# Molecular Cloning, Characterization and Functional Analysis of a 2C-methyl-D-erythritol 2, 4-cyclodiphosphate Synthase Gene from *Ginkgo biloba*

**Shi Gao[1], Juan Lin[1], Xuefen Liu[1], Zhongxiang Deng[1], Yingjun Li[1], Xiaofen Sun[1] and Kexuan Tang[1,2,*]**

[1]State Key Laboratory of Genetic Engineering, School of Life Sciences, Morgan-Tan International Center for Life Sciences, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Fudan University, Shanghai 200433, China,
[2]Plant Biotechnology Research Center, School of Agriculture and Biology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiao Tong University, Shanghai 200030, China

2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (MECPS, EC: 4.6.1.12) is the fifth enzyme of the non-mevalonate terpenoid pathway for isopentenyl diphosphate biosynthesis and is involved in the methylerythritol phosphate (MEP) pathway for ginkgolide biosynthesis. The full-length *mecps* cDNA sequence (designated as *Gbmecps*) was cloned and characterized for the first time from gymnosperm plant species, *Ginkgo biloba*, using RACE (rapid amplification of cDNA ends) technique. The full-length cDNA of *Gbmecps* was 874 bp containing a 720 bp open reading frame (ORF) encoding a peptide of 239 amino acids with a calculated molecular mass of 26.03 kDa and an isoelectric point of 8.83. Comparative and bioinformatic analyses revealed that GbMECPS showed extensive homology with MECPSs from other species and contained conserved residues owned by the MECPS protein family. Phylogenetic analysis indicated that GbMECPS was more ancient than other plant MECPSs. Tissue expression pattern analysis indicated that GbMECPS expressed the highest in roots, followed by in leaves, and the lowest in seeds. The color complementation assay indicated that GbMECPS could accelerate the accumulation of β-carotene. The cloning, characterization and functional analysis of GbMECPS will be helpful to understand more about the role of MECPS involved in the ginkgolides biosynthesis at the molecular level.

**Keywords:** Color complementation assay, 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (MECPS), *Ginkgo biloba*, Ginkgolides biosynthesis, RACE

---

---

*To whom correspondence should be addressed.
Tel: 86-21-65642772; Fax: 86-21-65643552
E-mail: kxtang1@yahoo.com; kxtang1@163.com

## Introduction

*Ginkgo biloba* is a deciduous and dioecious gymnosperm species originated in China and is the sole survivor of the ancient family of *Ginkgoaceae* (Carrier *et al.*, 1998). The active ingredients of *G. biloba* extracts are flavonoid glycosides and terpene lactones, containing ginkgolides A, B, C, J, M and bilobalides (Jacob *et al.*, 2000). The ginkgolides, only found in *G. biloba*, are the most potent antagonists of platelet-activating factor receptors which have been widely used as platelet-activating factor receptors to cure cerebrovascular and cardiovascular diseases (Hosford *et al.*, 1990). Ginkgolides have attracted great commercial interests as pharmaceuticals or nutraceuticals (Sacchettini and Poulter, 1997). However the contents of ginkgolides are very low in the native ginkgo plant materials (Van Beek *et al.*, 1991) and the ginkgo cell cultures hardly produce ginkgolides (Laurain *et al.*, 1997). Therefore, to map the ginkgolides biosynthetic pathway at the level of molecular genetics may lead ginkgolides production to a bright path (Schwarz and Arigoni, 1999).

Terpenoids such as ginkgolides are biosynthesized from a universal $C_5$ building block: isopentenyl diphosphate (IPP) (Ying-Shan Han *et al.*, 2003). IPP can be derived from two pathways: one is classical cytosolic mevalonate pathway (Bloch, 1992) and the other is plastidial methylerythritol 4-phosphate (MEP) pathway, which is mevalonate-independent (Zeidler *et al.*, 1997; Eisenreich *et al.*, 2001).

Plants use both MEP pathway and MVA pathway for isoprenoid biosynthesis, although they are localized in different compartments (Lange *et al.*, 2000; Rodriguez-Concepcion *et al.*, 2002). Isoprenoids are specially abundant and diverse in plants (Chappell *et al.*, 1995). Besides their primary (essential) roles in membrane structure, redox reactions, light harvesting and photoprotection, and regulation of growth and development, isoprenoids also participate as secondary metabolites in a variety of functions in plants.

Furthermore, the MEP pathway was confirmed essential for plastidial isoprenoid biosynthesis in plants (Rodriguez-Concepion *et al.*, 2004). After the identification of the enzymes and intermediates of the MEP pathway, it was observed that the pathway not only contributes to isoprenoid biosynthesis, but also produces the precursors for the synthesis of cytokinins such as *trans*-zeatin and isopentenyladenine derivatives. Since the MEP pathway is absent from archaebacteria, fungi and animals and the MEP pathway enzymes are highly conserved but show no homology to mammalian proteins, the use of specific inhibitors of MEP pathway should result in novel antimicrobial drugs with broad spectrum activity and little toxicity to human. In fact, after its discovery and elucidation, the MEP pathway has become one of the most promising targets for the identification of new herbicides and antibiotics in recent years (Rodriguez-Concepion *et al.*, 2004).

The MEP pathway starts with the formation of 1-deoxy-D-xylulose 5-phosphate (DXP) from D-glyceraldehyde 3-phosphate and pyruvate by the catalytic action of a 1-Deoxy-D-xylulose 5-phosphate synthase (Sprenger *et al.*, 1997; Bouvier *et al.*, 1998). Then, DXP is converted into IPP by a series of enzymes. The fifth-step enzyme in the MEP pathway is 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (MECPS, EC: 4.6.1.12) which is encoded by the *mecps* gene (Bick *et al.*, 2003). The participation of MECPS in the control of isoprenoid accumulation in plants is sustained by experimental results (Veau *et al.*, 2000). It has been demonstrated that there is a correlation between MECPS gene expression and terpenoid MIA accumulation (Veau *et al.*, 2000). Ginkgolides, as highly functional diterpenes, are biosynthesized via the newly

discovered plastidial MEP pathway (Eisenreich *et al.*, 2001). There have been a few reports on the cloning of genes involved in biosynthesis of ginkgolides, such as *Gbdxs* (GenBank accession No.: AY971576) and *Gbdxr* (GenBank accession No.: AY971576). However, until now there have been no reports on the cloning of *mecps* genes from gymnosperm plants including *G. biloba*. In the present study, we report for the first time the molecular cloning and characterization of the *mecps* gene from the living fossil plant *G. biloba* by RACE technique and validate its biological function in *Escherichia coli*, which will enable us to map and regulate an important step involved in ginkgolide biosynthetic pathway at the level of molecular genetics in the future.

## Materials and Methods

**Plant materials.** Young leaves and roots were collected from *G. biloba* plant grown in the greenhouse at Fudan University, Shanghai, China, and used as the starting material for total RNA isolation. Seeds of *G. biloba* stored at 4°C were also used to isolation total RNA.

**Cloning of *Gbmecps* core cDNA fragment.** Total RNA was isolated by CTAB method and lithium chloride precipitation (Liao *et al.*, 2004). The core cDNA fragment of *Gbmecps* gene was cloned according to the protocol of One Step RNA PCR Kit (AMV) (TaKaRa). The forward primer MFP1 and reverse primer MRP1 (Table 1) used for the cloning of *Gbmecps* core cDNA fragment were designed and synthesized according to the conserved regions of the *mecps* gene sequences of *Arabidopsis thaliana*, *Oryza sativa* and *Catharanthus roseus* deposited in GenBank. PCR

**Table 1.** Primers used in the cloning and analysis of *Gbmecps*

| Primer | Primer sequence (5'→3') |
| --- | --- |
| MFP1 | 5'-ATAGAGG(A/T/C)TGCGAGGCTCA(C/T)TC(T/C)GATG-3' |
| MRP1 | 5'-(T/A)GT(A/G)TG(T/A)GC(G/C/T)GC(A/T)A(T/C)(A/G)CTTC(G/T)(G/A)TT(C/T)TC-3' |
| MFP3 | 5'-ATGAGGCAGGGTATGAGCTTGGGAACT-3' |
| NEWMRP1 | 5'-AAGTTCCCAAGCTCATACCCTGCCTCA-3' |
| NEWMRP2 | 5'-GCCCAATGTCAGGAAGACCCAATGCTC-3' |
| M-BEGIN | 5'-GGTATTCAAGCATAGCAATTTCAGCTG-3' |
| M-END | 5'-TATGACTGGGAAACAGACACTATGGTA-3' |
| MORF-FP | 5'-ATGGCTGCCGCATTCTCTGCACTTGTG-3' |
| MORF-RP | 5'-TCACTTCTTCATCAAAAGTACAATTGT-3' |
| MORF-FPB | 5'-CGAGATCTATGGCTGCCGCATTCTCTG-3' |
| MORF-RPN | 5'-ATGCGGCCGCTCACTTCTTCATCAAAAG-3' |
| Universal Primers Mixed (UPM) | 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3' |
| | 5'-CTAATACGACTCACTATAGGGC-3' |
| Nested Universal Primer A (NUP) | 5'-AAGCAGTGGTATCAACGCAGAGT-3' |
| AP | 5'-GGCCACGCGTCGACTAGTAC (T) 17-3' |
| Universal Amplification Primer (UAP) | 5'-CUACUACUACUAGGCCACGCGTCGACTAGTAC-3' |
| Abridged Universal Amplification Primer (AUAP) | 5'-GGCCACGCGTCGACTAGTAC-3' |
| 18SF | 5'-ATGATAACTCGACGGATCGC-3' |
| 18SR | 5'-CTTGGATGTGGTAGCCGTTT-3' |

was carried out in a total volume of 50 μl containing 5 μl 10 × buffer, 10 μl MgCl$_2$, 5 μl 2.5 mmol dNTP Mix, 1 μl RNase inhibitor, 1 μl total RNA, 2 μl 10 pmol primer MFP1, 4 μl 10 pmol primer MRP1, 1 μl AMV Rtase XL, and 1 μl *AMV-Optimized Taq* polymerase. The amplification was performed in a GeneAmp PCR System 2400 for 30 min at 50°C, 2 min at 94°C, and 30 cycles with 30 sec at 94°C, 30 sec at 53°C, 1.5 min at 72°C. After the final cycle, the amplification was extended for 10 min at 72°C.

**5' RACE of *Gbmecps*.** According to the protocol of the SMART™ RACE cDNA Amplification Kit (Clontech), about 100 ng of total RNA was reversely transcribed with primer 5'-CDS primer coupled with (dC) tailing and SMART II A oligo, which annealed to the tail of the RNA and served as an extended template for PowerScript RT. The reverse primer used for the cloning of partial coding sequence of *Gbmecps* was designed and synthesized according to the core cDNA fragment of *Gbmecps* gene obtained earlier. The first round PCR was performed with NEWRP1 as the reverse primer and Universal Primer A Mix (UPM) as the forward primer (Table 1). PCR was carried out in a total volume of 50 μl containing 2.5 μl cDNA, 1 μl 10 pmol primer NEWRP1, 5 μl 10 pmol UPM, 41.5 μl Master Mix (34.5 μl PCR-Grade Water, 5 μl 10 × Advantage 2 PCR buffer, 1 μl 10 mmol dNTP Mix, 1 μl 50 × Advantage 2 Polymerase Mix) under the following condition: the template was firstly denatured at 94°C for 3 min and then subjected to 30 cycles of amplification (94°C for 30 sec, 68°C for 30 sec, 72°C for 3 min) followed by 10 min at 72°C. Subsequently, nested PCR was performed under normal PCR reaction conditions using NEWRP2 as reverse primer and the Nested Universal Primer A (NUP) as forward primer (Table 1).

**3' RACE of *Gbmecps*.** According to the protocol of the 3' RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL), an aliquot of isolated 100 ng RNA was reversely transcribed using the cDNA synthesis primer AP (Table 1), provided within the kit (GIBCO BRL). The gene-specific primer MFP1 was used as the forward primer and Universal Amplification Primer (UAP) (Table 1) was used as the reverse primer to amplify the 3' end cDNA. PCR was performed in a total volume of 50 μl containing 2 μl cDNA, 2 μl 1 pmol primer MFP1, 2 μl 10 pmol primer UAP, 1 μl 10 mmol dNTPs, 5 μl 10 × cDNA reaction buffer and 8 U *Taq* polymerase. The amplification was performed in a GeneAmp PCR System 2400 for 3 min at 94°C followed by 35 cycles with 30 sec at 94°C, 1 min at 48°C and 1 min at 72°C. After the final cycle, the amplification was extended for 10 min at 72°C. The gene-specific primer MFP3 was used as the forward primer and Abridged Universal Amplification Primer (AUAP) (Table 1) was used as the reverse primer in the nested PCR, which was carried out in a total volume of 50 μl containing 1 μl of the first round 3' RACE-PCR product under the following condition: the template was denatured at 94°C for 3 min and then subjected to 35 cycles of amplification (94°C for 30 sec, 57°C for 1 min, 72°C for 1 min) followed by 10 min at 72°C.

**Generation of the full-length cDNA of *Gbmecps*.** By comparing and aligning the sequences of the core fragment, the 5' RACE and 3' RACE products, the full-length cDNA sequence of *Gbmecps* was deduced and subsequently amplified with primers M-BEGIN and M-END (Table 1), using One Step PCR Kit (TaKaRa). RT-PCR

was carried out in a total volume of 50 μl containing 1 μl total RNA, 1 μl 10 pmol of M-BEGIN and M-END respectively, 1 μl 10 mmol dNTPs, 5 μl 10 × One Step RNA PCR buffer, 1 μl RNase Inhibitor, 1 μl AMV Reverse Transcriptase XL and 1 μl AMV-Optimized *Taq* polymerase under the following conditions: 50°C for 30 min and 94°C for 2 min followed by 30 cycles of amplification (94°C for 30 sec, 54°C for 30 sec, 72°C for 1 min). After the final cycle, the amplification was extended for 10 min at 72°C.

All RT-PCR products were purified and cloned into pMD 18-T Vector (TaKaRa), and sequenced.

**Comparative and bioinformatic analyses.** Comparative and bioinformatic analyses of *Gbmecps* were carried out online at the websites (http://www.ncbi.nlm.nih.gov and http://cn.expasy.org). The nucleotide sequence, deduced amino acid sequence and ORF (open reading frame) encoded by *Gbmecps* were analyzed and the sequence comparison was conducted through database search using BLAST program (http://www.ncbi.nlm.nih.gov). The chloroplast transit peptides (cTP) of GbMECPS was predicted at the website (http://www.cbs.dtu.dk/services/ChloroP/). The homology-based 3-D structural modeling of GbMECPS was accomplished by Swiss-Modeling (Schwede *et al.*, 2003). WebLab ViewerLite was used for 3-D structure displaying (homology-based modeling by Swiss-Model). The phylogenetic analysis of GbMECPS and MECPSs from other species was aligned with CLUSTAL W (1.82) using default parameters. Phylogenetic tree was constructed using MEGA version 2.1 (Kumar *et al.*, 2001) from CLUSTAL W alignments. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the tree.

**Southern blot analysis.** Genomic DNA (50 μg/sample) was isolated from leaves of *G. biloba* using a CTAB method, digested overnight respectively with *Dra*I and *Hin*dIII (Biolabs). The products were electrophoresed on 0.8% agarose gel and transferred onto nylon membrane (Amersham) by capillary blotting. Biotin labeling of the *Gbmecps* coding region as the probe was carried out by the random priming method using the Amersham Multiprime Labeling system. The blot was hybridized, blocked, incubated with antibody and washed, and signal generation and detection were performed according to the gene images CDP-star detection protocols (Amersham Pharmacia Biotech).

**Semi-quantitative RT-PCR analysis of *Gbmecps* expression in different tissues.** To investigate the *Gbmecps* expression pattern in different parts of *G. biloba*, total RNA was extracted from roots, leaves and seeds respectively, followed by incubation with RNase-free DNase I at 37°C for 30 min according to the manufacturer's instruction (Takara). An aliquot of 0.25 μg total RNA was used as the template in semi-quantitative RT-PCR using One Step RNA PCR Kit (TaKaRa), with MORF-FP and MORF-RP as primers (Table 1). The template was reversely transcribed at 50°C for 30 min and denatured at 94°C for 2 min, followed by 20 cycles of amplification (94°C for 30 sec, 53°C for 30 sec, 72°C for 1 min) and by extension at 72°C for 10 min. Two primers, 18SF and 18SR (Table 1), were also used to amplify 18S rRNA gene in the semi-quantitative RT-PCR as controls. The PCR products were separated on 1% agarose gels stained with ethidium bromide (10 μg/ml). Quantification of band strength was accomplished by scanning the

gel with the ultraviolet and visible light automobile analyzing instruments of FR-200A (FuRi, China) and using the image smart view software.

**Functional analysis of GbMECPS in *Escherichia coli*.** *E. coli* strain XL1-Blue, plasmids pAC-BETA and pTrc-ATIPI (provided by Dr. Francis X. Cunningham, Department of Cell Biology and Molecular Genetics, University of Maryland, USA) were used to test the biological function of GbMECPS. A complementation strategy was used in this experiment. The plasmid, pAC-BETA, retains functional genes for geranylgeranyl pyrophosphate synthase (*crtE*), phytoene synthase (*crtB*), phytoene desaturase (*crtL*) and lycopene cyclase (*crtY*). It also retains a chloramphenicol resistance gene. Cells of *E. coli* containing this plasmid produce and accumulate β-carotene, resulting in yellow colonies (Cunningham *et al.*, 1994, 1996). The plasmid, pTrc-ATIPI, retains an ampicillin resistance gene and an IPI gene whose product can accelerate the accumulation of β-carotene (Cunningham *et al.*, 2000). The plasmid pAC-BETA was introduced into *E. coli* XL1-Blue to create a β-carotene accumulating strain of *E. coli*. The coding region of *Gbmecps* was amplified by PCR using primers MORF-FPB and MORF-RPN (Table 1). Both of the fragment and the plasmid pTrc-ATIPI were digested with *Bgl*II and *Not*I for 10 h. Subsequently, the coding region of *Gbmecps* was cloned into the expression vector pTrc to obtain the plasmid Trc-MECPS. The pTrc-MECPS was transformed into the XL1-Blue containing pAC-BETA. Transformants were cultured on solid LB medium containing ampicillin (150 mg/l) and chloramphenicol (50 mg/l) at 28°C for 48 h. The color of the transformants can be used as a visible marker to test if GbMECPS can accelerate the accumulation of β-carotene.

## Results

**Cloning of the full-length cDNA of *Gbmecps*.** By comparing and aligning the sequences of the core fragment, the 5' RACE and 3' RACE products, the full-length cDNA sequence of *Gbmecps* was deduced and subsequently amplified using One Step PCR Kit (TaKaRa) with primers M-BEGIN and M-END (Table 1). The full-length cDNA of *Gbmecps* (GenBank accession No. AY971576) was 874 bp and contained an ORF of 720 bp, flanked by stretches of 38 bp and 116 bp at the 5'- and 3'- untranslated regions respectively. The stop codon (tga) and one putative polyadenylation signal site were recognized in the 3' untranslated region, which were followed by a short polyA tail. The ORF encoded a predicted polypeptide of 239 amino acid residues, with a predicted molecular mass of 26.0 kDa and pI of 8.83 (Fig. 1).

**Comparative and bioinformatic analyses of GbMECPS.** Sequence alignment using Vector NTI Suite 8.0 and BLAST research showed that the predicted GbMECPS from *G. biloba* had high identity with MECPSs from other plant species, such as *O. sativa* (Accession No. BAD29384), *A. thaliana* (Accession No. AAM62786) and *C. roseus* (Accession No. AAF65155), with the identity of 65.7, 64.0, and 63.2% respectively (Fig.

```
  1          ggtattcaagcatagcaatttcagctgtaatctgaagc
 39  atggctgccgcattctctgcacttgtattgcccaaatggacagat
     M  A  A  A  F  S  A  L  V  L  P  K  W  T  D
 84  gattcatctctcaaatctgcaatcagttgggttcataatcatcag
     D  S  S  L  K  S  A  I  S  W  V  H  N  H  Q
129  cttcgccccattaacttcagcacacacatctgtgtctccagagct
     L  R  P  I  N  F  S  T  H  I  C  V  S  R  A
174  ccaattttgcgcagaggaaatcaattcaaacggcctgtatgccaa
     P  I  L  R  R  G  N  Q  F  K  R  P  V  C  Q
219  tctgccaatccagcggttgagattaaaaaacagagtgctgtagca
     S  A  N  P  A  V  E  I  K  K  Q  S  A  V  A
264  aagagtgcgcctgtacttccgtttaggggttggacatggttttgat
     K  S  A  P  V  L  P  F  R  V  G  H  G  F  D
309  ttgcataggttggagcctgggctgcctctgataattggtgggata
     L  H  R  L  E  P  G  L  P  L  I  I  G  G  I
354  gatattcctcatgatagaggttgcgacgctcactcagatggtgat
     D  I  P  H  D  R  G  C  D  A  H  S  D  G  D
399  gtgctgcttcattgtgttgtggatgcaattttgggagcattgggt
     V  L  L  H  C  V  V  D  A  I  L  G  A  L  G
444  cttcctgacattgggcaactttttcctgataatgatcccaaatgg
     L  P  D  I  G  Q  L  F  P  D  N  D  P  K  W
489  cgtggagcagcttcttcagtttttcttgaaggaggcggtacgtctt
     R  G  A  A  S  S  V  F  L  K  E  A  V  R  L
534  atgcatgaggcagggtatgagcttgggaacttggatgcaacattg
     M  H  E  A  G  Y  E  L  G  N  L  D  A  T  L
579  atcttgcagagacccaagttgagtccccataaagaggctattcga
     I  L  Q  R  P  K  L  S  P  H  K  E  A  I  R
624  accaacctctgtgagttgctgggatctgatccttccgttattaat
     T  N  L  C  E  L  L  G  S  D  P  S  V  I  N
669  cttaaggctaagactcatgagaaggtcgacagtcttggggagaat
     L  K  A  K  T  H  E  K  V  D  S  L  G  E  N
714  cggagtatcgctgctcatacaattgtactttttgatgaagaag*tga*
     R  S  I  A  A  H  T  I  V  L  L  M  K  K  *
759  gttaatttgcattgttaatttgaattgtagtcagctatttgactg
804  taccatagtgtctgtttcccagtcataaaataataaatggcccat
849  tactatgacc*aaaaaaaaaaaaaaaaaa*874
```

**Fig. 1.** The full-length cDNA sequence and the deduced amino acid sequence of *G. biloba* 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (*mecps*). The start codon (atg) was underlined in bold and the stop codon (tga) was italically underlined in bold. The poly (A) tail was painted with shadow.

2). The results of BlastP also showed that the predicted GbMECPS belonged to the YgbB family, a single domain family. The ygbB protein is a putative enzyme of deoxy-xylulose pathway (terpenoid biosynthesis). Thus, the BLAST analysis indicated that GbMECPS belonged to the MECPS family. The crystal structure of the *E. coli* MECPS has been published (Steinbacher *et al.*, 2002). The homology-based 3-D structural modeling of GbMECPS was analyzed by Swiss-Modeling on the basis of the *E. coli* MECPS crystal structure and displayed by WebLab ViewerLite (Fig. 3). The 3-D
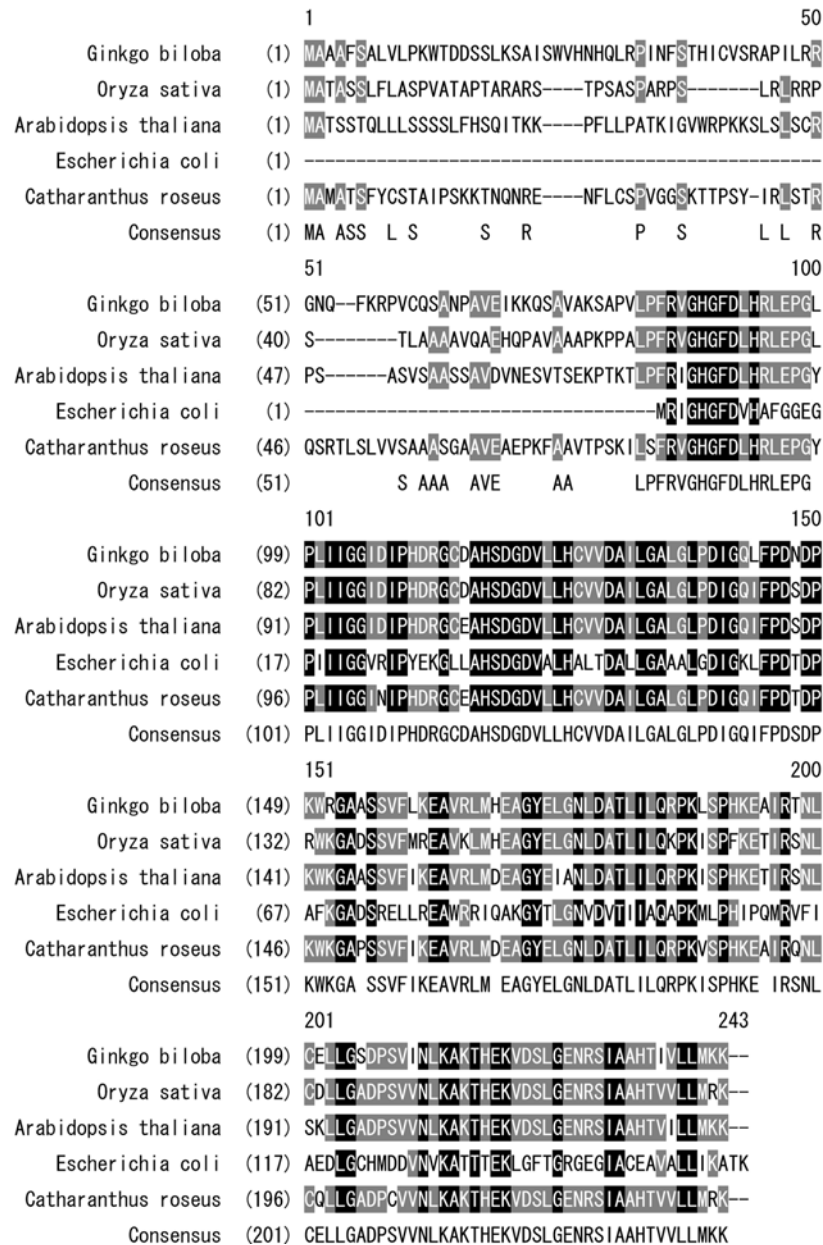
**Fig. 2.** Multi-alignment of amino acid sequences of GbMECPS and other MECPSs. The identical amino acids were showed in white with black background and the conserved amino acids were showed in white with gray background. The aligned MECPSs were from *Ginkgo biloba* (GenBank accession No. AY971576), *Oryza sativa* (GenBank accession No. BAD29384), *Arabidopsis thaliana* (GenBank accession No. AAM62786), *Catharanthus roseus* (GenBank accession No. AAF65155) and *Escherichia coli* (GenBank accession No. 1JY8A).

structure of *E. coli* MECPS consists of a large four-stranded β-sheet, a small two-stranded β-sheet and four α-helices (Steinbacher *et al.*, 2002). It has been proved that a metal ion zinc was coordinated in a tetrahedral geometry by Asp8, His10 and His42 in the *E. coli* MECPS, representing highly conserved residues within the protein family (Steinbacher *et al.*, 2002). The 3-D structure of GbMECPS is very similar to the *E. coli* MECPS, and sequence alignment showed that the highly conserved residues also existed in the GbMECPS (Fig. 2). Structure analysis and molecular modeling revealed that

GbMECPS strongly resembled the *E. coli* MECPS for two histidine residues and an aspartic acid provided three ligands in each protein, suggesting that they had potential catalytic similarities (Fig. 3). Protoplast targeting analysis predicted a chloroplast transit peptide (cTP) consisting of 59 residues at the N-terminal of GbMECPS.

**Molecular evolution analysis.** *Gbmecps* was the first *mecps* gene cloned from gymnosperm plants. Therefore it would be interesting to investigate its evolutionary position among the
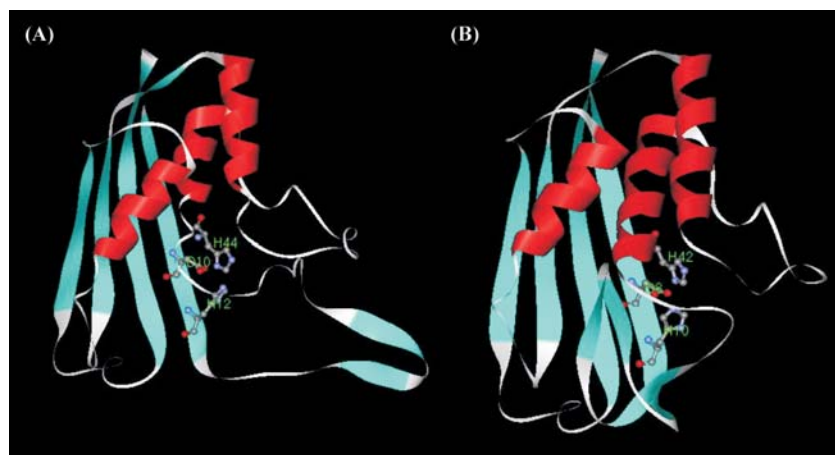
**Fig. 3.** The 3-D structures of GbMECPS (A) and *E. coli* MECPS (B). A metal ion was co-ordinated in a tetrahedral geometry by two histidine residues (H) and an aspartic acid (D) in each protein.
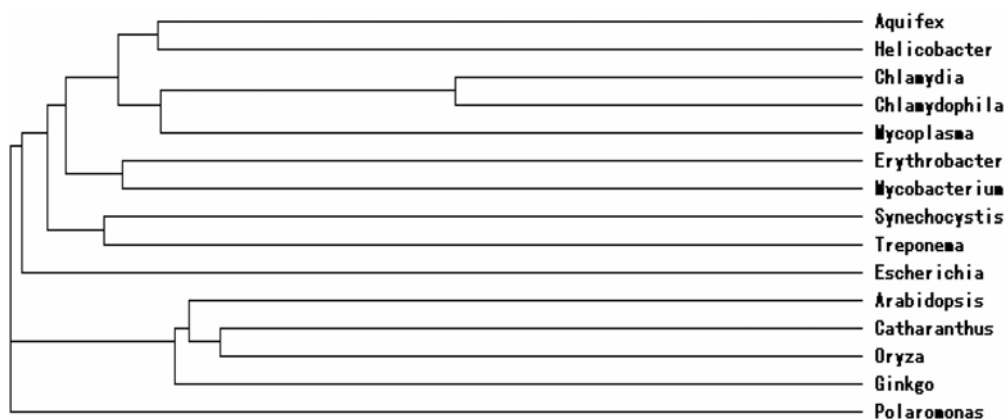


**Fig. 4.** Phylogenetic analysis of MECPSs from *G. biloba* and other species by MEGA version 2.1 from CLUSTAL W alignments. The neighbor-joining method was used to construct the tree. The MECPSs used in phylogenetic tree analysis were those from *Aquifex aeolicus* (GenBank accession No. O67089), *Helicobacter pylori J99* (GenBank accession No. Q9ZM19), *Chlamydia trachomatis* (GenBank accession No. O84441), *Chlamydophila pneumoniae* (GenBank accession No. Q9Z805), *Mycoplasma pneumoniae* (GenBank accession No. P75118), *Erythrobacter litoralis HTCC2594* (GenBank accession No. ZP 00376525), *Mycobacterium tuberculosis* (GenBank accession No. P96863), *Synechocystis sp. PCC 6803* (GenBank accession No. P73426), *Treponema pallidum* (GenBank accession No. O83525), *Escherichia coli* (GenBank accession No. 1JY8A), *Arabidopsis thaliana* (GenBank accession No. AAM62786), *Catharanthus roseus* (GenBank accession No. AAF65155), *Oryza sativa* (GenBank accession No. BAD29384), *Ginkgo biloba* (GenBank accession No. AY971576) and *Polaromonas sp. JS666* (GenBank accession No. EAM38704).

phylogenetic tree of various MECPSs. Using MEGA version 2.1 from CLUSTAL W alignments, a phylogenetic tree of MECPSs was constructed from different organisms including plants and bacteria. According to the phylogenetic tree, GbMECPS had higher identity with plant MECPSs than bacterium MECPSs and it was more ancient than other plant MECPSs (Fig. 4), which was consistent with the fact that *G. biloba* was one of the most ancient gymnosperm plant species. All the analysis results strongly suggest that GbMECPS is a plant MECPS protein involved in the mevalonate-independent biosynthesis.

**Southern blot analysis.** In order to determine the copy number of the *Gbmecps* in the genome of *G. biloba*, Southern blot analysis was carried out by digesting the genomic DNA of *G. biloba* with *Dra*I and *Hin*dIII respectively, followed by hybridization with the 720 bp *Eco*RI-*Hin*dIII fragment from the *Gbmecps* cDNA ORF sequence which was generated by PCR with primers MORF-FP and MORF-RP (Table 1). In the probe region, there are no *Dra*I and *Hin*dIII restriction sites. The result showed that only one hybridized band was present in each lane (Fig. 5), indicating that *Gbmecps* was a single copy gene.

**Tissue specific expression of *Gbmecps*.** To investigate the *Gbmecps* expression pattern in different parts of *G. biloba* plant, total RNA was extracted from roots, leaves and seeds respectively, and subjected to semi-quantitative RT-PCR
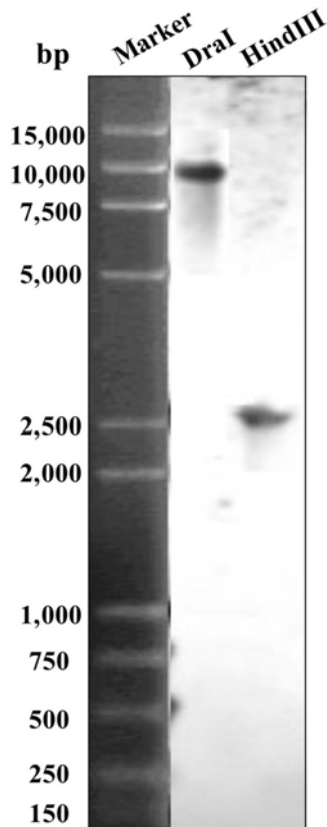
**Fig. 5.** DNA gel blot analysis of *Gbmecps*. Genomic DNA isolated from *G. biloba* leaves was digested overnight respectively with *Dra*I and *Hin*dIII, followed by hybridization with the coding sequence of *Gbmecps* as the probe.



**Fig. 6.** Expression pattern of *Gbmecps* in different tissues of *G. biloba* by semi-quantitative RT-PCR analysis. Total RNA samples were isolated from roots, leaves and seeds respectively, and subjected to semi-quantitative RT-PCR analysis (upper panel). The 18S rRNA gene was used as the control to show the normalization of the amount of templates in PCR reactions (lower panel).

analysis. The result showed that *Gbmecps* constitutively expressed in all the tested tissues but expressed the highest in roots, followed by in leaves, and the lowest in seeds (Fig. 6).

**Biological function analysis of Gbmecps in *E. coli*.** In the color complementation assay, we used the *E. coli* strain XL1-Blue to test if *Gbmecps* encoded the anticipated functional enzyme. The plastidial methylerythritol 4-phosphate (MEP) pathway exists in the *E. coli*. As a result, the XL1-Blue
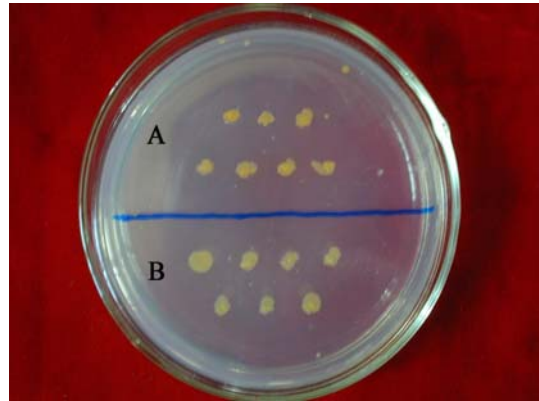


**Fig. 7.** Functional demonstration of GbMECPS activity, using *E. coli* strain L1-Blue. *E. coli* cells were transformed with pAC-BETA and pTrc-MECPS (A) and pAC-BETA (B).

containing pAC-BETA can manufacture and accumulate β-carotene and form yellow colonies. When pTrc-MECPS containing the coding region of *Gbmecps* was transformed into this β-carotene accumulating *E. coli* XL1-Blue, the color of the bacteria turned from yellow to orange-yellow, indicating that GbMECPS could accelerate the accumulation of β-carotene (Fig. 7). Therefore, the functioning of the GbMECPS has been confirmed by the change in the color of *E. coli* cultures.

## Discussion

The *mecps* genes have been cloned and characterized from bacterium (Deckert *et al.*, 1998; Steinbacher *et al.*, 2002), Chlamydia (Stephens *et al.*, 1998; Kalman *et al.*, 1999) and a few plant species (Veau *et al.*, 2000; Sasaki *et al.*, 2002). However there are no reports on the cloning of *mecps* genes from gymnosperm plant species including *G. biloba*. In the present study, we reported, for the first time, on the cloning and characterization of the gene encoding 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (MECPS), the fifth-step enzyme involved in ginkgolide biosynthesis in *G. biloba* and analyzed its function through color complementation assay. The deduced amino acid sequence of GbMECPS showed extensive similarity to the counterparts of other species.

The 3-D structure of GbMECPS is very similar to the *E. coli* MECPS. It has been demonstrated that a metal ion was coordinated in a tetrahedral geometry by Asp8, His10 and His42 in the *E. coli* MECPS, representing highly conserved residues within the protein family. This ion was identified as tightly bound zinc. The importance of the coordinating residues and of the bound metal ion was demonstrated by site-directed mutants Asp8Ser and His42Ser, which were enzymatically inactive (Steinbacher *et al.*, 2002). These highly conserved residues also existed in the GbMECPS, suggesting it has similar biological function. The results of BlastP showed that

the predicted GbMECPS from *G. biloba* belonged to the YgbB family, a single domain family. Both GbMECPS and *E.coli* MECPS had this conserved domain. However, GbMECPS was longer than *E.coli* MECPS at the N-terminal. Protoplast targeting analysis showed that a chloroplast transit peptide (cTP) consisting of 59 residues was predicted at the N-terminal of GbMECPS. This was understandable for GbMECPS was an enzyme of the MEP pathway which was a plastidial pathway in plants. The cTP helps the transportation of GbMECPS into the chloroplast. Therefore the N-terminal of GbMECPS may be important for its function.

Phylogenetic analysis indicates that GbMECPS is more ancient than other plant MECPSs, which supports the fact that *G. biloba* is one of the most ancient gymnosperm plant species. All the molecular analyses showed that *Gbmecps* was very similar to other *mecps*s, suggesting that *Gbmecps* belonged to *mecps* superfamily.

Plants can use both classical cytosolic mevalonate (MVA) pathway and plastidial methylerythritol 4-phosphate (MEP) pathway to produce terpenoids. It is found that that sterols, ubiquinone and sesquiterpenes are mainly biosynthesized from MVA pathway while isoprene, monoterpenes, diterpenes, carotenoids and abscisic acid are mainly biosynthesized from MEP pathway (Bick and Lange, 2003). The main active ingredients in seeds of *G. biloba* are bilobalides which are kinds of sesquiterpenes. They are mainly biosynthesized from MVA pathway. Therefore it is understandable that the expression of *Gbmecps* in seeds of *G. biloba* is the lowest.

Color complementation assay demonstrated that GbMECPS helped to accelerate the accumulation of β-carotene. Since both β-carotene and ginkgolide, a kind of diterpenes, are mainly biosynthesized through MEP pathway (Bick and Lange, 2003), it can be expected that GbMECPS helps to accelerate the accumulation of ginkgolide. Therefore, *Gbmecps* may be a target gene in metabolic engineering to improve the contents of ginkgolides.

MECPS catalyzes the conversion of 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate into 2C-methyl-D-erythritol 2, 4-cyclodiphosphate, which is reported to be a metabolic intermediate of the alternative non-mevalonate pathway for terpenoid biosynthesis. Detailed knowledge of the mechanisms and regulation of the pathway will benefit the biotechnological production of commercially interesting isoprenoids, such as terpenoid MIA (Veau *et al.*, 2000) and novel herbicides (Lichtenthaler *et al.*, 2000). Furthermore, the MEP pathway is present in plants and bacteria but not in humans, making MECPS an ideal target for the development of novel classes of highly specific antibiotics and antimalarial agents (Rohmer *et al.*, 1998). For example, many human pathogens are in the non-mevalonate pathway group, including *Plasmodium falciparum*, the causative agent of malaria, which belongs to the phylum *Apicomplexa*, a group of obligate endoparasites. In the latter, the non-mevalonate pathway resides in the apicoplast, a non-photosynthetic plastid system that is probably acquired by secondary endosymbiosis of a cyanobacterial-like

prokaryote but remains fundamentally bacterial in nature and is different from the cytoplasmic, eukaryotic systems of their hosts. Therefore, biochemical pathways located in the apicoplast have been recognized as attractive drug targets in general (Steinbacher *et al.*, 2002).

MECPS is considered to play an important role in the control of ginkgolides biosynthesis and is an important potential target for the control of flux of the MEP pathway. On the basis of the molecular genetics of ginkgolides biosynthesis, ginkgolides might be obtained alternatively through metabolic engineering. Overexpression of the *mecps* may lead to increased flux through the pathway with consequent increase in the rate of formation of plastidial monoterpene ($C_{10}$), diterpene ($C_{20}$) and tetraterpene ($C_{40}$) products, as well as related prenylated metabolites. The cloning and characterization of *Gbmecps* will be helpful to understand more about the role of *mecps* involved in the ginkgolides biosynthesis at the molecular level. Based on the cloning and characterization of *Gbmecps*, plant expression vector containing the *Gbmecps* has been constructed and genetic transformation of *G. biloba* is undergoing in order to test its potential role in improving ginkgolides production by genetic engineering.

## References

Bick, J. A. and Lange, B. M. (2003) Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.* **415**, 146-154.

Bloch, K. (1992) Sterol molecule - structure, biosynthesis, and function. *Steroids* **57**, 378-383.

Bouvier, F., d'Harlingue, A., Suire, C., Backhaus, R. A. and Camara, B. (1998) Dedicated roles of plastid transketolases during the early onset of isoprenoid biogenesis in pepper fruits. *Plant Physiol.* **117**, 1423-1431.

Carrier, D. J., van Beek, T. A., van der Heijden, R. and Verpoorte, R. (1998) Distribution of ginkgolides and terpenoid biosynthetic activity in *Ginkgo biloba*. *Phytochemistry* **48**, 89-92.

Chappell, J. (1995) Biochemistry and molecular-biology of the isoprenoid biosynthetic-pathway in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 521-547.

Cunningham, F. X., Sun, Z. R., Chamovitz, D., Hirschberg, J. and Gantt, E. (1994) Molecular structure and enzymatic function of lycopene cyclase from the *Cyanobacterium synechococcus* Sp Strain Pcc7942. *Plant Cell* **6**, 1107-1121.

Cunningham, F. X., Pogson, B., Sun, Z. R., McDonald, K. A., DellaPenna, D. and Gantt, E. (1996) Functional analysis of the beta and epsilon lycopene cyclase enzymes of arabidopsis

reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell* **8**, 1613-1626.

Cunningham, F. X. and Gantt, E. (2000) Identification of multigene families encoding isopentenyl diphosphate isomerase in plants by heterologous complementation in *Escherichia coli. Plant Cell Physiol.* **41,** 119-123.

Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. and Swanson, R. V. (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus. Nature* **392**, 353-358.

Eisenreich, W., Rohdich, F. and Bacher, A. (2001) Deoxyxylulose phosphate pathway to terpenoids. *Trends Plant Sci.* **6**, 78-84.

Han, Y. S., Roytrakul, S., Verberne, M. C., van der Heijden, R., Linthorst, H. J. M. and Verpoorte, R. (2003) Cloning of a cDNA encoding 1-deoxy-D-xylulose 5-phosphate synthase from *Morinda citrifolia* and analysis of its expression in relation to anthraquinone accumulation. *Plant Sci.* **164**, 911-917.

Hosford, D. J., Domingo, M. T., Chabrier, P. E. and Braquet, P. (1990) Ginkgolides and platelet-activating factor binding sites. *Methods Enzymol.* **187**, 433-446.

Jacobs, B. P. and Browner, W. S. (2000) *Ginkgo biloba*: A living fossil. *Am. J. Med.* **108**, 341-342.

Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, L., Hyman, R. W., Olinger, L., Grimwood, L., Davis, R. W. and Stephens, R. S. (1999) Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis. Nat. Genet.* **21**, 385-389.

Kumar, S., Tamura, K., Jakobsen, I. B. and Nei, M. (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244-1245.

Lange, B. M., Rujan, T., Martin, W. and Croteau, R. (2000) Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. USA* **97**, 13172-13177.

Laurain, D., Tremouillaux, G. J., Chenieux, J. C. and van Beek, T. A. (1997) Production of ginkgolide and bilobalide in transformed and gametophyte derived cell cultures of *Ginkgo biloba. Phytochemistry* **46**, 127-130.

Liao, Z., Chen, M., Gong, Y., Guo, L., Tan, Q., Feng, X., Sun, X., Tan, F. and Tang, K. (2004) A new geranylgeranyl diphosphate synthase gene from *Ginkgo biloba*, which intermediates the biosynthesis of the key precursor for ginkgolides. *DNA Seq.* **15**, 153-158.

Lichtenthaler, H. K., Zeidler, J., Schwender, J. and Muller, C. (2000) The non-mevalonate isoprenoid biosynthesis of plants as a test system for new herbicides and drugs against pathogenic bacteria and the malaria parasite. *Z. Naturforsch. (C)* **55**, 305-313.

Rodriguez-Concepion, M. (2004) The MEP pathway: A new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr. Pharm. Des.* **10**, 2391-2400.

Rodriguez-Concepcion, M. and Boronat, A. (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol.* **130**, 1079-1089.

Rohmer, M. (1998). Isoprenoid biosynthesis via the mevalonate-independent route, a novel target for antibacterial drugs? *Prog. Drug Res.* **50,** 135-154.

Sacchettini, J. C. and Poulter, C. D. (1997) Biochemistry - creating isoprenoid diversity. *Science* **277**, 1788-1789.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.

Sasaki, S., Obara, M., Kashiba, K., Sato, T., Yano, M., Ebitani, T. and Yamaya, T. (2002) Linkage analysis and characterization for QTL on chromosome 2 that associated with cytosolic glutamine synthetase content and panicle weight in rice (*Oryza sativa* L.). *Plant Cell Physiol.* **43**, 71-71.

Schwarz, M. and Arigoni, D. (1999) Ginkgolide biosynthesis; in *Comprehensive Natural Product Chemistry*, Cane, D. (ed.), pp. 367-399, Pergamon, Oxford, UK.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381-3385.

Sprenger, G. A., Schorken, U., Wiegert, T., Grolle, S., deGraaf, A. A., Taylor, S. V., Begley, T. P., Bringermeyer, S. and Sahm, H. (1997) Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc. Natl. Acad. Sci. USA* **94**, 12857-12862.

Steinbacher, S., Kaiser, J., Wungsintaweekul, J., Hecht, S., Eisenreich, W., Gerhardt, S., Bacher, A. and Rohdich, F. (2002) Structure of 2C-methyl-D-erythritol-2,4-cyclodiphosphate synthase involved in mevalonate-independent biosynthesis of isoprenoids. *J. Mol. Biol.* **316**, 79-88.

Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L., Zhao, Q. X., Koonin, E. V. and Davis, R. W. (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis. Science* **282**, 754-759.

Vanbeek, T. A., Scheeren, H. A., Rantio, T., Melger, W. C. and Lelyveld, G. P. (1991) Determination of Ginkgolides and Bilobalide in *Ginkgo biloba* Leaves and Phytopharmaceuticals. *J. Chromatography* **543**, 375-387.

Veau, B., Courtois, M., Oudin, A., Chenieux, J. C., Rideau, M. and Clastre, M. (2000) Cloning and expression of cDNAs encoding two enzymes of the MEP pathway in *Catharanthus roseus. Biochim. Biophys. Acta* **1517**, 159-163.

Zeidler, J. G., Lichtenthaler, H. K., May, H. U. and Lichtenthaler, F. W. (1997) Is isoprene emitted by plants synthesized via the novel isopentenyl pyrophosphate pathway? *Z. Naturforsch. (C)* **52**, 15-23.