Plant Gene and Trait 2021, Vol.12, No.4, 1-7
http://genbreedpublisher.com/index.php/pgt

Research Report Open Access

Gene Identification, Evolution and Expression of *MCT* Gene in Mango (*Mangifera indica*)

Gai Jiangtao, Wang Peng, Ye Xiuxu, Chen Yeyuan
Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, 571101, China
Corresponding author email: gjt006616@163.com
Plant Gene and Trait, 2021, Vol.12, No.4 doi: 10.5376/pgt.2021.12.0004
Received: 23 Nov., 2021
Accepted: 28 Nov., 2021
Published: 08 Dec., 2021

Preferred citation for this article:
Gai J.T., Wang P., Ye X.X., and Chen Y.Y., Gene identification, evolution and expression of *MCT* gene in mango (*Mangifera indica*), Plant Gene and Trait, 12(4): 1-7 (doi: 10.5376/pgt.2021.12.0004)

Abstract 2-C-methyl-*D*-erythritol 4-phosphate cytidylyltransferase (MCT) is a key enzyme involved in the 2-*C*-methyl-*D*-erythritol-4-phosphate (MEP) pathway. In order to identify the MCT gene in mango, study the evolutionary route and expression pattern of *MCT* gene. In this study, with arabidopsis *MCT* gene as a reference, 14 MCT genes were searched from 12 plant genomes through the BLAST and HMMER. Then, we analyzed the conserved domains, phylogeny, and the expression of *MCT* gene in mango. The results showed that the MCT protein of mango contains a PF01128.19 domain, and the protein sequences of *MCT* gene is highly conserved. Phylogenetic analysis showed that the evolution of *MCT* genes were consistent with the evolutionary route of its species, the number of copies of the *MCT* genes was tightly controlled from algae to angiosperms. The results of expression analysis showed that the expression of *MCT* gene in the peels and flesh of developing mango were higher than that in the peels and flesh of mature mango in 'guire-82'and 'hongyu'. It is speculated that *MCT* gene play a more important role in the development of mango. These provide scientific basis for further elucidating the function of *MCT* gene in mango.
Keywords Mango (*Mangifera indica*); 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT); 2-C-methyl-D-erythritol 4-phosphate (MEP)

Mango (*Mangifera indica Linnaeus*) is native to India, which is a large evergreen tree in Anacardiaceae. Mango fruit is a popular fruit in human life. Mango fruit is rich in sugar, protein, carotene, vitamin C and other nutrients. Terpenoids are the largest natural products of plants at present. Secondary terpenoids produced by isoprenoid biosynthesis pathway play an extremely important role in the growth and development of plants (Trapp and Croteau, 2001). The MEP pathway locates in the plastids of plant cells and provides precursors for the synthesis of terpenoids. While 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT) is a catalytic enzyme in the third step of the MEP metabolic pathway, and one of the key enzymes in the MEP pathway. MCT enzyme catalyzes Cytidine triphosphate and 2-C-methyl-D-erythritol 4-phosphate (MEP) forms 4-diphosphocytidyl-2-C-methyl-D-erythritol (Rohdich et al., 1999), which plays an important role in the MEP pathway.

The gene of MCT enzyme has been cloned in a variety of plants. *MCT* gene has been cloned in *Lepidium apetalum* (Zhao et al., 2016), *Fritillaria cirrhosa* (Zhang et al., 2018), *Tripterygium wilfordii* (Tong et al., 2015), *Artemisia annua* (Zhang et al., 2016) and *Taxus chinensis* (Jian et al., 2019). Song et al. (2018) transferred the RNAi vector of *Tripterygium wilfordii MCT* gene into *Tripterygium wilfordii* suspension cells, and verified that interference with the expression of *MCT* gene significantly inhibited the accumulation of triptolide and celastrol in *Tripterygium wilfordii*. Jian et al. (2019) used real-time quantitative PCR experiment to prove that the expression level of *MCT* gene in the green bark of *Taxus chinensis* was significantly higher than that in other tissues, and the expression level in leaves was significantly higher than that in stems and roots. Zhang et al. (2016) transformed *Artemisia annua MCT* gene into *Arabidopsis thaliana* and conducted overexpression experiments. The results showed that the contents of chlorophyll a, chlorophyll b and carotenoid were significantly increased in *Arabidopsis thaliana*, indicating that *MCT* gene plays an important role in the biosynthesis of terpenoids.

Considering that *MCT* gene exists widely in plants and plays an important role in terpenoid synthesis, it is inferred that *MCT* gene may also play an important role in mango. However, functional analysis of *MCT* gene in mango has not been reported. In order to identify *MCT* gene in mango and explore the evolutionary route and expression pattern of *MCT* gene, this study used BLAST and HMMER software to search 14 *MCT* genes from 12 plant genomes, and constructed species evolutionary tree and gene evolutionary tree to analyze the evolutionary route of *MCT* gene in plants. The *MCT* gene in mango was identified and its expression was analyzed to provide scientific basis for further study on the function of *MCT* gene in mango.

1 Results and Analysis

1.1 Protein domain identification and sequence alignment of mango *MCT* gene

A conserved domain PF01128.19 was identified in the amino acid sequence of *Arabidopsis thaliana MCT* gene (AT2G02500.1) by searching the Pfam database and named *IspD* in the Pfam database. In the total length of 302 amino acids, this domain was located in the 79-299 amino acid region, covering 221 amino acids, indicating that the latter half of Arabidopsis *MCT* gene was very conserved (Figure 1).

Figure 1 The site of PF01128.19 domain in AT2G02500.1 and Mi04g14720.1 gene

Two genes containing PF01128.19 domain were also identified in the mango genome: *Mi11g18190.1* and *Mi04g14720.1*. The protein sequence of *Mi11g18190.1* gene contained 142 amino acids, and its domain location was 69-128 with a domain length of 60 amino acids, indicating that the domain segment of *Mi11g18190.1* gene was incomplete (Figure 2). The full length of the protein sequence of *Mi04g14720.1* gene was 293 amino acids, and its domain location was 64-284 with a domain length of 221 amino acids, indicating that the protein domain of *Mi04g14720.1* gene was complete and the latter half of the region was very conserved (Figure 1). Since a protein domain with incomplete structure could not produce physiological functions, we removed the incomplete *Mi11g18190.1* gene sequence and named the mango *Mi04g14720.1* gene with complete structure as *MiMCT* gene for subsequent analysis.

```
AT2G02500.1     MAMLQTNLGFITSPTFLCPKLKVKLNSYLWFSYRSQVQKL      40
Mi11g18190.1    ........................................       0
Mi04g14720.1    ..............MGLLLLNLTPPSLSLFPDTHRMLTK      25
Consensus

AT2G02500.1     DFSKRVNRSYKRDALLLSIKCSSSTGFDNSNVVVKEKSVS      80
Mi11g18190.1    ....................................MPELKE       6
Mi04g14720.1    TGHKNQIFGLPRKARIASSRICCSAKLDNSSVVVKEKSVS      65
Consensus

AT2G02500.1     VILLAGGQGKRMKMSMPKQYIP LGQPI  YSFFTFSRMP    120
Mi11g18190.1    IVVVCDPS....YQDIFEGFIS S.LAC  VCVFVLP...     38
Mi04g14720.1    VILLAGGKGKRMAASMPKQYLP LGQPI  YSFYTFSRLP    105
Consensus                              l     al

AT2G02500.1     EVKEIVVV DPFFRDIFEEYEES D   R AI GK         160
Mi11g18190.1    .LEHSTLC ICMLN...QTPKKK C   K SL EN          74
Mi04g14720.1    EVKEIVVV DPSYQDVFADSKEK S   K TL GK         145
Consensus               c              i vdl f  p  erqds

AT2G02500.1     YSGL EI V                  NTE  EKVL GSAV    200
Mi11g18190.1    YTRF AT S                  TSG  A... CWLI    111
Mi04g14720.1    HSGF AI S                  TAG  EKVL GWLI    185
Consensus       v    q  d nselvcihdsarplv   dv    kd

AT2G02500.1          A       VNSDSLVVKT DRKTLWEMQTPQVI    240
Mi11g18190.1         S       VT...IPDDL LAERILNM......    142
Mi04g14720.1         A       ANNESFVVRT DRKTLWEIQTPQVI    225
Consensus       gaavlgvp katike          l

AT2G02500.1     KPELLKKGFELVKSEGLEVTDDVSIVEYLKHPVYVSQGSY    280
Mi11g18190.1    ........................................    142
Mi04g14720.1    KPELLKKGFELVNREGLEVTDDVSIIEHLKHPVYITEGSY    265
Consensus

AT2G02500.1     TNIKVTTPDDLLLAERILSEDS.....                 302
Mi11g18190.1    ........................                    142
Mi04g14720.1    TNIKVTTPDDLLLAERILNMNLEPANK                 292
Consensus
```

Figure 2 The alignment of Arabidopsis MCT protein and mango MCT protein sequences

1.2 Identification of *MCT* gene

In order to further identify the mango *MCT* gene, 11 species whose genome sequences have been published were searched, all of which had high assembly quality genomes and covered the evolution process from lower plants to higher plants (Figure 3). Plants included are as follows: *Chlamydomonas reinhardtii* is the representative of green algae, *Physcomitrella patens* is the representative of bryophyte, *Selaginellae moellendorfii* is the representative of pteridophyte, *Ginkgo biloba* is the representative of gymnosperms, *Amborella trichopoda* is the the representative of the basal species of angiosperms, *Oryza sativa* and *Zea mays* are the representative of monocots, *Arabidopsis thaliana*, *Citrus sinensis*, *Malus domestica* and *Prunus persica* are the representative of dicotyledons.

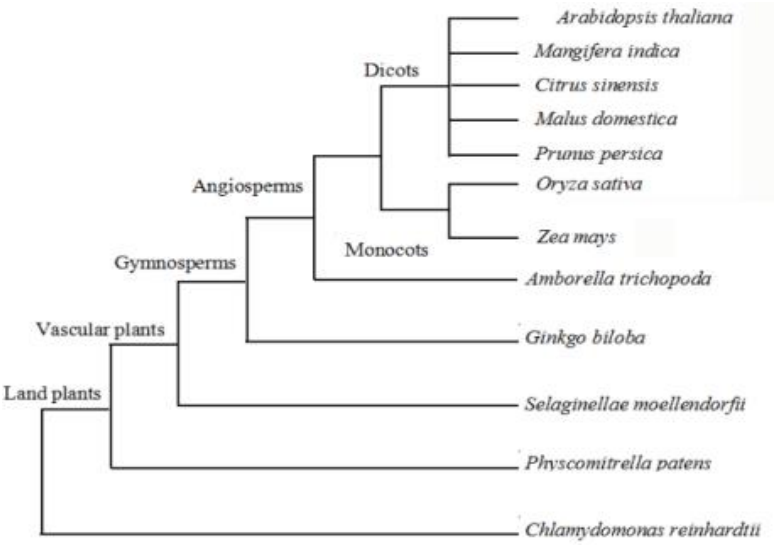

Figure 3 Phylogenetic tree of species used to identify *MCT* gene

1.3 Evolution of the *MCT* gene

In order to analyze the evolutionary relationship of *MCT* gene in plants, BLASTP was used to search the protein model files of each species genome, and the protein sequences containing PF01128.19 domain were retrieved. A total of 14 *MCT* gene protein sequences containing the complete domain were obtained. The phylogenetic tree of *MCT* gene was constructed by maximum likelihood estimate (MLE) (Figure 4). The number of *MCT* gene remained low in all species, and there was only one *MCT* gene in other species except for two protein sequences of *MCT* gene in *Physcomitrella patens* and *Zea mays*, indicating that *MCT* gene was strictly controlled during plant evolution. Moreover, green algae, bryophytes, pteridophyte, gymnosperms, dicotyledons, and monocotyledons clustered in one branch respectively, indicating that the evolutionary relationship of *MCT* gene was consistent with that of the species in which it was located. Moreover, the evolution of mango *MCT* gene Min_Mi04g147201.1 was closely related to the evolution of citrus *MCT* gene Csi_orange1.1g022657m.

1.4 Expression of *MCT* gene in Mango

Two mango varieties ‘guire-82’ and ‘hongyu’ were selected to compare the expression levels of *MiMCT* genes and to express heat maps from transcriptome data of mango *MiMCT* genes. The expression levels of flesh of ripening were very low in mango ‘hongyu’ and ‘guire-82’. In ‘guire-82’, the expression of peels of ripening, peels of developing, flesh of ripening was slightly higher than that of flesh of ripening. But on the whole, the expression level of *MiMCT* gene was lower in both peel and flesh of ‘guire-82’, while the expression level of *MiMCT* gene was the highest in flesh of developing of mango ‘hongyu’, followed by that in peel of developing. Therefore, the expression levels of developing peel and flesh were higher than that of ripening peel and flesh. The expression difference of *MiMCT* gene in ‘guire-82’ and ‘hongyu’ tissues was compared, and the results showed that the expression level of *MiMCT* gene in flesh of ripening of ‘guire-82’ was significantly lower than that in other tissues. However, the expression level of ‘hongyu’ was as follows: the flesh of developing was significantly higher than that of peel, the peel of developing was significantly higher than that of ripening, and the peel of ripening was significantly higher than that flesh of ripening (Figure 5).

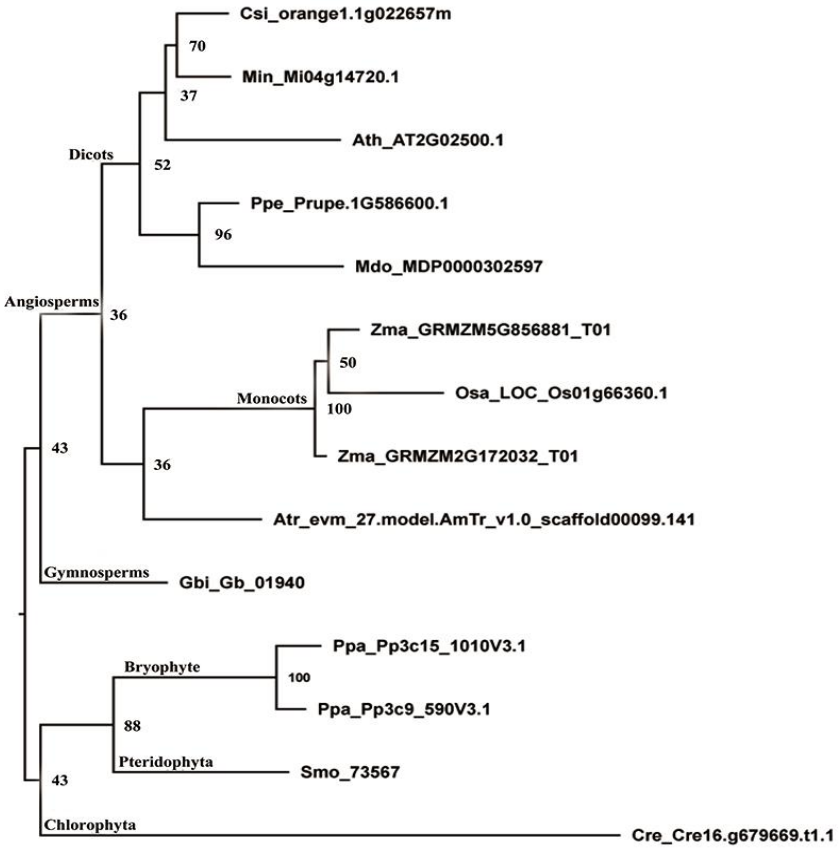

Figure 4 Phylogenetic tree of *MCT* genes

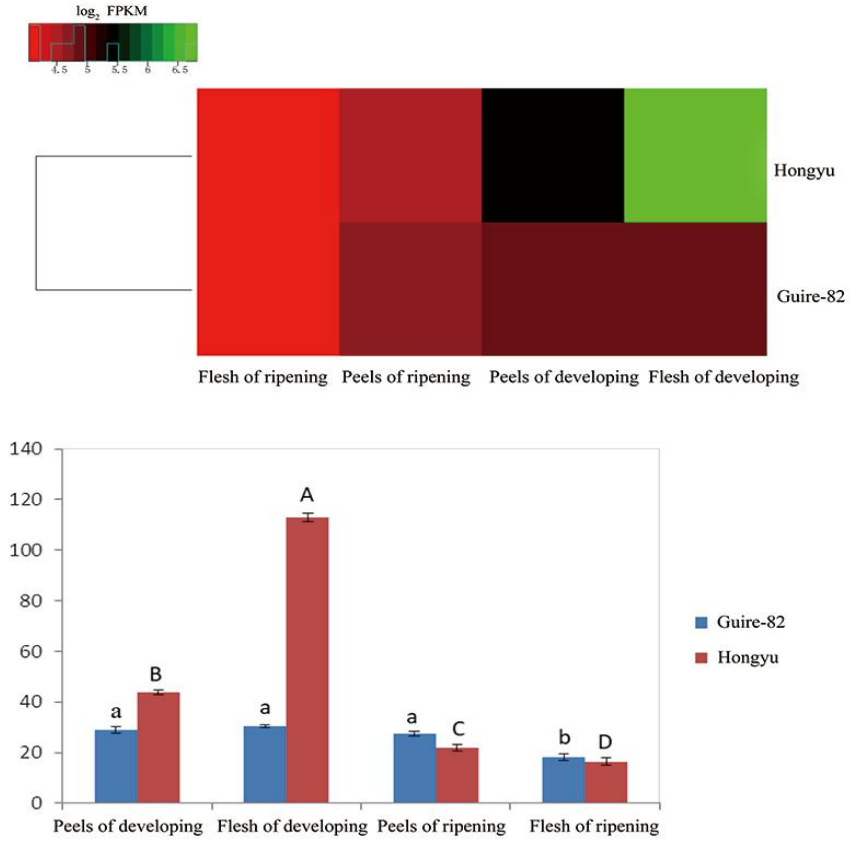

Figure 5 Expression of *MiMCT* gene in peels and flesh of 'hongyu' fruits as well as 'guire-82'

2 Discussion

2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT) is not only a key enzyme involved in the synthesis of isopentenyl pyrophosphate, but also a key enzyme in the MEP metabolic pathway of terpenoid biosynthesis (Rohdich et al., 1999). The results of protein domain analysis in this study showed that the conserved domain of mango MCT protein was consistent with that of other 11 species, including one PF01128.19 domain, belonging to glycosyltransferase (Liu and Mushegian, 2003), indicating that *MCT* gene was highly conserved. It is speculated that mngo *MCT* gene plays an important role in MEP pathway.

In order to systematically study the evolution of *MCT* genes in plants, this study searched *MCT* genes in 12 plant genomes from lower plants to higher plants based on the homology of *MCT* genes in different plants, and obtained 14 *MCT* genes in total. Among them, there was 1 *MCT* gene in each of 10 plant genomes, indicating that its copy number was strictly controlled during plant evolution.

From the phylogenetic tree of *MCT* gene, it can be seen that *MCT* gene exists in all stages of species from green algae to angiosperms, indicating that *MCT* gene evolves with plant evolution. And the *MCT* evolution process is consistent with the evolution of its species, which indicating that *MCT* genes evolves with the evolution of the species. And the number of *MCT* genes is very conservative, two *MCT* genes were only found in *Physcomitrella patens* in bryophytes and *Zea mays*, and only one *MCT* gene was found in the genomes of other species. This may be caused by the choice of adaptability of nature, speculated that *MCT* genes may have an integral role in the process of evolution.

According to the expression levels of *MCT* gene in two mango varieties ‘guire-82’ and ‘hongyu’, it can be seen that the expression levels of *MCT* gene in the developing stage are higher than those in the ripening stage, and the expression levels in the peel and flesh of ‘hongyu’ are higher than those in the peel and flesh of ‘guire-82’. Therefore, it can be inferred that the expression of *MCT* gene is different in different mango varieties and tissues, which may play a more important role in the development of mango and may also be related to chloroplast formation.

At present, there are no studies on the *MCT* gene in mango. In this study, representative species from lower plants to higher plants were selected to conduct identification and evolution research on the *MCT* gene in mango, and the *MCT* gene was preliminarily identified. The expression of *MCT* gene in two mango cultivars was analyzed to provide basis for further study on the function of mango *MCT* gene.

3 Materials and Methods

3.1 MCT protein sequence acquisition

In this study, genome data of *Chlamydomonas reinhardtii*, *Physcomitrella patens*, *Selaginella moellendorfii*, *Amborella trichopoda*, *Oryza sativa*, *Zea mays*, *Musa acuminata*, *Arabidopsis thaliana*, *Citrus sinensis*, *Malus domestica* and *Prunus persica* were downloaded from the Phytozome database (Goodstein et al., 2012). Ginkgo genome data download from http://gigadb.org/dataset/100209 (Guan et al., 2016). Mango genome was sequencing data in this study. BLAST was used to search the protein model database of each species, and the searched E value was set to 1e-10. Combined with HMMER software analysis, *MCT* gene protein sequence information of each species was obtained (Finn et al., 2011).

3.2 Identification and sequence modification of MCT protein conserved domain

The description line contents of protein sequences in each species were removed, only the sequence ID and sequence were retained, and the *MCT* gene protein sequences of each species were input into Pfam database. The default parameters were used as the standard to identify the domain and obtain the start and end site information (El-Gebali et al., 2018), and the sequences containing complete domain were retained. Sequence alignment was carried out to remove repeated sequences. The sequence ID of each species was labeled at the beginning as a combination of genus name and species name (Wang et al., 2013). For example, the *MCT* gene of mango is labeled Min_Mi04g147201.1, and the *MCT* gene of Arabidopsis is labeled Ath_AT2G02500.

3.3 MCT protein sequence alignment and phylogenetic tree construction

The determined protein sequences were compared by Probcons (Do et al., 2005), and the phylogenetic tree was constructed by PhyML. Bootstrap analysis of 1 000 generations was performed (Guindon et al., 2009). FigTree v1.4.4 was used to display the phylogenetic tree.

3.4 Expression analysis of *MCT* gene

In this study, mango varieties 'guire-82' and 'hongyu' were planted in the Mango Germplasm Resource Nursery of Danzhou, Hainan Province, and fresh peel and flesh samples of the two varieties at the development and maturity stages were collected respectively. The samples were three biological iterations, which were quick-frozen in liquid nitrogen and stored at -80℃ for reserve. Transcriptome sequencing data were extracted to compare the expression of mango *MCT* gene.

Authors' contributions

GJT is the experimental designer and executor of this study. GJT completed data analysis and wrote the first draft of the paper. YXX, CYY, and WP participated in experimental design and analysis of experimental results. GJT is the architect and principal of the project, directing experimental design, data analysis, paper writing and revision. All authors read and approved the final manuscript.

Acknowledgments

This study was supported by the Hainan Provincial Natural Science Foundation of China (319QN310) and Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (1630032019033).

Reference

Do C.B., Mahabhashyam M.S., Brudno M., and Batzoglou S., 2005, ProbCons: probabilistic consistency-based multiple sequence alignment, Genome Res., 15(2): 330-340
https://doi.org/10.1101/gr.2821705
PMid:15687296 PMCid:PMC546535

El-Gebali S., Mistry J., Bateman A., Eddy S.R., Luciani A., Potter S.C., Qureshi M., Richardson L.J., Salazar G.A., Smart A., Sonnhammer E.L.L., Hirsh L., Paladin L., Piovesan D., Tosatto S.C.E., and Finn R.D., 2018, The Pfam protein families database in 2019, Nucleic Acids Res., 47(D): 427-432
https://doi.org/10.1093/nar/gky995
PMid:30357350 PMCid:PMC6324024

Finn R.D., Clements J., and Eddy S.R., 2011, HMMER web server: interactive sequence similarity searching. Nucleic Acids Res., 39(2): W29-W37
https://doi.org/10.1093/nar/gkr367
PMid:21593126 PMCid:PMC3125773

Goodstein D.M., Shu S.Q., Howson R., Neupane R., Hayes R.D., Fazo J., Mitros T., Dirks W., Hellsten U., Putnam M., and Rokhsar D.S., 2012, Phytozome: a comparative platform for green plant genomics, Nucleic Acids Res., 40(D): 1178-1186
https://doi.org/10.1093/nar/gkr944
PMid:22110026 PMCid:PMC3245001

Guan R., Zhao Y.P., Zhang H., Fan G.Y., Liu X., Zhou W.B., Shi C.C., Wang J.H., Liu W.Q, Liang X.M., Fu Y.Y, Ma K.L., Zhao L.J., Zhang F.M., Lu Z.H., Lee S.M.Y., Xu X., Wang J., Yang H.M., Fu C.X., Ge S., and Chen W.B., 2016, Draft genome of the living fossil *Ginkgo biloba*, Gigascience, 5(1): 5-49
https://doi.org/10.1186/s13742-016-0154-1
PMid:27871309 PMCid:PMC5118899

Guindon S., Delsuc F., Dufayard J.F., and Gascuel O., 2009, Estimating maximum likelihood phylogenies with PhyML, Methods in Molecular Biology, 537: 113-137
https://doi.org/10.1007/978-1-59745-251-9_6
PMid:19378142

Jian D.Q., Zeng L.J., Yang Y.F., and Yang C.X., 2019, Molecular cloning, functional analysis and tissue expression specificity detection of *MCT* from *Taxus chinensis*, Journal of Southern Agriculture, 50(7): 1408-1416

Liu J., and Mushegian A., 2003, Three monophyletic superfamilies account for the majority of the known glycosyltransferases, Protein Sci., 12(7): 1418-1431
https://doi.org/10.1110/ps.0302103
PMid:12824488 PMCid:PMC2323934

Rohdich F., Wungsintaweekul J., Fellermeier M., Sagner S., Herz S., Kis K., Eisenreich W., Bacher A., and Zenk M.H., 1999, Cytidine 5'-triphosphate-dependent biosynthesis of isoprenoids: Ygbp protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol, Proc. Natl. Acad. Sci. USA, 96(21): 11758-11763
https://doi.org/10.1073/pnas.96.21.11758
PMid:10518523 PMCid:PMC18359

Song Y.D., Zhao Y.J., Chen S., Hu T.Y., Zhang R., Wang J.D., Lu Y., Wang X.J., Gao W., and Huang L.Q., 2018, Effect on biosynthesis of terpenoid active components through RNA interference with *MCT* gene in *Tripterygium wilfordii*,Acta Pharmaceutica Sinica, 53(8): 1209-1214

Tong Y.R., Su P., Zhang M., Zhao Y.J., Wang X.J., Gao W., and Huang L.Q., 2015, Cloning and expression analysis of 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase gene in *Tripterygium wilfordi*,China Journal of Chinese Materia Medica, 40(22): 4378-4383

Trapp S.C., and Croteau R.B., 2001, Genomic organization of plant terpene synthases and molecular evolutionary implications, Genetics, 158(2): 811-832
https://doi.org/10.1093/genetics/158.2.811
PMid:11404343 PMCid:PMC1461696

Wang P., Wang Z.N., Dou Y.C., Zhang X.X., Wang M.Y., and Tian X.M, 2013, Genome-wide identification and analysis of membrane-bound O-acyltransferase (MBOAT) gene family in plants, Planta, 238(5): 907-922
https://doi.org/10.1007/s00425-013-1939-4
PMid:23928653

Zhang M., Xiang L.E., Wang H., Lan X.Z, Chen M., and Liao Z.H., 2016, Molecular cloning and characterization of the 2-C-methyl-*D*-erythritol 4-phosphate cytidylyltransferase gene from *Artemisia annua* L.,Acta Pharmaceutica Sinica, 51(8): 1334-1339

Zhang T.T., Li R., Chen X.Y., and Zhao Q., 2018, Cloning and expression analysis of *MCT* gene in *Fritillaria cirrhosa* D. Don, Molecular Plant Breeding, 16(22), 7275-7280

Zhao L., Ma L.G., Li X.Y., Feng W.S., and Zheng X.K., 2016, Cloning, sequence analysis and prokaryotic expression of *LaMCT* gene from *Lepidium apetalum*,Guihaia), 36(10): 1225-1231