dewettinck, K.; Vandesompele, J. Selection and validation of reference genes for accurate RT-qPCR gene expression normalization in cacao beans during fermentation. *Tree Genet. Genomes* **2021**, *17*, 1–14.

75. Martínez-Atienza, J.; Jiang, X.; Garciadeblas, B.; Mendoza, I.; Zhu, J.-K.; Pardo, J.M.; Quintero, F.J. Conservation of the salt overly sensitive pathway in rice. *Plant Physiol.* **2007**, *143*, 1001–1012.

76. Manishankar, P.; Wang, N.; Köster, P.; Alatar, A.A.; Kudla, J. Calcium signaling during salt stress and in the regulation of ion homeostasis. *J. Exp. Bot.* **2018**, *69*, 4215–4226.

77. Faraji, S.; Filiz, E.; Kazemitabar, S.K.; Vannozzi, A.; Palumbo, F.; Barcaccia, G.; Heidari, P. The AP2/ERF Gene Family in Triticum durum: Genome-Wide Identification and Expression Analysis under Drought and Salinity Stresses. *Genes (Basel).* **2020**, *11*, 1464.

78. Zhang, J. Evolution by gene duplication: an update. *Trends Ecol. Evol.* **2003**, *18*, 292–298.

79. Hahn, M.W. Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* **2009**, *100*, 605–617.

80. Ahmadizadeh, M.; Heidari, P. Bioinformatics study of transcription factors involved in cold stress. *Biharean Biol.* **2014**, *8*.

81. Cakmak, I.; Kirkby, E.A. Role of magnesium in carbon partitioning and alleviating photooxidative damage. *Physiol. Plant.* **2008**, *133*, 692–704.

82. Tränkner, M.; Jákli, B.; Tavakol, E.; Geilfus, C.-M.; Cakmak, I.; Dittert, K.; Senbayram, M. Magnesium deficiency decreases biomass water-use efficiency and increases leaf water-use efficiency and oxidative stress in barley plants. *Plant Soil* **2016**, *406*, 409–423.