Inconsistent evolution and expression divergence of three key enzymes of the Met biosynthesis in plants: CGS, HMT and MMT

Man Zhao [Corresp., 1], **Wenyi Wang** [2], **Lei Wei** [3], **Peng Chen** [3], **Fengjie Yuan** [4], **Zhao Wang** [3], **Xiangxian Ying** [3]

[1] College of Biotechnology and Bioengineering, Zhejiang University of Technology, Hangzhou, zhejiang province, China
[2] College of Biotechnology and Bioengineering, Zhejiang University of Technology, Hangzhou, Zhejiang province, China
[3] College of Biotechnology and Bioengineering, Zhejiang University of Technology, hangzhou, Zhejiang province, China
[4] Institute of Crop Science, Zhejiang Academy of Agricultural Sciences, hangzhou, zhejiang province, China

Corresponding Author: Man Zhao
Email address: mzhao@zjut.edu.cn

Methionine (Met) is an essential sulphur-containing amino acid in humans. Cystathionine γ-synthase (CGS), methionine methyltransferase (MMT) and homocysteine methyltransferase (HMT) are committing enzymes, which synergistically synthesize Met through the aspartate (Asp) family pathway and the S-methylmethionine (SMM) cycle. The biological functions of *CGS*, *MMT* and *HMT* genes have been studied, yet how they cooperate with each other to synthesize methionine during evolution in plants is unknown. In the present study, a reconstruction of the evolutionary relationship of *CGS*, *MMT* and *HMT* gene families showed that they inconsistently diverged. *MMT*s were relatively conserved, while the grass and seed plant levels of *CGS*s and *HMT*s diverged. A gene structure analysis indicated that the protein motifs and intron–exon numbers of *MMT*s and *HMT*s were conserved, but the *CGS*s of Class 2 diverged. All of the genes were under strong negative selections, although their expression patterns in soybeans, as well as in the database, had varying levels of divergence. The e xpression patterns of the *MMT*s were basically conserved, but the *CGS*s and *HMT*s had tissue-specific divergence in their grass and seed plant levels. Taken with the evolution and expression results, the data implied that the *CGS*, *HMT* and *MMT* gene families have experienced inconsistent evolutionary and expression divergence, which might be vital to the supply of Met in the growth and development of different botanical lineages during evolution.

Research article

Inconsistent evolution and expression divergence of three key enzymes of the Met biosynthesis in plants: CGS, HMT and MMT

Man Zhao[1,*], Wenyi Wang[1], Lei Wei[1], Peng Chen[1], Fengjie Yuan[2], Zhao Wang[1], Xiangxian Ying[1,*]

1 College of Biotechnology and Bioengineering, Zhejiang University of Technology, 310014 Hangzhou, China;

2 Institute of Crop Science, Zhejiang Academy of Agricultural Sciences, 310014 Hangzhou, China

*Correspondence: mzhao@zjut.edu.cn (M.Z.). yingxx@zjut.edu.cn (X.Y.)

Author email list

M.Z.: mzhao@zjut.edu.cn

W.W.: hzwywang@163.com

L.W.: leiweistudy@163.com

P.C.: chenpengoffice@163.com

F. Y.: fjyuanhz@126.com

Z.W.: hzwangzhao@163.com

X.Y.: yingxx@zjut.edu.cn

ABSTRACT

Methionine (Met) is an essential sulphur-containing amino acid in humans. Cystathionine γ-synthase (CGS), methionine methyltransferase (MMT) and homocysteine methyltransferase (HMT) are committing enzymes, which synergistically synthesize Met through the aspartate (Asp) family pathway and the S-methylmethionine (SMM) cycle. The biological functions of *CGS*, *MMT* and *HMT* genes have been studied, yet how they cooperate with each other to synthesize methionine during evolution in plants is unknown. In the present study, a reconstruction of the evolutionary relationship of *CGS*, *MMT* and *HMT* gene families showed that they inconsistently diverged. *MMT*s were relatively conserved, while the grass and seed

plant levels of *CGS*s and *HMT*s diverged. A gene structure analysis indicated that the protein motifs and intron–exon numbers of *MMT*s and *HMT*s were conserved, but the *CGS*s of Class 2 diverged. All of the genes were under strong negative selections, although their expression patterns in soybeans, as well as in the database, had varying levels of divergence. The expression patterns of the *MMT*s were basically conserved, but the *CGS*s and *HMT*s had tissue-specific divergence in their grass and seed plant levels. Taken with the evolution and expression results, the data implied that the *CGS*, *HMT* and *MMT* gene families have experienced inconsistent evolutionary and expression divergence, which might be vital to the supply of Met in the growth and development of different botanical lineages during evolution.

Subjects Plant biology, Molecular biology, Evolutionary biology

Keywords Methionine biosynthesis; *CGS* gene family; *MMT* gene family; *HMT* gene family; Evolution; Gene expression

INTRODUCTION

Met is an essential amino acid, which is mainly obtained from human and animal foods. Met plays important functions, not only as a protein component and in the initiation of mRNA translation, but also in indirectly regulating various metabolic processes through its main catabolic product, *S*-adenosylmethionine (SAM, AdoMet) (Galili & Amir, 2013; Roje, 2006; Sauter et al., 2013). Despite its important functions, there is a limited quantity of Met in major crops (Galili & Amir, 2013).

The biosynthetic pathway of Met has been widely studied in plants and can be synthesized through the Asp family pathway as well as the SMM cycle. In the Asp family pathway process, homoserine is firstly converted into O-phosphohomoserine (OPH) by homoserine kinase (HSK). Then, the condensation reaction of OPH with cysteine is catalyzed into cystathionine by cystathionine gamma-synthase (CGS), and cystathionine is hydrolyzed into homocysteine through cystathionine beta-lyase. Next, Met is synthesized de novo through methionine synthase (Datko et al., 1974). As for the SMM cycle, MMT uses Met, synthesized by the Asp family pathway, and SAM to form SMM, then SMM and homocysteine (Hcy) are converted into two molecules of Met through the catalysis of HMT (Bourgis et al., 1999; Ranocha et al., 2008; Lee et al., 2008; Cohen et al., 2017a).

Furthermore, the biosynthesis process of Met during the development of plants is revealed

by genetic and biochemical experiments. First, Met is synthesized by the Asp family pathway in rosette leaves, in which it is converted into SMM by MMT; second, the SMM is translocated into reproductive tissues, such as siliques and seeds, and reconverted back into Met in the developing seeds by HMT (Bourgis et al., 1999; Ranocha et al., 2001; Lee et al., 2008; Cohen et al., 2017b; Kocsis et al., 2003). Additionally, the possible contribution of SMM to the stress effects was proposed (Cohen et al., 2017b). Above all, CGS, HMT and MMT are essential enzymes in the synthesis of Met.

CGS is the key regulatory enzyme in the Asp family pathway. In *Arabidopsis*, when CGS was constitutively over-expressed, the soluble Met and SMM accumulated in specific stages, such as flowers, siliques, seedling tissues and roots of mature plants (Kim et al., 2002). Contrarily, the repression of *CGSs* made the plants abnormal and produced partial methionine auxotrophy (Kim & Leustek, 2000). Interestingly, when the seed-specific repression of *CGS* was performed, more SMMs were transported from the leaves to reproductive organs, in which there were higher reconversion rates of SMM to Met, and more Met was accumulated in seeds (Cohen et al., 2017b). In addition, studies have reported that the expressions of *CGS* were in the negative feedback regulation of their products, Met or SAM, in wild-type *Arabidopsis* (Kim & Leustek, 2000; Thompson et al., 1982; Ranocha et al., 2000). Further, the MTO1 region in the first exon of *AtCGS* was proved to result in its negative feedback regulation (Chiba et al., 1999; Ominato et al., 2002). In *mto1* mutants of *AtCGS1*, both the enzyme levels and soluble Met levels were increased (Chiba et al., 1999). Afterwards, the seed-specific expression of the feedback-insensitive form of *AtCGSs* in plants were also studied, but with different results (Cohen et al., 2017b; Hanafy et al., 2013; Song et al., 2013; Cohen et al., 2016; Matityahu et al., 2013; Cohen et al., 2014). For example, in *Arabidopsis*, soybean and tobacco, the sulfur-associated metabolism was altered and the soluble Met was significantly elevated in seeds. However, there was no Met increase in azuki bean (Matityahu et al., 2013). Therefore, the *CGS* gene family might have diverged in different organisms during evolution.

HMT and MMT are essential in the SMM cycle (Cohen et al., 2017a; Zhao et al., 2018). The evolution and expression of *HMT*s have been studied, as detailed in our previous research. Research found that *HMT*s have diverged into two clades in seed plants and that their expression was tissue-specific. It has been proposed that the divergence of *HMTs* might be crucial to meeting the needs of plant development and growth (Zhao et al., 2018). As for MMT, it was only

studied in *Arabidopsis* by catalyzing the synthesis of SMM from Met and AdoMet (Ranocha, 2001). Nevertheless, the systematic evolution pattern and cooperative functional models of the three key enzymes CGSs, HMTs and MMTs in plants are unclear.

Soybean is an important economic crop, as it is a source of vegetable proteins in the human diet. In soybean seeds, major storage proteins consist of glycinin (11S) and conglycinin (7S), and 11S proteins account for approximately 30% (Nielsen et al., 1989; Harada et al., 1989). Considering the importance of CGS, MMT and HMT in the synthesis of Met, this study comprehensively analyzed their evolutionary history, including their phylogenetic relationship and gene structures, and examined their selection pressures. Their expression profiles in soybeans and other plants were also widely analyzed. Taken together, this research is helpful for understanding the evolutionary history and functional divergence of the *CGS*, *MMT* and *HMT* gene families in plants, and might also provide an overall picture of the evolutionary and functional model of the Met biosynthesis pathway in plants.

MATERIALS AND METHODS

Phylogenetic analysis

The gene sequences in full genomes of plants were examined with genes in *A. thaliana* as query. The sequences followed the criteria: E-value < 1×e-05 in the BLASTN and TBLASTN programs, and an amino acid identity above 40%, which were downloaded from the databases of Phytozome (http://www.phytozome.net/), congenie (http://congenie.org/) and NCBI (https://www.ncbi.nlm.nih.gov/). Altogether, 49 *CGS* sequences and 43 *MMT* sequences were obtained from the major plant lineages studied (Tables S1,S2 in File 3). Multiple alignments of gene sequences were executed in the Clustal X v1.81 program with default parameters and alignments, optimized via manual adjustments using BioEdit v 7.0.9.0 (Thompson et al., 1997; Hall, 1999). Maximum likelihood (ML) and Neighbor-Joining (NJ) trees were reconstructed using PhyML online with the GTR + G + I model and MEGA6 software (Guindon & Gascuel, 2003; Tamura et al., 2013). The resultant trees were represented using MEGA 6. The phylogenetic tree of *HMT* genes has been shown in our previous study (Zhao et al., 2018).

Analysis of gene structure

The intron and exon structures of *CGS* and *MMT* genes were analyzed according to their genome sequences and coding sequences. The length and numbers of introns and exons are

shown in Table S3 in File S3. In addition, the conserved motifs in proteins were detected using the Multiple Em for Motif Elicitation (MEME) server (http://meme-suite.org/tools/meme) (Bailey et al., 2009). The server was run using the default values and choices. We conducted the search for 16 motifs in proteins arbitrarily. The motifs retrieved by MEME were reported according to their statistical significance, and the most statistically significant (low E-value) ones were shown first. The E-value of a motif is based on its log likelihood ratio, width, sites, and the size of the set. The motifs of HMTs have been analyzed in our previous results (Zhao et al., 2018).

Detection of selection pressures

To estimate the selection pressures in the gene families, the codeml program from the PAML v4.4 package, on the basis of codon sequence alignments, was performed (Yang, 2007). The likelihood ratio test (LRT) is a general method for testing assumptions (model parameters) by comparing two competing hypotheses.

Plant materials and growth conditions

The cultivated soybean 'Chuandou 4' was grown at a farm in Fuyang (Hangzhou, China) during summer. Each materials of the leaves, stems, flowers, and 2-, 3-, 4-, 5- and 6-week post-fertilization fruits for the gene expression study were harvested at the same time. The harvested tissues were immediately stored in liquid N2 and then stored at −80 °C for total RNA extraction using TRIzol reagent (Invitrogen).

Realtime RT-PCR analyses

Two micrograms of total RNA were used to synthesize the first strand cDNA using a ReverTra Ace qPCR RT Kit cDNA Synthesis Kit (TOYOBO). Quantitative RT-PCR (qRT-PCR) was conducted using a ChamQ TM SYBR qPCR Master Mix (Vazyme) in a CFX Connect Real-Time system (BIO-RAD). ACTIN (Glyma.18G290800) was used as an internal control. Each experiment was performed using three independent biological samples. PCR was performed in a 25.0 μL reaction mixture containing 5 μL ChamQ TM SYBR qPCR Master Mix (Vazyme), 50 ng cDNA template, 0.4 μL of each primer (10.0 μM) and 3.2 μL of double distilled H_2O (dd H_2O). The optimized operational procedure was performed as follows: 2 min at 95 °C (1 cycle), 10 s at 95 °C, 30 s at 60 °C (40 cycles), 5 s at 65 °C and 5 s at 95 °C (1 cycle for the melting curve analysis). The relative gene expression was evaluated as previously described (Livak & Schmittgen, 2001).

The expression of genes in different tissues was analyzed in the PLEXdb database (http://www.plexdb.org/index.php) (Dash et al., 2012).

Promoters analysis

The promoter sequences (2000 bp upstream of the transcription initiation site), *GmaHMT*s, *GmaCGS*s and *GmaMMT*s, were obtained from Phytozome. To identify the putative *cis*-acting regulatory elements, the promoter sequences of *GmaHMT*s, *GmaCGS*s and *GmaMMT*s were submitted to PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002).

Statistical analysis

In this study, standard deviations were calculated based on a minimum of three independent replicates. Comparative statistical analyses of groups were performed using Student's t test.

RESULTS

Identification and phylogenetic analysis of *CGS*, *MMT* and *HMT* genes

In this study, we reconstructed the phylogenetic trees of the *CGS* and *MMT* genes in plants to understand their evolutionary history. The genes from representative whole-genome plants lineages, which contained monocots, eudicots, basal angiosperms, gymnosperms, basal land plants and chlorophyta, were surveyed (Tables S1,S2 in File S3). The phylogenetic trees were constructed using maximum likelihood (ML) and Neighbor-joining (NJ) methods (File S1). Due to the similar topologies of ML and NJ trees, ML trees were shown with higher support values (Fig. 1, Figs. S1A-B in File S2).

First, the *CGS* genes were widely separated in plant lineages from algeas to angiosperms. In total, 49 representative *CGS*s were used to reconstruct the phylogenetic tree, and the *CGS* genes in algeas, as the outgroup, were located at the base. The *CGS* genes in gymnosperms and basal land plants had not diverged and were grouped in Class 3, while the *CGS*s had diverged into two classes in angiosperms, Class 1 and Class 2 (Fig. 1, Fig. S1A in File S2). In Class 1, all of the genes were contained in angiosperms, whereas, in Class 2, only the genes contained in grasses were present. The results indicated the *CGS* genes might have diverged asynchronously in angiosperms.

However, *MMT* genes were not found in algeas and were relatively conserved, with only one or two copies, except for *PpaMMT*s with 3 copies. In total, 43 *MMT*s were surveyed.

Phylogenetically, the evolutionary relationship of *MMT* genes with their species relationship was relatively consistent (Fig. 1, Fig. S1B in File S2). As for the *HMT* genes, we have reported that they existed in kinds of plant lineages at the base of algeas in the evolutionary tree. Notably, they have diverged into two classes in seed plants (Fig. 1) (Zhao et al., 2018). Therefore, the phylogenetic relationships between *CGS*, *MMT* and *HMT* genes showed that the *MMT*s were conserved, while the *HMT*s and *CGS*s diverged in grasses and seed plants. The results implied that these genes have experienced inconsistent divergence during their evolution.

Intron–exon structures of *CGS*s, *MMT*s and *HMT*s

The divergence of genes is reflected in their structures to a certain extent. The intron–exon structures, as well as the number and length of introns and exons, in the *CGS*, *MMT* and *HMT* genes were analyzed (Table S3 in File S3).

The analysis of the number of introns and exons showed that most *CGS* genes in land plants contained an 11-exon and 10-intron pattern (Table S3A in File S3). For example, in Class 1, 26 (26/35 = 74.3%) genes maintained the pattern and the remaining 9 had different degrees of intron gains or losses. In Class 3 all of the *CGS*s belonged to the 11-exon and 10-intron pattern, except for the unknown *PglCGS* and *PsiCGS*. However, in Class 2 the number of exons was less than 6 and half of them (5/10 = 50%) contained only 2 exons. The results showed that the divergence in exon–intron numbers have occurred between Class 2 and other classes in *CGS* genes. In addition, 76.2% (32/42 = 76.2%) of *MMT* genes contained a 12-exon and 11-intron pattern, while the remaining 23.8% (10/42 = 23.8%) experienced intron gain or loss events of different degrees (Table S3B in File S3). Generally, the exon–intron pattern of *MMT* genes was relatively conserved during evolution. The intron–exon numbers in *HMT* genes were also relatively conserved: 70.67% *HMT*s had a 7-exon and 6-intron pattern (Zhao et al., 2018).

Besides the numbers, the length of exons and introns was also considered in our study. In the three gene families, the lengths of exons were basically consistent, except for the divergent clades, while the corresponding lengths of introns were divergent (Table S3 in File S3). Finally, the analysis of the exon–intron structures indicated that the structures of the *HMT* and *MMT* genes were basically conserved, while the intron–exon numbers in *CGS* genes diverged, especially in Class 2.

Protein motifs analysis in MEME

Protein structures were analyzed to survey the conserved protein motifs of CGSs, MMTs and HMTs in MEME. In total, 16 motifs were identified and shown in CGSs (Fig. 2, Fig. S2A in File S2). Among them, 12 motifs (motif 1 to motif 6 and motif 9 to motif 14), located in the middle and C-terminal of the CGS proteins, were found in all CGS proteins (Fig. 2, Fig. S2A in File S2). However, the motifs in N-terminals, such as motif 7, 8, 15 and 16, were divided among classes. For instance, motifs 7, 8 and 16 were in Class 1 and Class 3, while motif 15 was in Class 2. The MTO1 region is essential for the negative feedback regulation of *CGS* genes, which is located in the N-terminals of CGS (marked in red lines in Fig. 3 and File S1). In this study, the MTO1 region was only found in motif 8. Hence, the CGSs in Class 2 had lost their MTO1 regions during evolution (Fig. 3). In addition, some *CGS*s in Class 1, such as *AthCGS2* and *BraCGS2*, lost their MTO1 regions. Furthermore, three key sites in the MTO1 region (R77, S81 and G84 in *AtCGS1*) were not detected in Class 2, AthCGS2 or BraCGS2. In view of the functions of the MTO1 region, the results indicated that the negative feedback regulation might have been lost in Class 2, *AthCGS2* and *BraCGS2*.

However, in MMT proteins, 16 motifs were totally consistent in all of the MMTs, except BdiMMT1 and RcoMMT, which indicated that the protein motifs of MMTs were conserved during evolution (Fig. 4, Fig. S2B in File S2). Similarly, the protein motifs of HMTs were also basically conserved (Zhao et al., 2018). Based on the results above, the divergence of protein motifs has occurred in CGS proteins, but not in HMTs and MMTs.

The selection pressure of the *CGS*, *HMT* and *MMT* family

Selection pressure refers to the function change of genes during evolution. To analyze the selection pressure of the gene families, the ω values ($\omega = dN/dS$) were estimated, and the ω value was defined as the ratio of nonsynonymous and synonymous substitution. The results showed that the ω values of *CGS*s, *MMT*s and *HMT*s were 0.19, 0.17 and 0.16, respectively. The selection pressures showed that they were under stringent negative selection during evolution, and hence their functions were relatively conserved during evolution.

qRT-PCR analysis of *CGS*, *MMT* and *HMT* genes in soybean

The expression of genes could reflect their functional divergence to some extent. To verify their expression patterns, we analyzed the expression of *CGS*s, *MMT*s and *HMT*s in soybean (Fig. 5A–C). In this study, the organs of leaves, stems, flowers and 2w–6w pods were collected and analyzed.

The expression patterns of *GmaCGS1* and *GmaCGS2* were similar. Both of them were highly expressed in leaves and flowers, but significantly decreased during the development process of pods (Fig. 5A). Similarly, the expression models of *GmaMMT*s were analogous, significantly highly expressed in stems, leaves, flowers and 2-week pods, and gradually decreased during the development of pods (Fig. 5B). However, in *GmaHMT*s, the expression patterns were varied (Fig. 5C). For example, the expression of *GmaHMT1* and *GmaHMT3* was significantly higher in the pods and flowers than in the leaves and stems. On the contrary, the *GmaHMT2* was fluctuant in different organs, such as leaves, stems, flowers and pods, and the expression levels of *GmaHMT4* were significantly higher in leaves than in flowers, stems and different pods. Above all, the expression patterns within *GmaCGS*s and *GmaMMT*s were relatively consistent, yet the *GmaHMT*s were distinct in different copies. The distinct expression patterns of the three gene families might be essential to supplying methionine for the growth and development of soybeans.

Expression profiles of *CGS*s, *HMT*s and *MMT*s in PLEXdb

To further investigate the gene expression patterns, the tissue expression profiles of *CGSs*, *HMTs* and *MMT*s were analyzed in *Arabidopsis*, soybean and rice in the PLEXdb database (Fig. S3A–C in File S2). First, the expression patterns of *CGS*s were analyzed. In *Arabidopsis*, both *AthCGS1* and *AthCGS2* were fluctuant in all of the tissues, while their expression levels were generally higher in vegetative tissues than in productive tissues (Fig. S3A in File S2). Nevertheless, the expression intensity of *AthCGS1* (11–14) and *AthCGS2* (3–7) was different. In soybean, only *GmaCGS1* was detected in the database. The expression trends of *GmaCGS1* were similar in the qRT-PCR results, and it was highly expressed in leaves but gradually decreased in seeds and pods (Fig. S3B in File S2). However, the expression of *OsaCGS*s was varied. For instance, the expression of *OsaCGS1* was fluctuant in vegetative and productive tissues. The *OsaCGS3* was highly expressed in vegetative tissues, such as leaves, roots and seedlings, while the *OsaCGS5* was higher in endosperms than in vegetative tissues. It is worth noting that their expression intensities were also different, and the highest was found in *OsaCGS3* (intensity from 11 to 14), followed by *OsaCGS1* (intensity from 5 to 8) and *OsaCGS5* (intensity from 1 to 4).

In *MMT* genes, the expression of *OsaMMT* was not detected (Fig. S4A,B in File S2). In *Arabidopsis*, the expression levels of *AthMMT* were basically consistent in different tissues, except for seeds. In seeds, the expression level of *AthMMT* was lower than in other vegetative

and productive tissues (Fig. S4A in File S2). In soybean, the expression of *GmaMMT1* was high in vegetative tissues and the early stage of seeds, but low in fully grown pods. However, the *GmaMMT2* in different tissues was basically consistent, with small fluctuations (Fig. S4B in File S2). The expression of *HMTs* in the database has also been comprehensively analyzed in our previous article (Zhao et al., 2018). The results showed that *HMT*s have distinct expression patterns, which is confirmed by the qRT-PCR results in this study. Some *HMT*s were widely expressed in different tissues, while others were particularly highly expressed in specific tissues, such as seeds or leaves. It is worth noting that the expression divergence of *HMT*s was not clade-specific. Generally, the expression of the three key enzymes of CGS, HMT and MMT has experienced varying degrees of divergence.

Promoter analysis of *CGS*s, *HMT*s and *MMT*s in soybean, *Arabidopsis* and rice

To understand the expression regulation and divergence, the promoters of *CGS*s, *HMT*s and *MMT*s were examined and the *cis*-acting regulatory elements were predicted in silico. A global analysis of regulatory elements in the promoters of *CGS*s, *HMT*s and *MMT*s in soybean, *Arabidopsis* and rice are shown in Table S4 in File S3. In this study, we divided the motifs into 2 groups: Group 1 (related to levels and locations of expression) and Group 2 (related to responses to stresses) (Table S4 in File S3).

First, the numbers of motifs of *GmaCGS* in Group 1 and Group 2 were similar. However, unlike *GmaCGS2*, *GmaCGS1* had two specific motifs, a 5UTR Py-rich stretch and TA-rich region, related to high expression levels, which indicated that the expression levels of *GmaCGS1* might be higher than *GmaCGS2* (Table S4A in File S3). As for *AthCGS*s, the numbers of motifs in the two groups were different. In Group 1, *AthCGS2* (10 motifs) had more motifs than *AthCGS1* (5 motifs), but the opposite was the case in Group 2. Considering their similar spatio-temporal expression patterns, the differences in Group 2 might suggest differences in their responses to different stresses (Table S4A in File S3). In rice, the *OsaCGS*s were divided into two classes, *OsaCGS1-4* in Class 2 and *OsaCGS5* in Class 1. *OsaCGSs* in Class 2 (19 in *OsaCGS1*, 24 in *OsaCGS2*, 23 in *OsaCGS2* and 19 in *OsaCGS2*) had more elements responsive to stresses than *OsaCGS5* (6 elements), suggesting that the *OsaCGS*s in Class 2 might have an important role in responses to stresses. In view of expression levels, *OsaCGS2, OsaCGS4* and *OsaCGS5* had one 5UTR Py-rich stretch, and *OsaCGS3* had one TA-rich region. In our study, the expression intensity of *OsaCGS3* was higher than that of *OsaCGS1* and *OsaCGS5*, which

indicated that the TA-rich region might be necessary to the high expression levels in *OsaCGS*s (Table S4A in File S3). The *MMT*s were relatively conserved, with one or two copies. For example, in rice and *Arabidopsis*, there was only one copy. However, in soybean, there were two copies, and there was a greater number of motifs of *GmaMMT2* than of *GmaMMT1* in Group 1 (18) and Group 2 (16) (5 and 8, respectively) (Table S4B in File S3). Moreover, in *GmaMMT2*, there were 15 enhancers in the promoter, which might be the reason why the expression intensity of *GmaMMT2* was higher than that of *GmaMMT1*.

As for *HMT*s, the *AthHMT*s have been analyzed in our previous study (Table S4C in File S3). In Group 1 and Group 2, the motifs of *AthHMTs* were different. In soybean, there was a greater number of motifs of *GmaHMT4* than of *GmaHMT1–3* in Group 1 (25) (5, 7 and 1, respectively), while in Group 2, there were fewer motifs of *GmaHMT4* (7) than the others (14, 30 and 15, respectively). Similarly, the motifs of *OsaHMT*s were varied in Group 1 and Group 2 (Table S4C in File S3). Therefore, just as their expression patterns were distinct, their promoters were varied.

DISCUSSION

The divergence of *CGS*, *HMT* and *MMT* genes was inconsistent.

CGS, *MMT* and *HMT* genes are vital to the synthesis of methionine in plants (Datko et al., 1974; Bourgis et al., 1999; Ranocha et al., 2001; Lee et al., 2008; Cohen et al., 2017a). In this study, their evolutionary histories were reconstructed. Their phylogenetic relationships were distinct, and the *MMT*s were conserved during evolution, yet the *CGS* and *HMT* gene families in grasses and seed plants diverged in varying degrees. Similarly, the gene structures of the *MMT*s and *HMT*s were conserved, but the structure of *CGS*s diverged in the N-terminals and intron–exon numbers. Further, the divergence in the N-terminals and intron–exon structure in *CGS*s was mainly present in Class 2. Therefore, the evolution of *CGS*, *HMT* and *MMT* gene families was inconsistent.

Although varying degrees of divergence has been detected in *CGS*s, *HMT*s and *MMT*s, they were all under stringent negative selection pressures. The results indicated that the three families did not experience adaptive evolution. However, a partial subfunctionalization might have occurred. Subfunctionalization in evolution often results from changes in gene expression (Gallego-Romero et al., 2012; Wang et al., 2012). In our previous results, the

subfunctionalization of *HMT*s has occurred in their tissue-specific expression, which might be vital to supplying methionine for the development seeds and growth of plants (Zhao et al., 2018). However, *MMT*s were similar in their expression patterns. Nevertheless, the expression of *CGS* genes in dicots was also basically consistent. However, in rice, the expression of *CGS* genes was tissue-specific. In view of the functions of the MTO1 region, which destabilizes the *CGS* mRNA, it seemed that the tissue-specific expression of the *OsaCGS*s might be influenced by the loss of the MTO1 region (Chiba et al., 1999). Furthermore, according to the analysis of promoters in *CGS* genes, the *OsaCGS*s without its MTO1 region were rich in the motifs related to stress responses (Table S4A in File S3). The results indicated that the divergence of *OsaCGS*s might be related to its response to different stresses in rice. However, it is worth noting that the tissue-specific expression did not occur in *AthCGSs*, although *AthCGS2* also lost its MTO1 region. Moreover, there were fewer stress response motifs in *AthCGS2* than in *AthCGS1*. Thus, it seemed that the loss of the MTO1 region in *CGS*s was independent of stress response motifs. In any case, the *CGS*s, *HMT*s and *MMT*s genes have experienced inconsistent divergence in evolution and expression.

Synergistically functional models of *CGSs*, *HMTs* and *MMTs* for methionine synthesis in plants.

Gene duplication supplies the raw materials for evolution. Functional innovations in evolution often result from the expressional changes of duplicated genes (Lynch & Conery, 2000; Zhang, 2003; Gallego-Romero et al., 2012; Wang et al., 2012). During evolution, the *CGS*, *HMT* and *MMT* genes were duplicated in different ways, and their phylogenetic relationship and expressional divergence were inconsistent. Therefore, combining the previous results, we completed evolutionary and functional models for supplying methionine in plants (Fig. 6).

In algeas, only *CGS* and *HMT* genes were found, which suggested that the methionine in algeas was only synthesized by the de novo Asp family pathway. Afterwards, land plants began to appear in the world. The evolutionary history of *MMT*s suggested that *MMT*s might have been the ancestor of land plants. In basal land plants, such as moss, *Selaginella moellendorffii* and so on, the *CGS*s, *HMT*s and *MMT*s were grouped together during evolution, which indicated that the three gene families might not be divergent. Therefore, in basal land plants, the methionine was supplied by the Asp family pathway and the SMM cycle together. However, due to the loss of expression data in this study, their specific functional patterns were unknown.

In seed plants, there have been varying degrees of divergence in the three gene families. First, the *MMT* and *CGS* genes had a high expression in vegetative tissues and a low one in reproductive tissues. The high expression of *MMT*s and *CGS*s in vegetative tissues suggested that the Asp family pathway probably plays a major role in supplying methionine during the early vegetative growth of seed plants, which was consistent with Cohen's results (Cohen et al., 2017b). It is worth mentioning that *HMT* genes have diverged into two clades in seed plants, and their expression has obviously experienced seed- or leaf-specific divergence, which indicated that divergence in the synthesis of Met might have occurred in different organs. Finally, combining the evolution and expression of *CGS*, *HMT* and *MMT* genes together, we inferred their co-functional models in seed plants as follows (Fig. 5). In vegetative tissues like leaves, an amount of Met is synthesized, mainly by CGS through the Asp family pathway. Afterwards, a considerable amount of Met enters into the SMM cycle, in which the Met is converted into SMM by MMT. Then, most of the SMMs are transported into reproductive tissues, such as seeds through phloems (Bourgis et al., 1999; Ranocha et al., 2001; Lee et al., 2008; Cohen et al., 2017b). At the same time, some SMMs are reconverted back into the SMM cycle in leaves. In the seeds, the transported SMMs are reconverted into Met by the seed-specific expressed HMTs, which is the main way in which Met is supplied for seed development, especially in the late stage of seed development (Cohen et al., 2017b). Additionally, the Asp family pathway synthesizes the methionine for seed development (Cohen et al., 2017b). It is worth noting that no matter the tissue, *R*,*S*-SAMs are always recovered by HMTs. Therefore, adequate methionine is supplied for the growth and development of seed plant as a consequence of the synergistic function of CGSs, HMTs and MMTs.

CONCLUSIONS

In this study, the three key enzymes of CGS, MMT and HMT in the biosynthesis of Met were investigated in detail. The evolution of the three gene families has experienced inconsistent divergence; *MMT*s were conserved, while the grass and seed plant levels of *CGS*s and *HMT*s diverged. The gene structures were relatively conserved, except for *CGS* genes. As for gene expression, the change trends of *MMT*s, *CGS*s and *HMT*s were consistent with their evolution. The *MMT*s were relatively conserved, and there was a tissue-specific expression divergence between *CGS*s and *HMT*s in grass and seed plant levels. Finally, it might be the inconsistent

403 divergence between *CGS*s, *HMT*s and *MMT*s that contributed to their functions that supply the
404 methionine, essential to the growth and development of plants.

405 **ADDITIONAL INFORMATION AND DECLARATIONS**

406 **Funding**

407 This work was supported by National Natural Science Foundation of China (Grant No.
408 31600181), Zhejiang Provincial Natural Science Foundation of China (Grant Nos.
409 LQ16C020003 and LY17B020012), and Zhejiang Provincial Major Agriculture Science and
410 Technology Special Sub-project (Grant No. 2016C02050-10-3).

411 **Competing Interests**

412 The authors declare no competing financial interest.

413 **Authors' contributions**

414 X.Y. and M.Z. conceived and designed the analyses. W. W. characterized the sequences and
415 carried out most of the analysis. P. C. and L. W. performed evolutionary and expression analyses.
416 M. Z. and X. Y. analyzed and interpreted the data. F. Y. and Z.W. coordinated the work. M.Z.
417 and X.Y. drafted the manuscript. All authors have read and approved the final manuscript.

418

419 **Supplementary information**

420 **File S1**: The amino acids alignments of *CGS*s and *MMT*s used in the study.

421 **File S2:**

422 **Figure S1.** Phylogenetic tree of the *MMT* (A) and *CGS* (B) gene family in plants. The trees were
423 constructed with maximum likelihood (ML) and neighbor-joining (NJ) methods based on the
424 amino acid (aa). Support values (>50% of ML and NJ) for the two trees are shown on the
425 branches, respectively. The black star indicates the divergence of Class 1 and Class 2. Gene
426 names and identifiers are shown in Table S1 in File S3.

427 **Figure S2.** Sequences of the conserved motifs detected by the MEME analysis on the CGS (A)
428 and MMT (B) homologs across plants. The height of the each letter denotes the probability of the
429 letter at that position, and total height of the stack represents the information content of that
430 position.

431 **Figure S3A.** Tissue-specific expression data of *AthCGS*s from PLEXdb. The green, red and
432 orange boxes represent vegetative tissues such as roots, stems and leaves, flowers and seeds,
433 respectively.

Figure S3B. Tissue-specific expression data of *GmaCGS*s from PLEXdb. The green and blue boxes represent root apical meristem and leaf non-meristem. The pink, orange, red and purple boxes represent bean 2mm, bean 5mm, pod elongation and fully grown, respectively.

Figure S3C. Tissue-specific expression data of *OsaCGS*s from PLEXdb. The pink, yellow, green, orange and red boxes represent emb6D, endo6D, leaf, root and seedling, respectively.

Figure S4A. Tissue-specific expression data of *AthMMT*s from PLEXdb. The green, red, orange boxes represent vegetative tissues such as roots, stems and leaves, flowers, and seeds, respectively.

Figure S4B. Tissue-specific expression data of *GmaMMT*s from PLEXdb. The green and blue boxes represent root apical meristem and leaf non-meristem. The pink, orange, red and purple boxes represent bean 2mm, bean 5mm, pod elongation and fully grown, respectively.

File S3:

Table S1. List of *CGS* (A) and *MMT* (B) sequences used in this study.

Table S2. Distribution of *CGS* (A) and *MMT* (B) genes in different classes.

Table S3. Size variations of exons and introns in *CGS* (A) and *MMT* (B) genes.

Table S4. Analysis of the motifs in the promoters of *CGS*s (A), *MMT*s (B) and *HMT*s (C).

Table S5. Primers used in this study.

REFERENCES

Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37:**202-208.

Bourgis F, Roje S, Nuccio ML, et al. 1999. S-methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. *The Plant Cell* **11:**1485-1498.

Chiba Y, Ishikawa M, Kijima F, Tyson RH, Kim J, Yamamoto A, et al. 1999. Evidence for autoregulation of cystathionine γ-synthase mRNA stability in *Arabidopsis*. *Science* **286:**1371-1374.

Cohen H, Hacham Y, Panizel I, Rogachev I, Aharoni A, Amira R. 2017b. Repression of CYSTATHIONINE γ-SYNTHASE in seeds recruits the S-methylmethionine cycle. *Plant Physiology* **174:**1322-1333.

Cohen H, Israeli H, Matityahu I, Amir R. 2014. Seed-specific expression of a feedback-insensitive form of CYSTATHIONINE-γ-SYNTHASE in *Arabidopsis* stimulates metabolic and transcriptomic responses associated with desiccation stress. *Plant Physiology* **166:**1575-1592.

Cohen H, Salmon A, Tietel Z, HachamY, Amir R. 2017a. The relative contribution of genes operating in the S-methylmethionine cycle to methionine metabolism in *Arabidopsis* seeds. *Plant Cell Reports* **36:**731-743.

Cohen H, Shir OM, Yu Y, Hou W, Sun S, Han T, Amir R. 2016. Genetic background and environmental conditions drive metabolic variation in wild type and transgenic soybean (*Glycine max*) seeds. *Plant Cell and Environment* **39:**1805-1817.

Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA. 2012. PLEXdb: gene expression resources for plants and plant pathogens. *Nucleic Acids Research* **40(D1):**D1194-D1201.

Datko AH, Giovanelli J, Mudd SH. 1974. Homocysteine biosynthesis in green plants: O-Phosphorylhomoserine as the physiological substrate for cystathionine γ-synthase. *The Journal of Biological Chemistry* **249 (4):**1139-1155.

Galili G, Amir R. 2013. Fortifying plants with the essential amino acids lysine and methionine to improve nutritional quality. *Plant Biotechnology Journal* **11:**211-222.

Gallego-Romero I, Ruvinsky I, Gilad Y. 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews* **13:**505-516.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52:**696-704.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:**95-98.

Hanafy MS, Rahman SM, Nakamoto Y, Fujiwara T, Naito S, Wakasa K, Ishimoto M. 2013. Differential response of methionine metabolism in two grain legumes, soybean and azuki bean, expressing a mutated form of *Arabidopsis* cystathionine γ-synthase. *Journal of Plant Physiology* **170:**338-345.

Harada JJ, Barker SJ, Goldberg RB. 1989. Soybean β-conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes. *The Plant Cell* **1:**415-425.

Kim J, Lee M, Chalam R, Martin MN, Leustek T, Boerjan W. 2002. Constitutive overexpression of cystathionine γ-synthase in *Arabidopsis* leads to accumulation of soluble

496 methionine and S-methylmethionine. *Plant Physiology* **128:**95-107.
497 **Kim J, Leustek T. 2000.** Repression of cystathionine γ-synthase in *Arabidopsis thaliana*
498 produces partial methionine auxotrophy and developmental abnormalities. *Plant Science* **151:**9-
499 18.
500 **Kocsis MG, Ranocha P, Gage DA, Simon ES, et al. 2003.** Insertional inactivation of the
501 methionine S-methyltransferase gene eliminates the S-methylmethionine cycle and increases the
502 methylation ratio. *Plant Physiology* **131:**1808-1815.
503 **Lee M, Huang T, Toro-Ramos T, Fraga M, Last RL, Jander G. 2008.** Reduced activity of
504 *Arabidopsis thaliana* HMT2, a methionine biosynthetic enzyme, increases seed methionine
505 content. *The Plant Journal* **54:**310-320.
506 **Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S.**
507 **2002.** PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in*
508 *silico* analysis of promoter sequences. *Nucleic Acids Research* **30(1):**325-327.
509 **Livak KJ, Schmittgen TD. 2001.** Analysis of relative gene expression data using real-time
510 quantitative PCR and the 2-ΔΔCt method. *Methods* **25:**402-408.
511 **Lynch M, Conery JS. 2000.** The evolutionary fate and consequences of duplicate genes.
512 *Science* **290:**1151-1155.
513 **Matityahu I, Godo I, Hacham Y, Amir R. 2013.** Tobacco seeds expressing feedback-
514 insensitive cystathionine γ-synthase exhibit elevated content of methionine and altered primary
515 metabolic profile. *BMC Plant Biology* **13:**206.
516 **Nielsen NC, Dickinson CD, Cho TJ, Thanh VH, Scallon BJ, Fischer RL, Sims TL, Drews**
517 **GN, Goldberg RB. 1989.** Characterization of the glycinin gene family in soybean. *The Plant*
518 *Cell* **1:**313-328.
519 **Ominato K, Akita H, Suzuki A, Kijima F, Yoshino T, Yoshino M, et al. 2002.** Identification
520 of a short highly conserved amino acid sequence as the functional region required for
521 posttranscriptional autoregulation of the cystathionine γ-synthase gene in *Arabidopsis*. *The*
522 *Journal of Biological Chemistry* **277:**36380-36386.
523 **Ranocha P, Bourgis F, Ziemak MJ, Rhodes D, Gage DA, Hanson AD. 2000.** Characterization
524 and functional expression of cDNAs encoding methionine-sensitive and -insensitive
525 homocysteine S-methyltransferases from *Arabidopsis*. *The Journal of Biological Chemistry*
526 **275:**15962-15968.

Ranocha P, McNeil SD, Ziemak MJ, Li C, Tarczynski MC, Hanson AD. 2001. The S-methylmethionine cycle in angiosperms: ubiquity, antiquity and activity. *The Plant Journal* **25:**575-584.

Ranocha P, McNeil SD, Ziemak MJ, Li C, Tarczynski MC, Hanson AD. 2001. The S-methylmethionine cycle in angiosperms: ubiquity, antiquity and activity. *The Plant Journal* **25(5):**575-584.

Roje S. 2006. S-Adenosyl-L-methionine: beyond the universal methyl group donor. *Phytochemistry* **67:**1686-1698.

Sauter M, Moffatt B, Saechao MC, Hell R, Wirtz M. 2013. Methionine salvage and S-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis. *Biochemical Journal* **451:**145-154.

Song S, Hou W, Godo I, Wu C, Yu Y, Matityahu I, Hacham Y, Sun S, Han T, Amir R. 2013. Soybean seeds expressing feedback-insensitive cystathionine γ-synthase exhibit a higher content of methionine. *Journal of Experimental Botany* **64:**1917-1926.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:**2725-2729.

Thompson GA, Datko AH, Mudd SH, Giovanelli J. 1982. Methionine biosynthesis in Lemna: studies on the regulation of cystathionine γ-synthase, O-phosphohomoserine sulfhydrase, and O-acetylserine sulfhydrase. *Plant Physiology* **69:**1077-1083.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25:**4876-4882.

Wang Y, Wang X, Paterson AH. 2012. Genome and gene duplications and gene expression divergence: a view from plants. *Annals New York Academy Sciences* **1256:**1-14.

Yang Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24:**1586-1591.

Zhang J. 2003. Evolution by gene duplication: An update. *Trends in Ecology & Evol*ution **18:**292-298.

Zhao M, Chen P, Wang W, et al. 2018. Molecular evolution and expression divergence of *HMT* gene family in plants. *International Journal of Molecular Sciences* **19 (4)**.

Figure 1(on next page)

The diagrams of the evolutionary relationships of CGS, HMT and MMT gene families.

The diagrams were based on their phylogenetic tree in Figure S1 and Zhao et al. 2018. The red, blue, purple, green and orange triangles represent algae, basal land plants, basal angiosperms and gymnosperms, monocots and dicots, respectively.

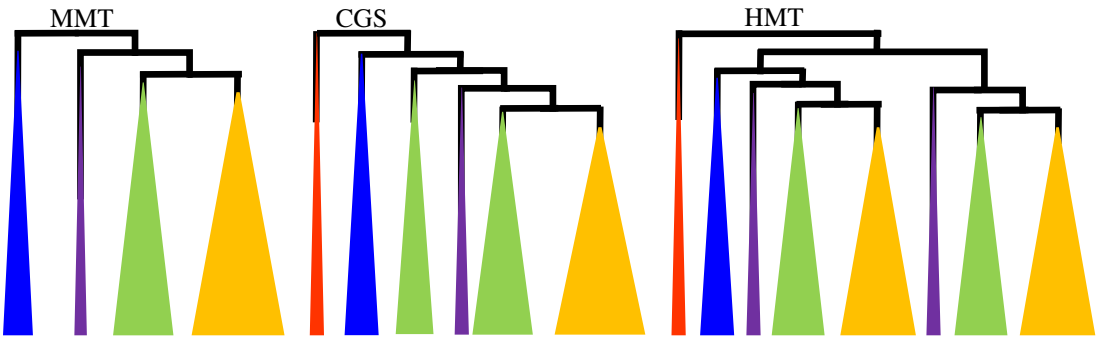

MMT
CGS
HMT

Figure 2(on next page)

Conserved motifs of CGS proteins identified on the MEME analysis across plants.

Each motif is represented by a colored box numbered on the bottom. The amino acid sequences of these motifs are presented in Figure S1 in File2. The black lines represent unique sequences. The scale bar indicates number of amino acids. Names to the left indicate the clades to which the sequences belong to according to Figure S1 in File2.

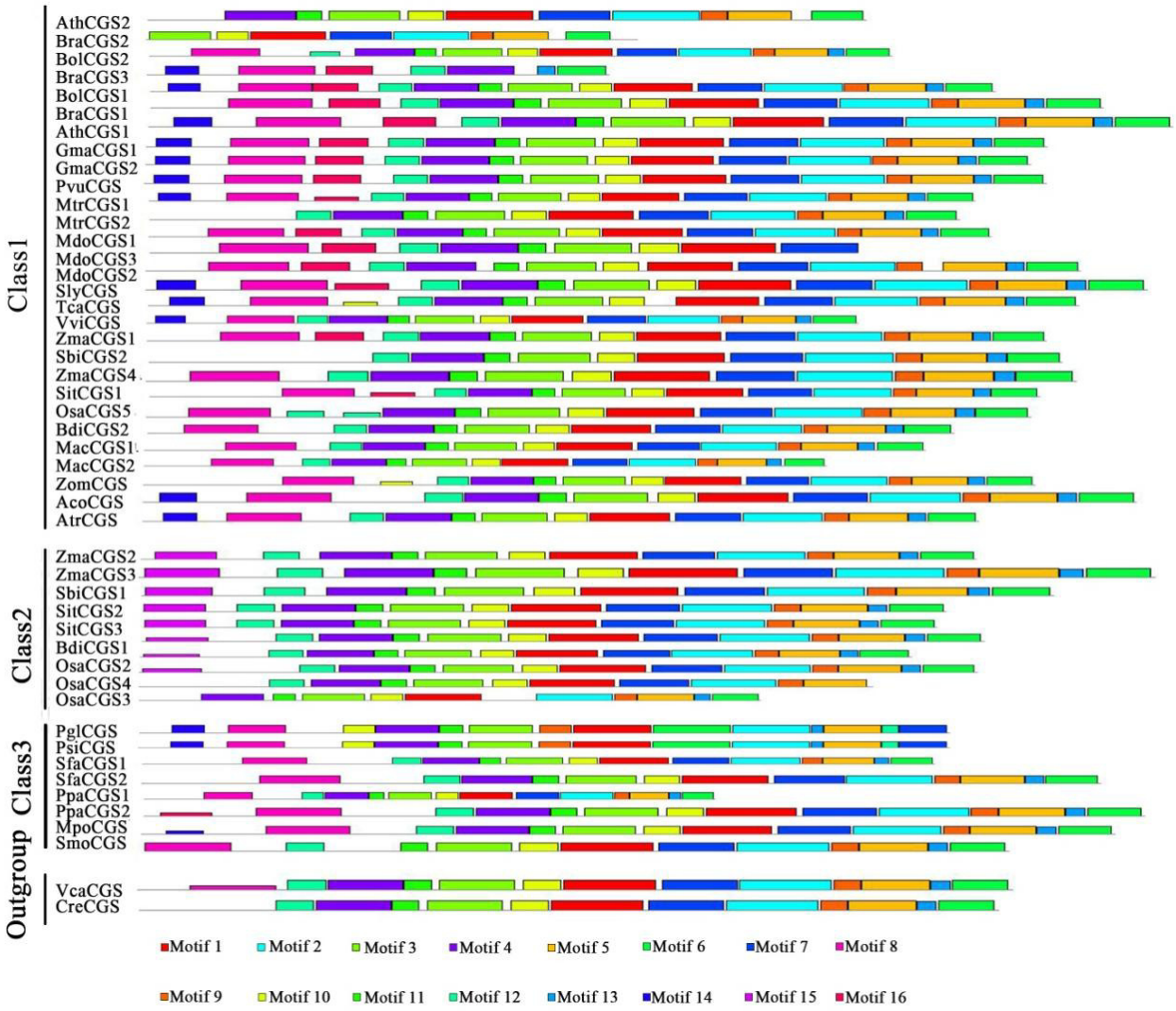

Class1
AthCGS2
BraCGS2
BolCGS2
BraCGS3
BolCGS1
BraCGS1
AthCGS1
GmaCGS1
GmaCGS2
PvuCGS
MtrCGS1
MtrCGS2
MdoCGS1
MdoCGS3
MdoCGS2
SlyCGS
TcaCGS
VviCGS
ZmaCGS1
SbiCGS2
ZmaCGS4
SitCGS1
OsaCGS5
BdiCGS2
MacCGS1
MacCGS2
ZomCGS
AcoCGS
AtrCGS
Class2
ZmaCGS2
ZmaCGS3
SbiCGS1
SitCGS2
SitCGS3
BdiCGS1
OsaCGS2
OsaCGS4
OsaCGS3
Class3
PglCGS
PsiCGS
SfaCGS1
SfaCGS2
PpaCGS1
PpaCGS2
MpoCGS
SmoCGS
Outgroup
VcaCGS
CreCGS
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif 9
Motif 10
Motif 11
Motif 12
Motif 13
Motif 14
Motif 15
Motif 16

Figure 3(on next page)

The sequence composition of the conserved regions in Class 1, Class 2 and Class 3 in CGS family.

The positions of MTO1 region are marked by red lines. The height of the each letter represents the probability of the letter at that position, and total height of the stack represents the information content of that position.

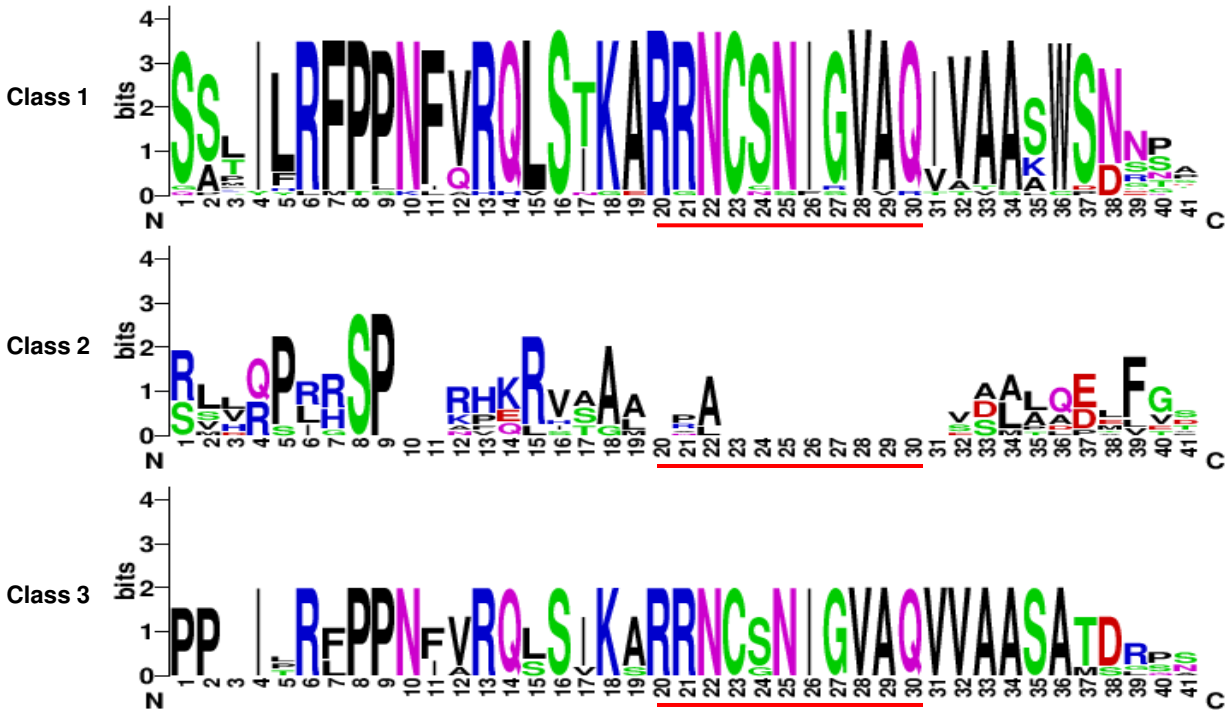

Class 1
Class 2
Class 3
bits
N
C

Figure 4(on next page)

Conserved motifs of MMT proteins identified on the MEME analysis across plants.

Each motif is represented by a colored box numbered on the bottom. The amino acid sequences of these motifs are presented in Figure S1 in File2. The black lines represent unique sequences. The scale bar indicates number of amino acids. Names to the left indicate the clades to which the sequences belong to according to Figure S1 in File2. ˜

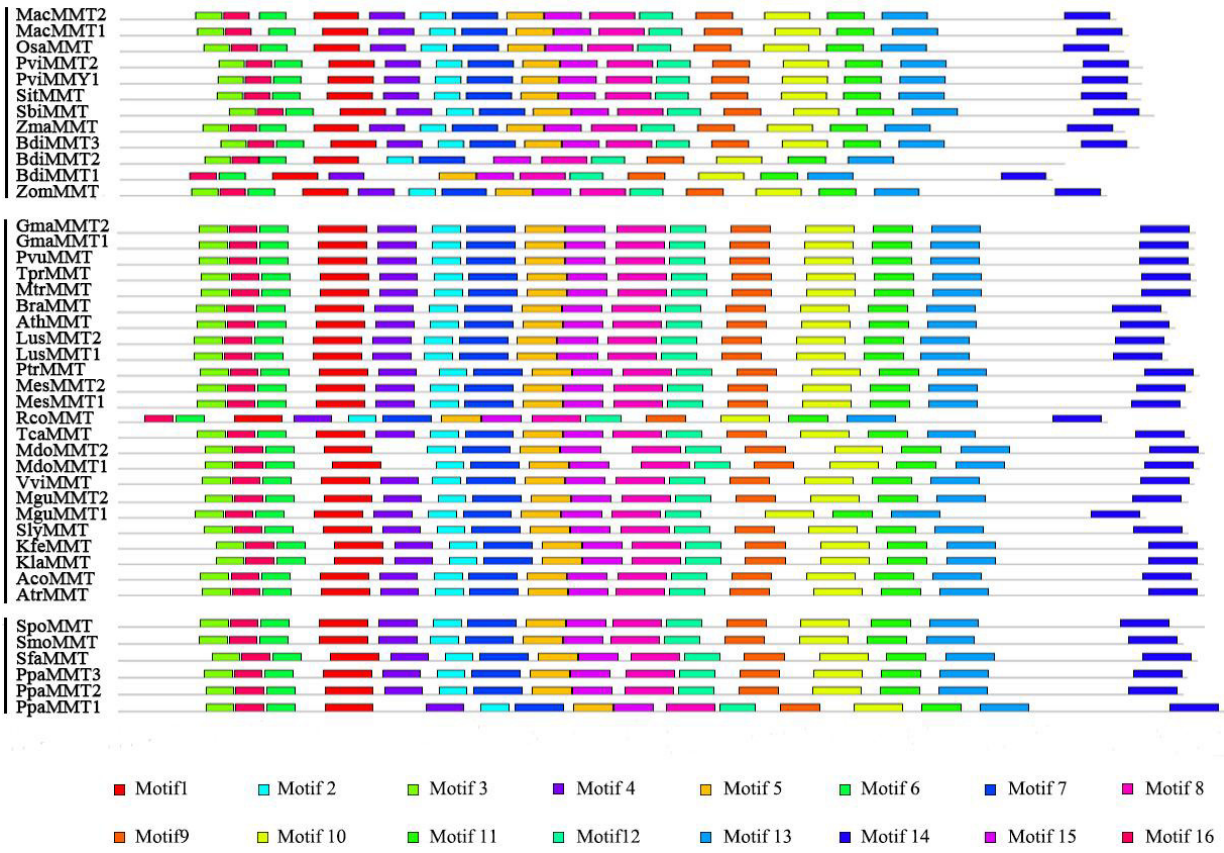

MacMMT2
MacMMT1
OsaMMT
PviMMT2
PviMMY1
SitMMT
SbiMMT
ZmaMMT
BdiMMT3
BdiMMT2
BdiMMT1
ZomMMT
GmaMMT2
GmaMMT1
PvuMMT
TprMMT
MtrMMT
BraMMT
AthMMT
LusMMT2
LusMMT1
PtrMMT
MesMMT2
MesMMT1
RcoMMT
TcaMMT
MdoMMT2
MdoMMT1
VviMMT
MguMMT2
MguMMT1
SlyMMT
KfeMMT
KlaMMT
AcoMMT
AtrMMT
SpoMMT
SmoMMT
SfaMMT
PpaMMT3
PpaMMT2
PpaMMT1
Motif1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
Motif 7
Motif 8
Motif9
Motif 10
Motif 11
Motif12
Motif 13
Motif 14
Motif 15
Motif 16

Figure 5(on next page)

Expression of the GmaCGS, GmaMMT and GmaHMT genes during soybean development. 192.1528681021

A-C The spatio-temporal expression of GmaCGS1 and GmaCGS2 (A), GmaMMT1 and GmaMMT2 (B), and GmaHMT1, GmaHMT2, GmaHMT3 and GmaHMT4(C). The total RNAs were isolated from stems, leaves of 14-day-old seedlings, flowers, and 2-, 4- and 6-week-old pods after fertilization. The ACTIN gene was used as an internal control. The experiments were repeated using three independent biological samples. Error bar: standard deviation. The significance was tested in comparison with the expression of each gene in leaves. The * means significance at a $P < 0.05$ level, and the** represents the significance at a $P < 0.01$ level.

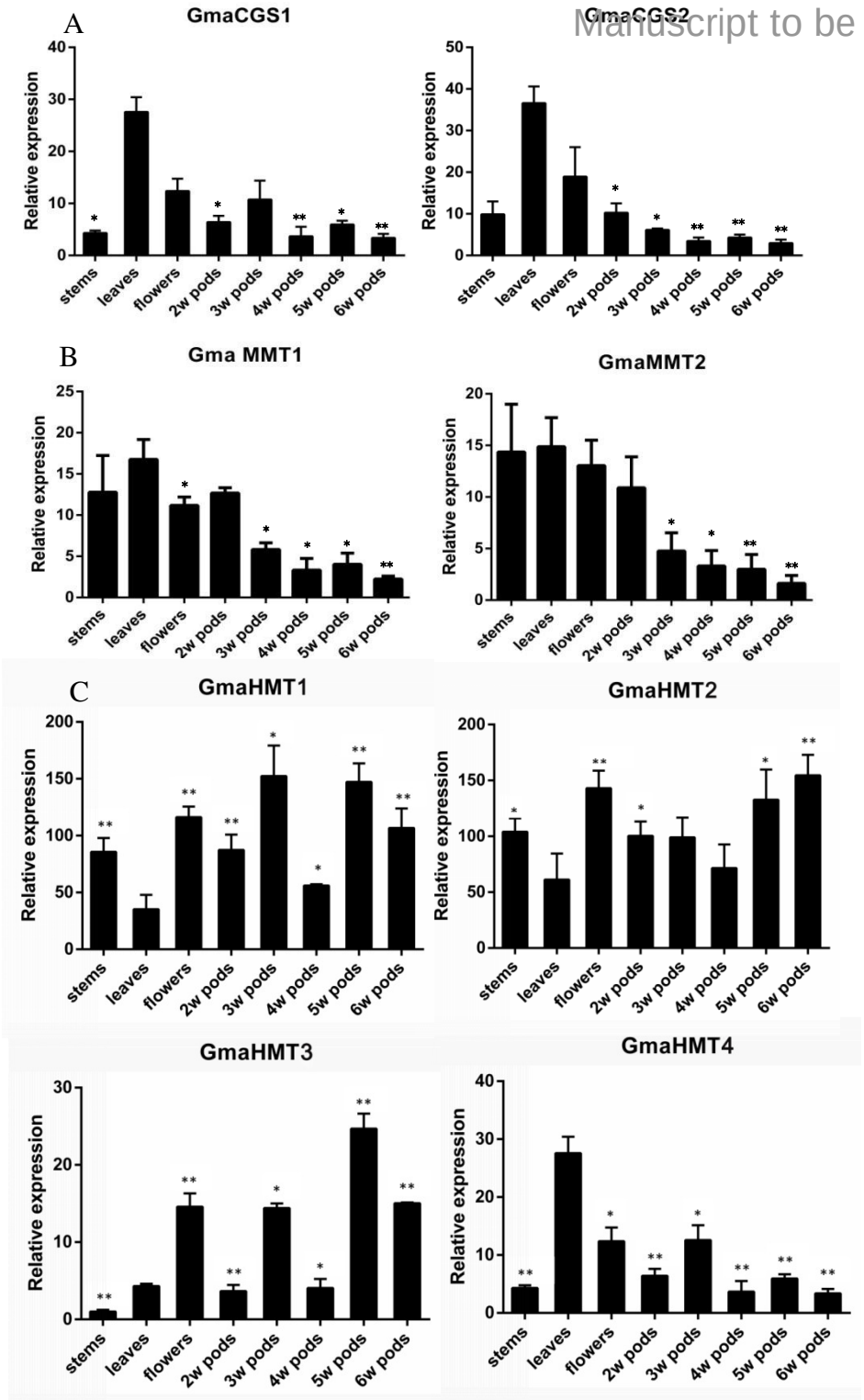
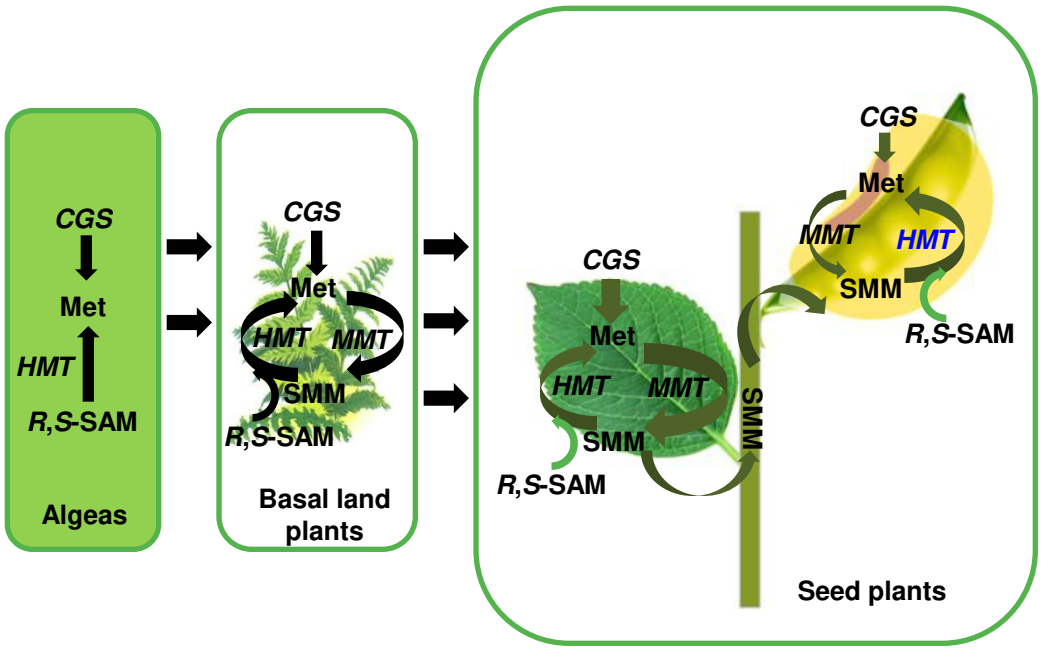

A
GmaCGS1
GmaCGS2
B
Gma MMT1
GmaMMT2
C
GmaHMT1
GmaHMT2
GmaHMT3
GmaHMT4
Relative expression
stems
leaves
flowers
2w pods
3w pods
4w pods
5w pods
6w pods

Figure 6(on next page)

The functional model of CGSs, MMTs and HMTs to synthesize the methionine in plants.

The enzyme of CGS, MMT and HMT in aspartate family pathway and SMM were italics. Black arrows indicate the direction of evolution or flux of reaction. The green arrows indicate methionine flux during the growth and seeds development of seed plants. The thickness of green arrows indicates the strength of flow. The blue HMT were specific expressed in seeds. CGS, cystathionine g-synthase; HMT, homocysteine S-methyltransferases; MMT, met S-methyltransferase; Met, Methionine; R,S-SAM, R,S-adenosylmethionine; SMM, S-methyl-methionine.


CGS
Met
HMT
R,S-SAM
Algeas
CGS
Met
HMT
MMT
SMM
R,S-SAM
Basal land plants
CGS
Met
HMT
MMT
SMM
R,S-SAM
SMM
CGS
Met
MMT
HMT
SMM
R,S-SAM
Seed plants