 doi:10.20944/preprints202206.0278.v1

Article

Genome-Wide Identification, and Molecular Evolution of The Magnesium Transporter (MGT) Gene Family in *Citrullus Lanatus* and *Cucumis Sativus*

Parviz Heidari [1,*], Fatemeh Puresmaeli [1], Freddy Mora-Poblete [2,*]

[1] Faculty of Agriculture, Shahrood University of Technology, Shahrood 3619995161, Iran; heidarip@shahroodut.ac.ir (P.H.); fesmaeilabadi@gmail.com (F.P.)
[2] Institute of Biological Sciences, University of Talca, Talca 3460000, Chile; fmora@utalca.cl (F.M-P.)
* Correspondence: heidarip@shahroodut.ac.ir; fmora@utalca.cl

Abstract: Magnesium transporters (MGTs) play a prominent role in the absorption, transport and storage of magnesium in plant cells. In the present study, MGT gene family members were identified and characterized in two species of cucurbitaceae, including *C. sativus* and *C. lanatus*. 20, 19, and 20 *MGT* genes were recognized in *C. lanatus*, *C. sativus*, and *C. melo*, respectively. According to physicochemical properties, the members of each sub-class of MGTs in the species of cucurbitaceae showed the close relationship. Proteins from NIPA were identified as hydrophilic proteins with high stability. Based on phylogenetic analysis, MGT family members were classified into three groups, and NIPAs showed more diversity. Besides, duplication events were not identified between the *MGT* members in *C. lanatus*, and *C. sativus*. According to pocket analysis, residues such as L, V, S, I, and A were frequently observed in the binding sites of MGT proteins in both species. The prediction of post-translation modifications revealed that MSR2 proteins have high phosphorylation potential than other sub-classes in both studied plants. The expression profile of *MGTs* showed that *MGTs* are more expressed in root tissues. In addition, *MGTs* showed differential expression in response to abiotic/biotic stresses as well as hormone application and *NIPAs* were more induced in response to stimuli in watermelon. The results of this study, as the primary work of MGT gene family, can be used in programs related to cucurbitaceae breeding.

Keywords: Magnesium; Evolution analysis; Plant gene families; Gene sequence analysis; Stresses

1. Introduction

Magnesium (Mg) is a vital element for living cells involved in many critical cellular activities [1–3]. For instance, Mg as a cofactor is essential for the activity of many enzymes (> 300 enzymes) such as kinase, polymerase, and H + -ATPase [4–6]. In addition, Mg, as the key atom of chlorophyll, affects the photosynthesis rate and plant growth [7]. Magnesium transporters (MGTs) are present in plants for Mg uptake, translocation, and cell storage. CorA protein as a member of MGTs was firstly identified in bacteria, *Salmonella typhimurium* [8], and in plants for the first time, MGTs were studied in the model plant, *Arabidopsis* [9]. Based on sequence structure, MGTs have been classified into three groups, including MRS2, CorA, and NIPA [10]. MRS2 and CorA proteins are recognized by a tripeptide conserved region, GMN (Glycine-Methionine-Asparagine), and two or three transmembrane (TM) domains in their C-terminal ends [11], while NIPAs contain several TMs in its structure [3,10]. However, our knowledge of NIPA class is limited. Due to the important role of Mg in plants, members of the MGT gene family have been identified and studied in different plants such as *Arabidopsis* [12], *Triticum turgidum* and *Camelina sativa* [10], pear [13], cacao [3], rapeseed [14], maize [15], citrus [16], tomato [17], rice [11], and sugarcane [18]. Besides, the function of recognized *MGT* genes was experimentally characterized in plant species.

 doi:10.20944/preprints202206.0278.v1

MGTs are distributed in various plant organs, including root, flower, leaf, and stem, to balance magnesium concentrations [10]. Previous studies revealed that some *MGTs* in root tissues are involved in uptaking Mg from the soil, including *OsMGT1* in *Oryza sativa*, and *AtMGT6* in *Arabidopsis thaliana* [19,20]. In addition, *AtMGT9* was identified as an Mg transporter translocating Mg from root tissues to shoot tissues in *Arabidopsis* [11]. Furthermore, *MGTs*, such as *AtMGT5* and *AtMGT9*, have been recognized to be involved in pollen development in *Arabidopsis thaliana* [21–23]. Moreover, some MGTs are located in the membranes of cellular organelles and are involved in the distribution and accumulation of Mg within the cell. For instance in *Arabidopsis*, AtMGT2 and AtMGT3 accumulate Mg in the vacuole [24], while AtMGT10 maintains Mg homeostasis in chloroplasts [21]. It has also been reported that *MGTs* increase plant tolerance to environmental stresses. For example, *OsMGT1* was identified as a gene related to response to salt stress in rice [25]. In addition, a positive correlation has also been reported between aluminum (Al) stress tolerance in plants and the expression of *MGT* genes. Furthermore, *OsMGT1* in rice was recognized to be involved in the tolerance to Al stress [25]. Besides, the transgenic lines for *AtMGT1* in *Nicotiana benthamiana* showed a reduction in Al toxicity [26]. It seems, that increasing MGT activity and more Mg uptake play an important role in reducing the negative effects of some elements and ions.

Cucurbitaceae are the most diverse plant species, with more than 800 species known worldwide [27,28]. The important vegetable crops, including cucumber (*Cucumis sativus* L.), melon (*C. melo* L.), watermelon (*Citrullus lanatus*), and squash (*Cucurbita L.*) belong to cucurbitaceae. Due to the important role of the MGT gene family, none of the studies have been conducted to identify and structurally evaluate members of this gene family in species of the cucurbitaceae. In the current study, we focus on identifying and characterizing the members of MGT family in two species of cucurbitaceae, *C. sativus* and *C. lanatus*. Besides, phylogenetic relationships, protein structure, phosphorylation sites, expression patterns, and promoter region analysis of members of the MGT family were conducted in both species. The present results improve our understanding of the structure, regulatory systems and function of the *MGTs* in two candidate species of cucurbitaceaecan. As a primary study can be used in future studies related to functional analysis of *MGTs* in *Citrullus lanatus,* and *Cucumis sativus*.

2. Results

2.1. Identification of MGT gene family in watermelon, cucumber, and melon

By searching in three species of cucurbitaceae, 20, 19, and 20 *MGT* genes were identified in *Citrullus lanatus*, *Cucumis sativus*, and *Cucumis melo*, respectively (Table 1, Table S1). In addition, subclasses of MGT, including MRS2, NIPA, and CorA, were identified based on their specific domain distribution (Table S1). Based on physicochemical properties, the MGTs in all three studied species of cucurbitaceae were close to each other and almost similar. Protein length of MGTs in *C. lanatus* ranged from 323 amino acids (aa) to 548 aa, *C. sativus* from 305 to 567 aa, and *C. melo* from 175 to 546 aa. Moreover, the MW of MGTs in *C. lanatus* varied from 35.44 to 62.78 kDa, from 34.89 to 62.89 kDa in *C. sativus*, and from 20.17 to 63.23 kDa in *C. melo*. In addition, the pI of MGTs ranged from 4.86 to 8.32 in *C. lanatus*, from 4.87 to 9.63 in *C. sativus*, and from 4.60 to 9.47 in *C. melo*. The GRAVY value of MGTs was between -0.35 and 0.77 in *C. lanatus*, between -0.37 and 0.87 in *C. sativus*, and between -0.45 and 1.07 in *C. melo*. Interestingly, proteins from subclasses MSR2 and CorA showed a positive GRAVY index, while most NIPAs showed a negative GRAVY. Moreover, according to the instability index, 45% of MGTs in *C. lanatus*, 53% in *C. sativus*, and 50% in *C. melo* were predicted as stable proteins. Besides, most proteins of subclass NIPA were predicted as stable proteins than subclasses MSR2 and CorA proteins in all studied species. Based on gene structure data, MGTs have between 4 to 13 exons in *C. lanatus*, between 4 to 12 exons in *C. sativus*, and between 3 to 14 exons in *C. melo* (Table S1).

 doi:10.20944/preprints202206.0278.v1

Table 1. Summary of MGTs properties in *Citrullus lanatus*, *Cucumis sativus*, and *Cucumis melo*. Full details are provided in Table S1.

Attributes	***C. lanatus***	***C. sativus***	***C. melo***
Number of gene	20	19	20
Protein length (aa)	323-548	305-567	175-546
Exon number	4-13	4-12	3-14
pI	4.86-8.32	4.87-9.63	4.60-9.47
MW (KDa)	35.44-62.78	34.89-62.89	20.17-63.23
GRAVY	-0.35 - 0.77	-0.37 - 0.87	-0.45 – 1.07
Instability index	0.45% stable	0.53% stable	0.50% stable

2.2. Phylogenetic analysis of the MGT Family

To better understand the evolutionary relationships of the MGT family, a phylogenetic tree was constructed for 85 members of this family in five different species, *Arabidopsis thaliana*, *Oryza sativa*, *Citrullus lanatus*, *Cucumis sativus*, and *Cucumis melo* (Fig. 1). According to phylogenetic tree, MGT family members can be classified into three groups, including group I, II, and III. Subclass NIPA proteins showed more diversity and they were divided into two groups. Besides, subclasses CorA and MSR2 were more closely related and were put together in group III.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

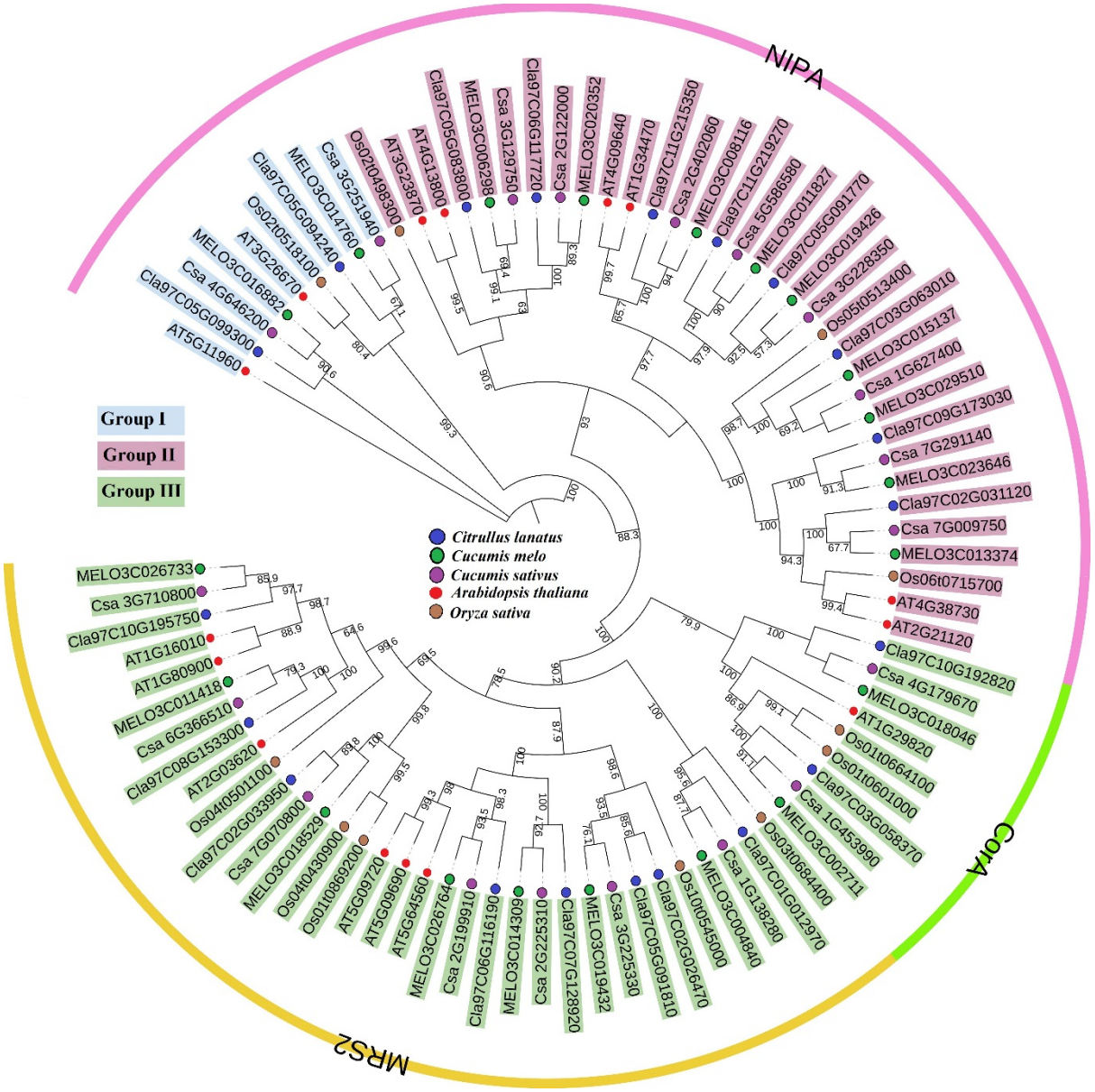

Figure 1. Phylogenetic analysis of MGT families in five different species, *Arabidopsis thaliana*, *Oryza sativa*, *Citrullus lanatus*, *Cucumis sativus*, and *Cucumis melo*. The phylogenetic tree was constructed by the maximum likelihood method. Different colored backgrounds indicate other groups.

2.3. Evolutionary process in MGT genes in Citrullus lanatus, and Cucumis sativus

Duplication events were not identified between the MGT members in *C. lanatus* and *C. sativus*. To investigate the duplication events between *MGTs* in *C. lanatus* and *C. sativus*, the Ks, Ka, and Ka/Ks for all gene pairs were calculated (Fig. 2). The Ks value of *MGTs* in *C. lanatus* was detected between 1.8 and 2.8 (Fig. 2a), and in *C. sativus*, it was observed between 1.8 and 2.6 (Fig. 2b). Besides, the frequency of Ka/Ks of *MGTs* in *C. lanatus* was more observed between 0.5 and 0.6 (Fig. 2c), while in *C. sativus*, Ka/Ks ratio was frequently detected, ranging from 0.4 to 0.5 (Fig. 2d). It seems that a similar evolutionary process has occurred in *C. lanatus* and *C. sativus* for the MGT gene family, and changes (mutations and duplication events) in members of this gene family may have occurred before the derivation of these two species.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

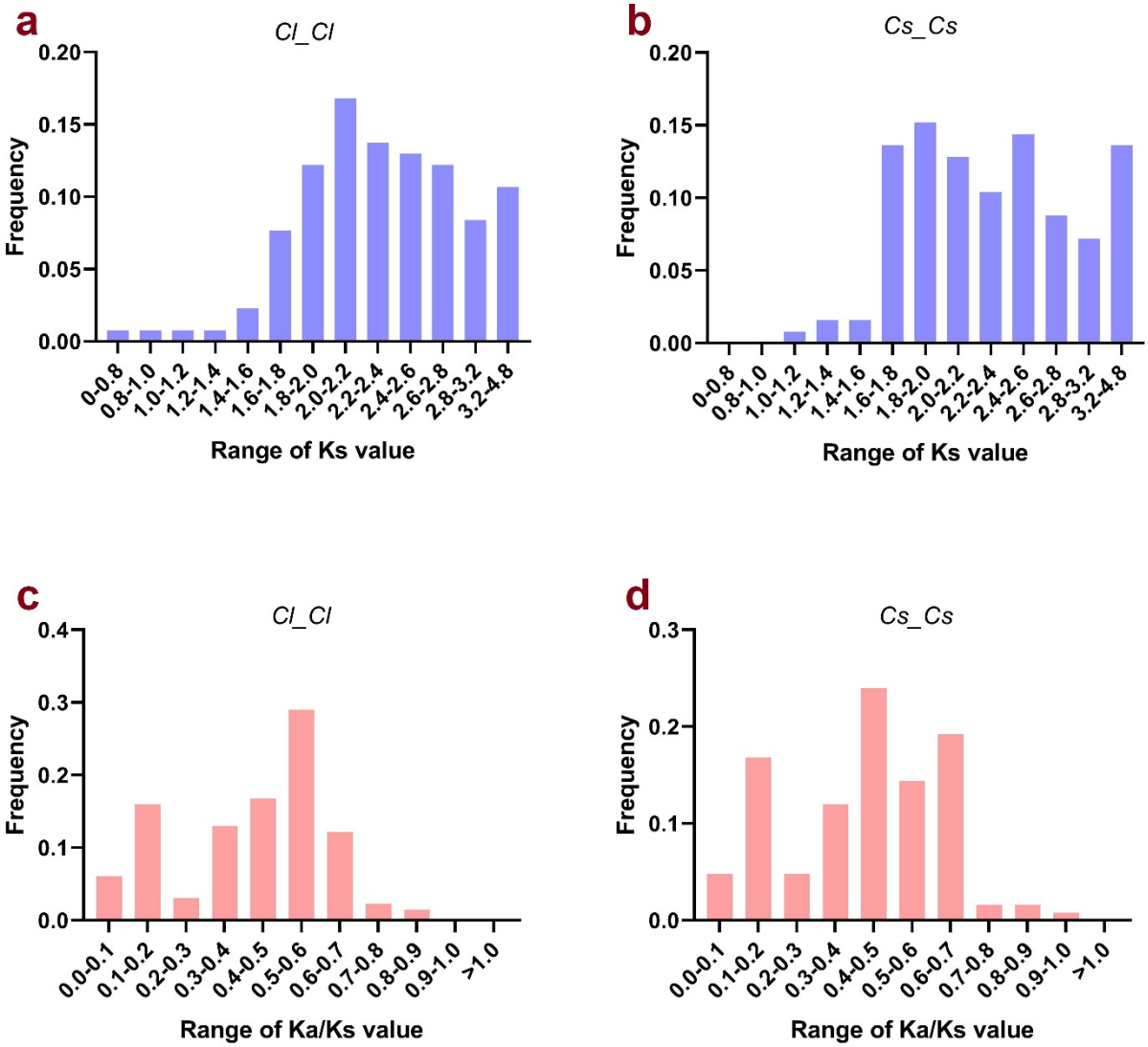

Figure 2. Ks value and Ka/Ks in *MGTs* in *C. lanatus* (Cl), and *C. sativus* (Cs). Frequency of Ks value between *MGTs* in *C. lanatus* (a), and *C. sativus* (b). Frequency of Ka/Ks ratio between *MGTs* in *C. lanatus* (b), and *C. sativus* (c).

2.4. Protein structure of MGTs in C. lanatus, and C. sativus

The three-dimensional structure of the MGT proteins, along with their binding sites in the two plants, *C. lanatus*, and *C. sativus* was predicted (Fig. 3). Results showed that the members of each subclass, NIPA, CorA, and MSR2, have an almost similar structure in both studied plants. Besides, various amino acids were predicted as ligand-binding residues in the MGT structures (Fig. 4; Table S2). L, V, S, I, and A were more observed in the binding sites of MGT proteins in *C. lanatus*, and *C. sativus*, indicating that these residues in molecular function of MGTs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

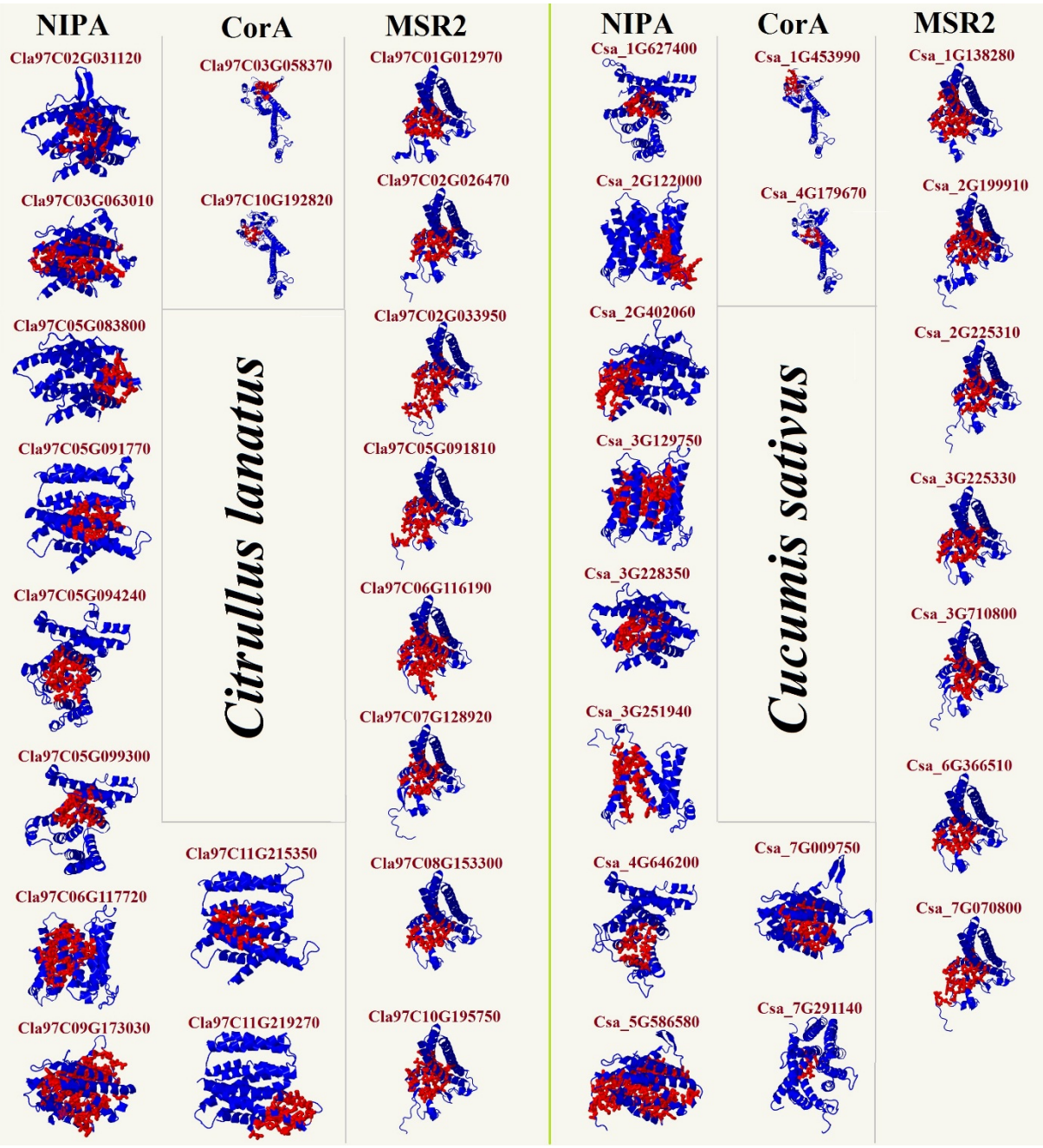

Figure 3. The predicted 3D structure of MGT proteins in *C. lanatus*, and *C. sativus*. The region of ligand-binding sites in the predicted 3D structure of MGTs is highlighted by a red circle. More details are provided in Table S2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

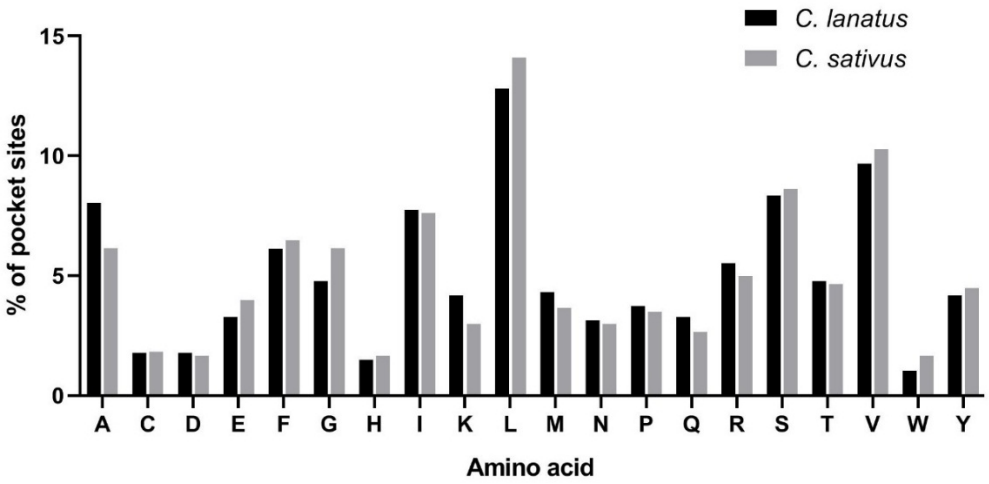

Figure 4. Percentage of amino acids in the predicted pocket sites of MGT proteins in *C. lanatus*, and *C. sativus*.

2.5. Transmembrane structure of MGTs

The transmembrane structure of three subclasses of MGTs was predicted in *C. lanatus*, and *C. sativus* (Fig.5). The most transmembrane helices were observed in NIPA proteins in both species with nine helices, while two helices in CorAs and between two and three helices in MSR2s were predicted. Both C- and N-terminal of the transmembrane structure of CorA proteins were observed in the cytoplasm, while C- and N-terminal of some MSR2 proteins were predicted to locate in extracellular. In addition, a signal peptide was expected in the N-terminal of MSR2 proteins. Overall, the results indicate that MGT proteins have an almost conserved transmembrane structure in both plants.

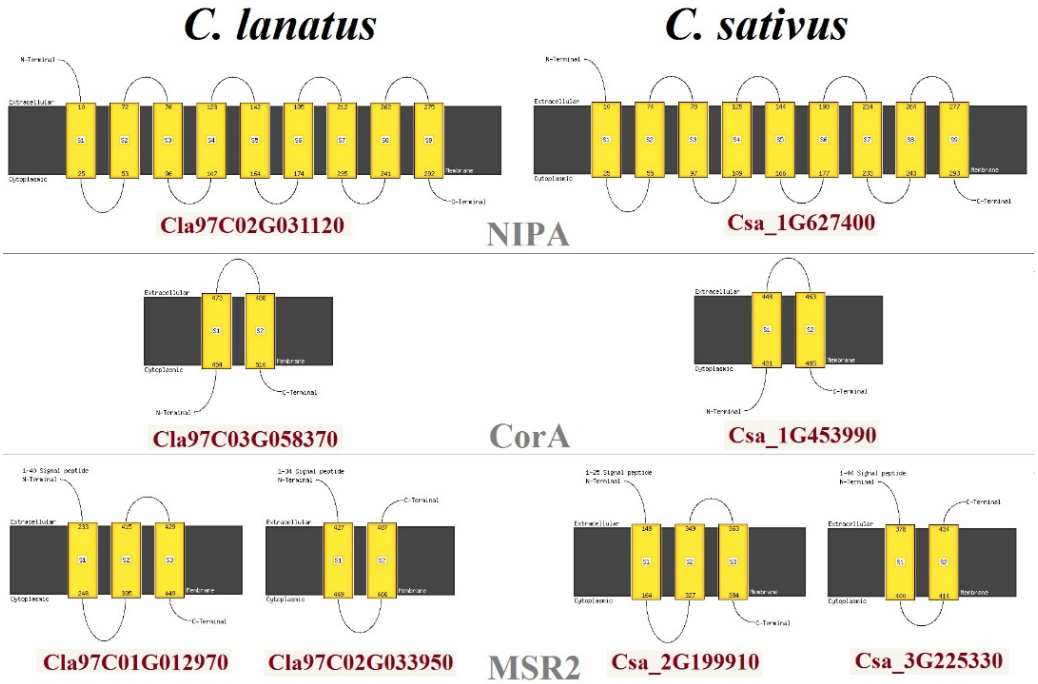

Figure 5. The distribution of transmembrane helices in various subclasses of MGTs in *C. lanatus*, and *C. sativus*.

2.6. Prediction of the phosphorylation site in MGT proteins

The phosphorylation sites into MGTs of *C. lanatus*, and *C. sativus* were predicted based on three amino acids, including serine (S), threonine (T), and tyrosine (Y) (Figure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

6). The predicted phosphorylation sites in MGTs of *C. lanatus* varied from 4 to 18 sites and from 3 to 19 in *C. sativus*. MSR2 proteins showed high phosphorylation potential than NIPAs in both studied plants. Besides, Cla97C07G128920, as a MSR2 protein, in *C. lanatus*, and Csa_1G453990, as a CorA protein, were identified as the MGT proteins with high potential phosphorylation potential.

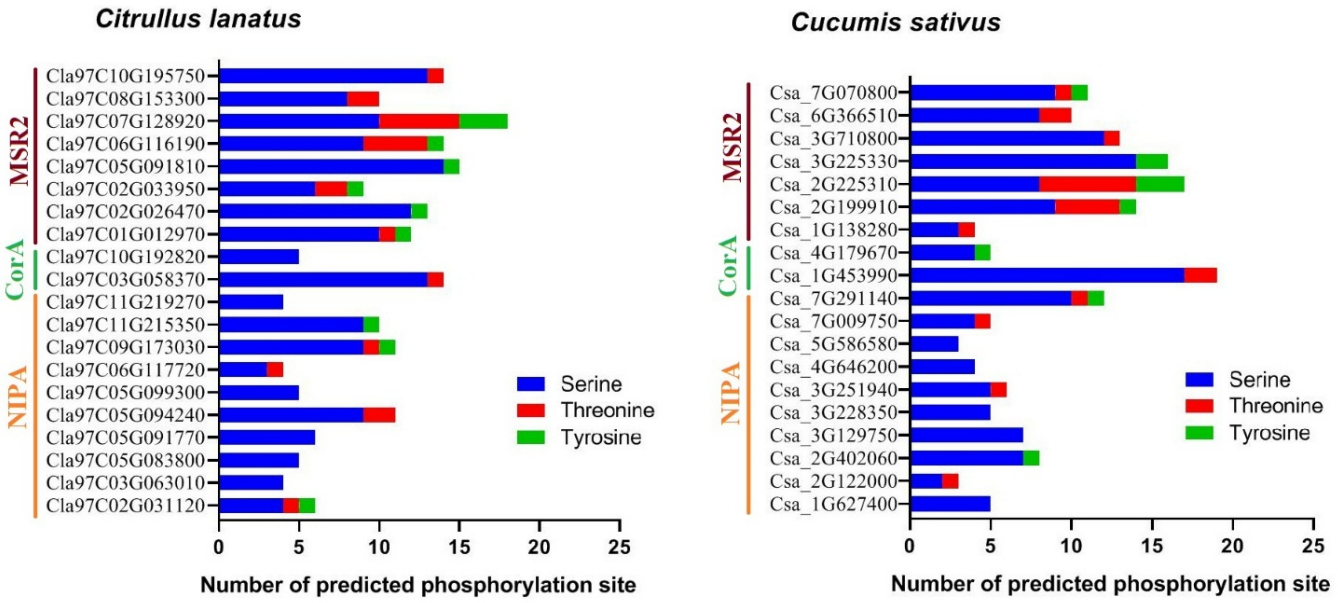

Figure 6. Prediction of phosphorylation site of MGT proteins in *C. lanatus*, and *C. sativus*.

2.7. Co-expression network of MGTs in cucumber

To understand the potential interaction of *MGT* genes with other genes as well as the involved pathways, an interaction network of MGT proteins was drawn in cucumber, *C. sativus* (Figure 7). The results enclosed that NIPAs highly interacted with the protein phosphatase C (PP2C). In addition, some interactions were observed between NIPA and auxin-induced proteins. Besides, MSR2 proteins showed strong interaction with acetyltransferase and CaCA antiporter proteins. Gene ontology (GO) analysis based on interaction network revealed molecular function terms including 8-amino-7-oxononanoate synthase activity, chlorophyllide a oxygenase activity, and transmembrane transporter activity were significantly involved in the MGT-interaction network. Moreover, biological processes such as divalent metal ion transporter process, localization process, and cellular process were linked with the MGT-interaction network.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

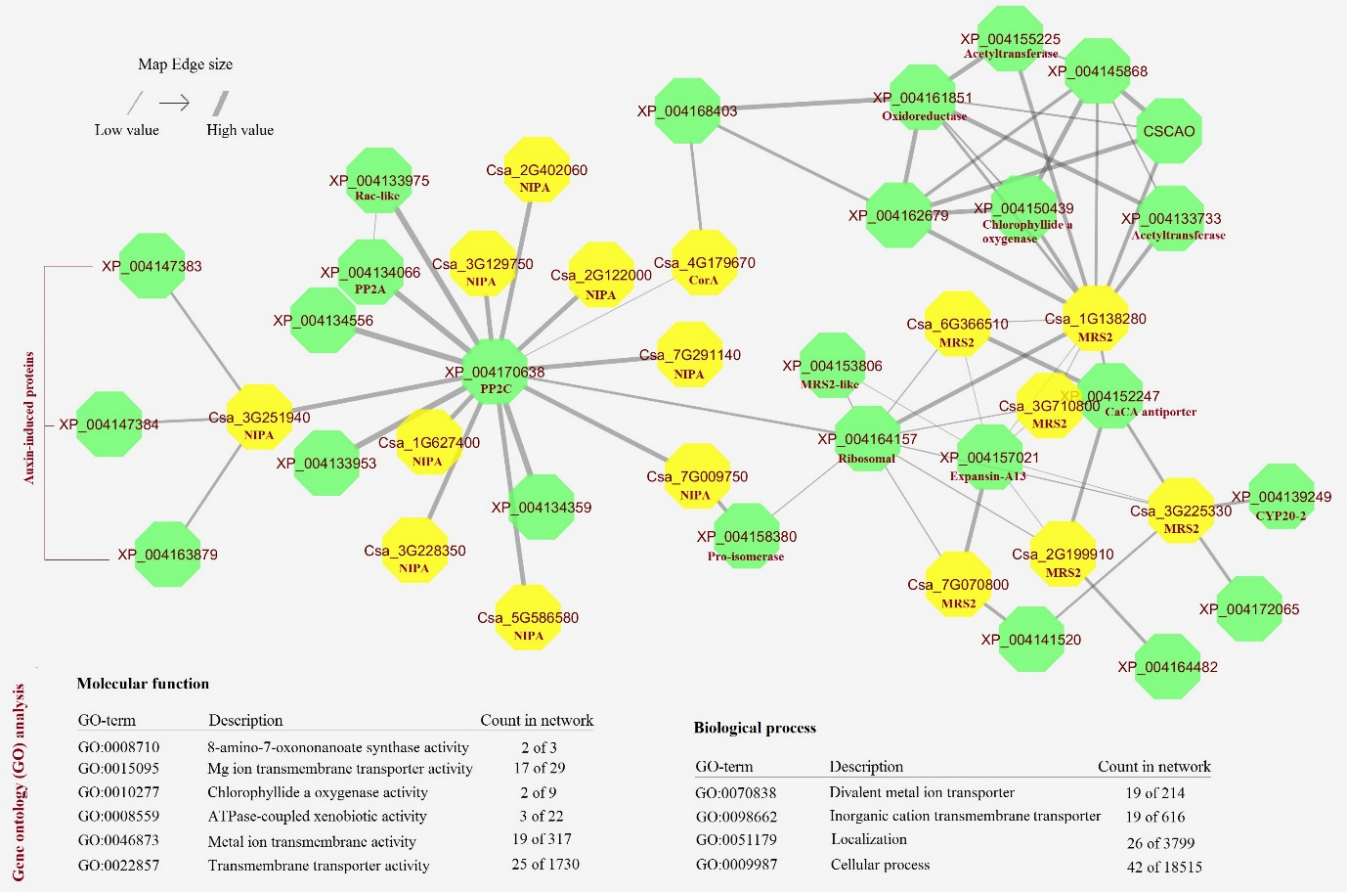

Figure 7. The co-expression network of *MGTs* with other genes in *C. sativus*. The significant GO terms (FDR <0.05) related to the network are provided based on molecular functions and biological process terms.

2.8. Upstream analysis of MGT genes

The upstream region, 1500 bp, of the *MGT* genes *C. lanatus*, and *C. sativus* was analyzed as the promoter region to identify cis-regulatory elements. All identified cis-elements were grouped according to their functions, including light-responsive elements (REs), hormone REs, stress REs, growth REs, and MYB binding site (Figure 8). The cis-regulatory elements related to stress REs were more observed in the promoter region of *MGTs* in *C. lanatus* (Figure 8a). In contrast, hormone REs were more observed in the upstream region of *MGTs* from *C. sativus* (Figure 8b). Besides, abscisic acid (ABA) REs were frequently identified in the promoter region of *MGTs*, more in the cucumber *MGTs*, while auxin REs and methyl jasmonate (MeJA) REs were more observed in in promoter region of *MGTs* from watermelon (Figure 8b and e). Stress REs were classified into four groups, including anaerobic, biotic, drought, and low-temperature stress (Figure 8c and f). Besides, regulatory elements of anaerobic were recognized with high frequency in the promoter region of *MGT* genes. Moreover, low-temperature REs were observed most often in the upstream region of *MGTs* from watermelon (Figure 8c), while regulatory elements responsive to drought stress were more observed in promoter region of cucumber *MGTs* (Figure 8f).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

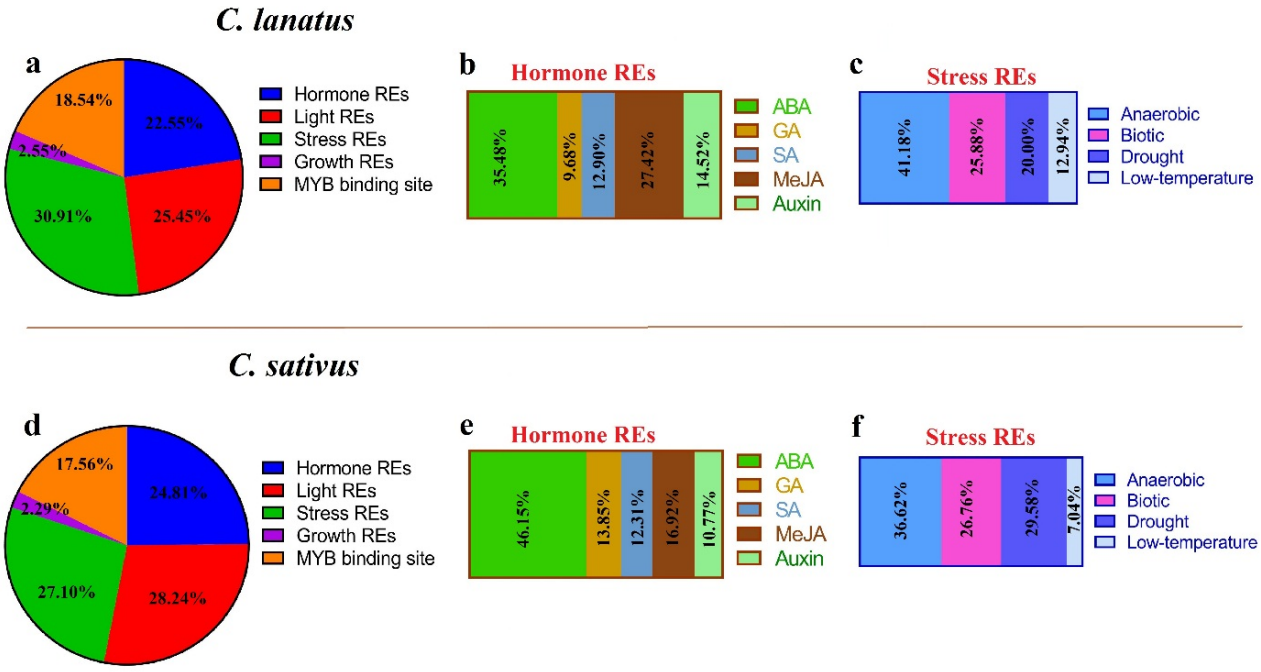

Figure 8. Proportion of cis-regulatory elements in upstream site (promoter regions) of *MGT* genes. The cis-regulatory elements were classified in hormone-responsive elements (REs), light REs, stress REs, growth REs, and MYB binding site in *C. lanatus* (a), and *C. sativus* (d). The percentage of cis-regulatory elements related to hormone REs in *C. lanatus* (b), and *C. sativus* (e), and stress REs in *C. lanatus* (c), and *C. sativus* (f). More details are provided in Table S3.

2.9. Expression profile of MGTs in C. lanatus and C. sativus

The expression patterns of *MGT* genes in different tissues of *C. lanatus* and *C. sativus* were provided based on the RNA-Seq datasets (Figure 9). The *MGTs* in both studied plants showed low expression in seed, while high expression levels of *MGTs* were observed in root tissues. Three *MSR2* genes, including *Csa_2G225310*, *Csa_2G19910*, and *Csa_7G070800*, and three *NIPAs*, *Csa_7G291140*, *Csa_5G586580*, and *Csa_3G251940* showed high expression in shoot tissues of cucumber (Figure 9a). Besides, an *MSR2* gene, *Cla97C02G033950*, showed substantial expression levels in all studied tissues in watermelon (Figure 9b), suggesting the important role of this gene during watermelon growth and expansion. The expression levels of *MGTs* were also investigated according to available RNA-seq datasets related to biotic and abiotic stresses and hormone treatments (Figure 10). The expression levels of a *NIPA*, *Csa_3G129750*, and an *MSR2* gene, *Csa_1G138280*, of cucumber were upregulated in response to NaCl stress (Figure 10a). Besides, in response to the nematode, the expression pattern of *Csa_2G225310*, as an *MSR2* gene, was upregulated in cucumber and *Csa_4G646200*, as a *NIPA* gene, showed an upregulation one day after infection (dai) by *Pseudoperonospora cubensis* (PC) (Figure 10a). In the watermelon, two *NIPA* genes, including *Cla97C05G083800* and *Cla97C03G063010*, were upregulated in response to drought stress and under mosaic virus stress, two *NIPAs*, *Cla97C05G083800* and *Cla97C02G031120*, also showed an upregulation (Figure 10b). Besides, two *NIPAs* genes from watermelon were more induced in response to melatonin treatment and under low nitrogen content in the leaf (Figure 10b). It seems that *NIPAs* are more induced in response to stimuli in watermelon.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2022 doi:10.20944/preprints202206.0278.v1

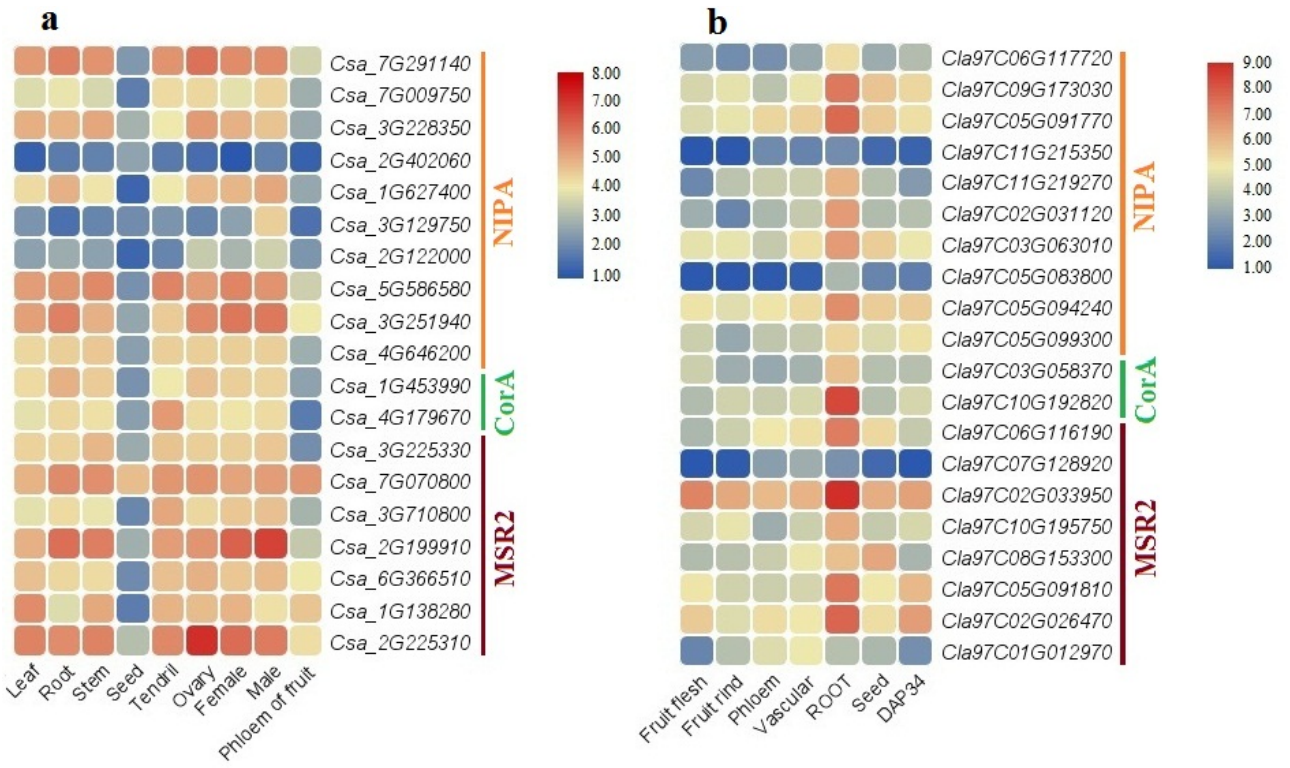

Figure 9. Heatmaps of expression of *MGT* genes in different tissues of *C. lanatus* (a), and *C. sativus* (b).

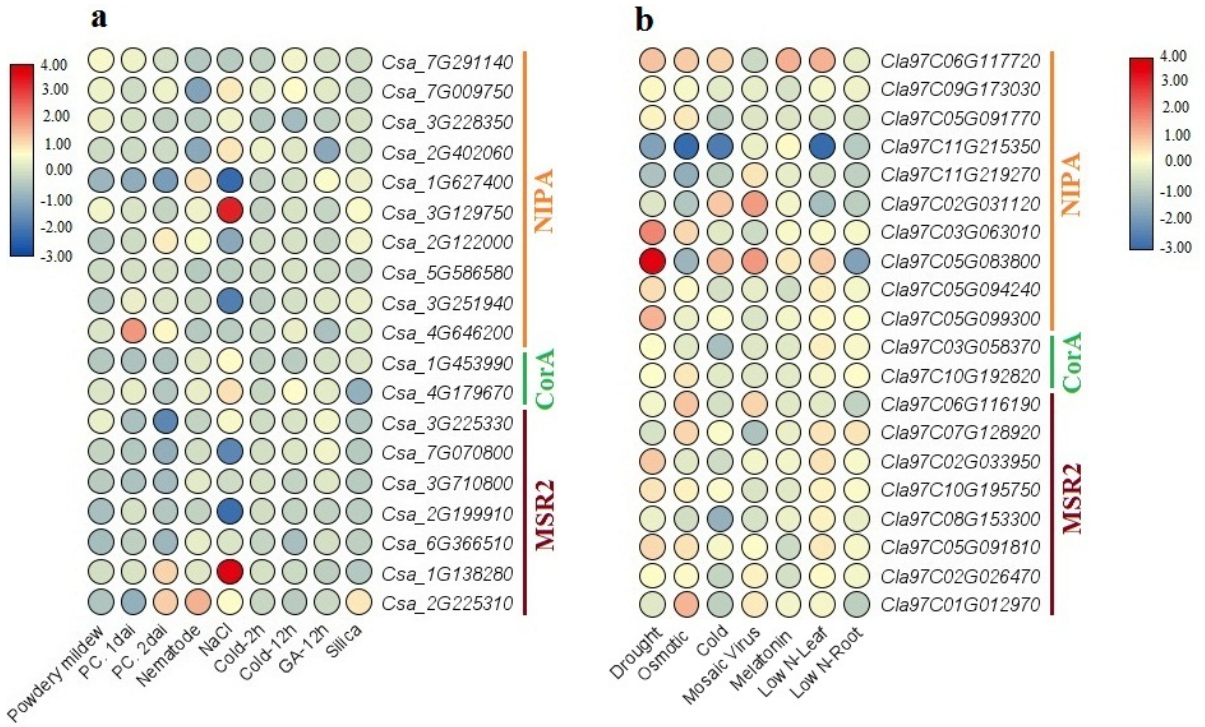

Figure 10. Heatmaps of expression of *MGT* genes respond to biotic and abiotic in *C. lanatus* (a), and *C. sativus* (b).

 doi:10.20944/preprints202206.0278.v1

3. Discussion

Magnesium (Mg), in addition to being a basic element for plant growth as an essential cofactor, is also involved in the activity of enzymes, and in metabolic and photosynthetic processes [29]. Magnesium transporters (MGTs) are fundamental in transmitting and maintaining Mg concentrations in various organelles and cell tissues. In the current study, as the first report, 20, 19, and 20 *MGT* genes were identified and characterized in three species of Cucurbitaceae, including *C. lanatus*, *C. sativus*, and *C. melo*, respectively. This number of genes is less than the number in *G. hirsutum* (41 *MGTs*) [3], *Camelina sativa* (62 *MGTs*) [10], *Triticum turgidum* (41 *MGTs*) [10], and *Brassica napus* (36 *MGTs*) [14], however the number of identified genes is greater than that reported in the cacao (18 *MGTs*) [3], *C. capsularis* (16 *MGTs*) [3], *Pyrus bretschneideri* (16 *MGTs*) [13], *Zea mays* (12 *MGTs*) [15], *Poncirus trifoliata* (8 *MGTs*) [16], and *Fagaria vesca* (12 *MGTs*) [17]. The evolutionary events such as duplication and polyploidization have increased the number of *MGTs* in some plant species [30,31]. However, we did not find any duplication between the MGT family members in *C. lanatus*, and *C. sativus*. We hypothesize that the duplication events probably did not occur after the derivation of these two species. Also, based on the structure of genes and physicochemical characteristics, an almost conserved state was observed between the studied species, further strengthening this hypothesis. Three subclasses of MGTs, including MSR2, CorA, and NIPA, were identified and compared, and NIPAs showed a significant difference from the other two classes. For instance, NIPAs were predicted as stable proteins. In addition, based on GRAVY as a solubility index [31,32], NIPAs were predicted as hydrophilic proteins. There is limited information about the NIPA class in plants, and due to the different structures, it is necessary to do more studies in the field of their functional analysis. Also, based on the analysis of gene structure, members of the MGT family had variations in the number of exons, especially in *NIPAs*. Exon number can increase the diversity of coding protein of a gene by affecting the post-transcriptional processes such as alternative splicing [33,34]. Those with fewer exons can activate rapidly in response to stress, and these genes play a stronger role in the process of adapting to adverse environmental conditions [35]. Evolutionary events such as duplication can affect the structure of genes. Our results reveal that the Ka/Ks ratios of all *MGTs* in both studied species, *C. lanatus* and *C. sativus* were less than 1.0 indicating that purifying (negative) selection was the most important force for motiving the evolution of *MGTs*. Besides, it was stated that most *MGTs* emerged before the disclosure of angiosperms [18]. According to the phylogenetic tree, MGT gene family members can be classified into three main groups; that more diversity was observed among NIPAs orthologous than CorA and MSR2. It seems that the evolutionary processes in *NIPAs* were different from the other two groups, *CorA* and *MSR2*.

The prediction of the 3D structure of MGT proteins showed that each subclass had an almost conserved structure in *C. lanatus*, and *C. sativus*, however differences in their binding sites were observed, indicating a parallel evolutionary trend in *MGTs* of both species. Besides, leucine, valine, serine, isoleucine, and alanine were frequently predicted in pocket sites, suggesting that these residues are more related to possible interactions of MGTs [36,37]. The interaction network of MGTs in *C. sativus* showed that NIPAs interact more with PP2C and auxin-induced proteins. PP2C is a phosphatases involved in ABA signaling and it was reported previously that Mg could induce Mg-dependent phosphatases PP2C heterodimer in response to heat stress [38]. In addition, hormone response elements were observed in upstream regions of the *MGT* genes in both plants, indicating an interaction between mg and phytohormones. Also, interactions between MSR2s and acetyltransferase and CaCA antiporter proteins were observed, that further studies are needed on how they interact in cucumber cells. The phosphorylation process is one of the key post-translational modifications that significantly affect the activity and stability of target proteins and also affect the regulation of cellular signaling pathways in response to adverse conditions [31,39,40]. Results reveal that MSR2 sub-class members have more potential sites for phosphorylation than NIPAs in both studied plant species, *C. lanatus*, and

 doi:10.20944/preprints202206.0278.v1

C. sativus, suggesting that MSR2s have more potential to interact with kinases and other signaling components. Based on expression profile, *MGTs* were expressed in various tissues in *C. lanatus* and *C. sativus* indicating that they are involved in different biological processes. However, most *MGT* genes were expressed in the root tissues, indicating *MGTs* are more involved in the uptake of mg in the root and then the distribution of mg in other tissues. Besides, *MGTs* showed various expression patterns in response to biotic, and abiotic stresses as well as hormone application. Interestingly, results disclose that *NIPA* members are more induced in response to stimuli in watermelon than *MSR2s* and *CorAs*. Findings suggest that this sub-class of *MGTs* may be an appropriate target group for further molecular breeding to release the watermelon-resistant lines. The specialized expression of *MGT* genes can also be related to their promoter region. These genes appear to be more influenced and induced by the pathways dependent on the phytohormones such as ABA, auxin and MeJA. Besides, it was revealed that cell signal transduction associated with hormone concentrations is induced in response to mg toxicity/deficiency [41,42]. Also, previous studies disclosed that *MGTs* interact with Ca^{2+} sensors to induce the downstream signals correlated with plants' reaction to adverse environmental conditions [43]. Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

4. Materials and Methods

4.1. Identification and characterization of MGT genes in C. lanatus, and C. sativus

To identify all sequences related to the MGT family, the MGT proteins of *Arabidopsis thaliana* were used as queries in the BLAST program in Ensembl Plants [44] against the genome of *C. lanatus*, and *C. sativus*. In addition, the orthologue of *MGTs* was identified in *Cucumis melo*, and *Oryza sativa* in the same way. The non-redundant sequences of MGTs were checked using CDD search [45], and Pfam database [46] to validate the presence of MGT domains. To predict the physicochemical properties, including instability index, GRAVY, isoelectric points (pI), and molecular weight (MW) of MGTs, the ProtParam tool was applied [47].

4.2. Phylogenetic Analysis of MGTs

The amino acid sequences of MGTs from *C. lanatus, C. sativus, C. melo, Arabidopsis thaliana* (as a model plant from dicots), and *Oryza sativa* (as a model plant from monocots) were used to construct a phylogenetic tree. In the first step, all sequences were aligned by a multiple alignment tool, Clustal-Omega [48]. Then, the output of the Clustal-Omega was submitted to the IQ-TREE webserver [49] to estimate the phylogenetic relationships of MGTs using the Maximum likelihood (ML) method under 1000 bootstrap replicates. In the final step, the phylogenetic tree of MGT proteins was prepared by the interactive tree of life (iTOL version 5) tool [50].

4.3. Prediction of Ka and Ks

To recognize the duplicated genes, cDNA sequences of MGT genes in *C. lanatus* and *C. sativus* by the ClustalX v.21 program [51]. According to identity matrix, the gene pairs with more than 90% identity were screened as a duplicated gene pairs [52]. In the present study to understand mutations that affected protein sequencing during the evolutionary process, synonymous substitution (K_s) and nonsynonymous substitution (K_a) were investigated for all paired genes of MGT family in *C. lanatus* and *C. sativus*. Ks, Ka, and Ka/Ks were calculated using TBtools software [53].

4.4. Transmembrane structure and pocket site analysis of MGTs

To predict the 3D structure and transmembrane structure of MGTs in *C. lanatus* and *C. sativus*, the amino acid sequences were submitted to the Phyre2 server [54], and the

 doi:10.20944/preprints202206.0278.v1

predicted models with the highest similarity were selected. The pocket sites of each MGT were identified using the Phyre investigator tool of the Phyre2 server.

4.5. Prediction of phosphorylation sites into MGTs

The phosphorylation sites of each MGT protein in *C. lanatus* and *C. sativus* were predicted based on three amino acids, including serine (S), tyrosine (Y), and threonine (T), using NetPhos 3.1 Server [55]. To predict the sites to a high percentage of confidence, the score was adjusted to the scores of more than 0.90.

4.6. Protein-protein interaction network

To construct the protein-protein interaction network between MGTs in *C. sativus*, the sequences of all MGTs were submitted to the STRING v11.5 database [56]. A maximum number of interactors was adjusted to no more than 5 interactors for the first shell and no more than 20 interactors for the second shell. Finally, the interaction networks were illustrated using Cytoscape v3.8.2 [57].

4.7. Promoter Analysis of MGT genes

To identify the known cis-regulatory elements related to the response of hormones and stresses as well as involved in growth, the upstream region (1500 bp before the start codon) of each *MGT* gene in *C. lanatus* and *C. sativus* was screened by the PlantCARE tool [58]. Finally, cis-regulatory elements were grouped according to their function.

4.8. Gene expression profile of MGT genes

To extract the expression patterns of MGT genes in *C. lanatus* and *C. sativus*, the available RNA-seq data from CuGenDBv1 (http://cucurbitgenomics.org/) was used. Three RNA-seq datasets of different tissues of *C. sativus*, including PRJNA80169 (leaf, stem, root, tendril, ovary, female, and male), PRJNA319011 (seed), and PRJNA263870 (phloem of fruit) were analyzed to extract the expression levels of MGTs. Besides, the RNA-seq data of *C. lanatus* related to root tissue (PRJNA209092), 34 days after pollination (DAP) in fruit tissues (PRJNA221197), fruit flesh and fruit rind (SRP012849), seed (PRJNA319011), phloem and vascular tissues (SRP012853) were used to find out the expression profile of *MGTs*. In addition to understanding the response of *MGTs* to biotic/abiotic stresses and exogenous application of hormones/elicitors, the RNA-seq datasets of *C. sativus* related to cold stress after 2h and 12h (PRJNA438923), NaCl (PRJNA437579), silica (PRJEB7612), GA at 12h (PRJNA376073), nematode infection (PRJNA419665), powdery mildew infection (PRJNA321023), and one and two days after infection with *Pseudoperonospora cubensis* (PRJNA285071) and the RNA-seq datasets of *C. lanatus*, including PRJNA326331 (osmotic stress), PRJNA454040 (drought stress), PRJNA389184 (mosaic virus), PRJNA328189 (cold stress and melatonin application), and PRJNA422970 (low nitrogen (N) stress in leaf and root) were used and analyzed. The expression data of *MGTs* were extracted based on FPKM values. The expression profile of the *MGTs* was illustrated in heatmaps based on log2 transformed method of FPKM+1 for expression in tissues and the log2 fold change in response to stresses and hormones application using TBtools software.

5. Conclusions

In the present study, MGT gene family members were identified and analyzed in two candidate species of Cucurbitaceae, *C. sativus* and *C. lanatus*, as the first report. Results reveal that a similar evolutionary process for the MGT gene family members has probably occurred in *C. lanatus* and *C. sativus*, and duplication events between *MGTs* may have occurred before the derivation of these two species. The NIPA class showed great structural diversity and different expression patterns from the MSR2 and CorA groups that should be considered more in future studies.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, **Table S1.** List of the identified *MGT* genes and their characteristics in watermelon *(Citrullus lanatus),* cucumber (*Cucumis sativus),* and melon (*Cucumis melo).* **Table S2.** List of ligand-binding sites in the predicted 3D structure of MGTs in *C. lanatus*, and *C. sativus***. Table S3.** Promoter important cis elements engaged in various developmental and stress responsive pathways in MGT genes of watermelon and cucumber.

Author Contributions: Conceptualization, P.H and F.P.; methodology, P.H..; software, P.H., F.P. and F.M-P.; writing—original draft preparation, P.H. and F.M-P.; writing—review and editing, P.H., F.P. and F.M-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Please refer to suggested Data Availability Statements in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics. If the study did not report any data, you might add "Not applicable" here.

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. Heidari, P.; Abdullah; Faraji, S.; Poczai, P. Magnesium transporter Gene Family: Genome-Wide Identification and Characterization in Theobroma cacao, Corchorus capsularis and Gossypium hirsutum of Family Malvaceae. *Agronomy* **2021**, *11*, 1651.

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

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

6. Shaul, O. Magnesium transport and function in plants: the tip of the iceberg. *Biometals* **2002**, *15*, 307–321.

7. Black, J.R.; Yin, Q.; Casey, W.H. An experimental study of magnesium-isotope fractionation in chlorophyll-a photosynthesis. *Geochim. Cosmochim. Acta* **2006**, *70*, 4072–4079.

8. Smith, R.L.; Gottlieb, E.; Kucharski, L.M.; Maguire, M.E. Functional similarity between archaeal and bacterial CorA magnesium transporters. *J. Bacteriol.* **1998**, *180*, 2788–2791.

9. Li, L.; Tutone, A.F.; Drummond, R.S.M.; Gardner, R.C.; Luan, S. A novel family of magnesium transport genes in Arabidopsis. *Plant Cell* **2001**, *13*, 2761–2775.

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

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. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. 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.

23. 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.

24. 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.

25. 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.

26. 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.

27. Schaefer, H.; Renner, S.S. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). *Taxon* **2011**, *60*, 122–138.

28. Rolnik, A.; Olas, B. Vegetables from the Cucurbitaceae family and their products: Positive effect on human health. *Nutrition* **2020**, *78*, 110788.

29. Guo, W. Magnesium homeostasis mechanisms and magnesium use efficiency in plants. In *Plant macronutrient use efficiency*; Elsevier, 2017; pp. 197–213.

30. 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.

31. Rezaee, S.; Ahmadizadeh, M.; Heidari, P. Genome-wide characterization, expression profiling, and post-transcriptional study of GASA gene family. *Gene Reports* **2020**, *20*, 100795.

32. Ahmadizadeh, M.; Rezaee, S.; Heidari, P. Genome-wide characterization and expression analysis of fatty acid desaturase gene family in Camelina sativa. *Gene Reports* **2020**, *21*, 100894.

33. Iñiguez, L.P.; Hernández, G. The evolutionary relationship between alternative splicing and gene duplication. *Front. Genet.* **2017**, *8*, 14.

34. Koralewski, T.E.; Krutovsky, K. V Evolution of exon-intron structure and alternative splicing. *PLoS One* **2011**, *6*, e18055.

35. Jeffares, D.C.; Penkett, C.J.; Bähler, J. Rapidly regulated genes are intron poor. *Trends Genet.* **2008**, *24*, 375–378.

36. 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.

37. 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.

38. Timucin, E.; Sezerman, O.U. Thermostability of the PYL–PP2C Heterodimer Is Dependent on Magnesium: In Silico Insights into the Link between Heat Stress Response and Magnesium Deficiency in Plants. *J. Chem. Inf. Model.* **2018**, *58*, 661–672.

39. Heidari, P.; Ahmadizadeh, M.; Izanlo, F.; Nussbaumer, T. In silico study of the CESA and CSL gene family in Arabidopsis thaliana and Oryza sativa: Focus on post-translation modifications. *Plant Gene* **2019**, *19*.

40. 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.

41. Sheng, Y.; Ding, Y.; Fu, Y.; Yang, D.; Liang, Z.; Guo, W. The research development of the response mechanisms to magnesium stresses in plants 2015.

42. Guo, W.; Cong, Y.; Hussain, N.; Wang, Y.; Liu, Z.; Jiang, L.; Liang, Z.; Chen, K. The remodeling of seedling development in response to long-term magnesium toxicity and regulation by ABA–DELLA signaling in Arabidopsis. *Plant Cell Physiol.* **2014**, *55*, 1713–1726.

43. 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.

44. 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.

45. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222–D226.

46. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J. Pfam: the protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230.

47. 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.

48. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.

49. 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.

50. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296.

51. 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.

52. Faraji, S.; Heidari, P.; Amouei, H.; Filiz, E.; Poczai, P. Investigation and Computational Analysis of the Sulfotransferase (SOT) Gene Family in Potato (Solanum tuberosum): Insights into Sulfur Adjustment for Proper Development and Stimuli Responses. *Plants* **2021**, *10*, 2597.

53. 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.

54. 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.

55. Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **1999**, *294*, 1351–1362.

56. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613.

57. Franz, M.; Lopes, C.T.; Huck, G.; Dong, Y.; Sumer, O.; Bader, G.D. Cytoscape. js: a graph theory library for visualisation and analysis. *Bioinformatics* **2016**, *32*, 309–311.

58. 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.