

Cloning and Expression Analysis of a Farnesyl Diphosphate Synthase (FPPS) Gene from *Chamaemelum nobile*

Xiao-Meng LIU¹, Ting-Ting TAO¹, Xiang-Xiang MENG¹,
Wei-Wei ZHANG¹, Jie CHANG^{2,3}, Feng XU^{1*}

¹Yangtze University, College of Horticulture and Gardening, Jingzhou, Hubei 434025, China; LiuXM925@163.com;
18010872961@163.com; dreamer9212@163.com; wwzhangchn@163.com; xufeng198@126.com (*corresponding author)

²Hubei Collaborative Innovation Center of Targeted Antitumor Drug, Jingmen, Hubei 448000, China; 18986662379@163.com

³College of Chemical Engineering and Pharmacy, Jingchu University of Technology, Jingmen, Hubei 448000, China

Abstract

Farnesyl diphosphate synthase (FPPS), an isopentenyl transferase, catalyzes the condensation reaction of five carbon isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) to form fifteen carbon farnesyl pyrophosphate (FPP), which is the key precursor for sesquiterpene biosynthesis. In this study, a FPPS gene (*CnFPPS*) was cloned from *Chamaemelum nobile*. The full-length cDNA of *CnFPPS* is 1239 bp and contains an open reading frame (ORF) of 1029 bp encoding 342 amino acids. The theoretical molecular weight and pI of the CnFPPS protein are 39.38 kDa and 5.59, respectively. Multiple alignment analysis showed the protein sequence of CnFPPS had a high homology with FPPS proteins from other plants. The deduced amino acid of CnFPPS contained five conservative domains such as substrate binding pocket, substrate-Mg²⁺ binding site, catalytic site, aspartate-rich region 1 and 2, suggesting CnFPPS is one member of FPPS family in *C. nobile*. Phylogenetic analysis based on the amino acid sequences of FPPSs showed that CnFPPS was closely related to the FPPS of *Matricaria chamomilla*. The result of qRT-PCR revealed that *CnFPPS* gene was constitutively expressed in different tissues of *C. nobile*, with the highest expression in the root. These findings improve the understanding of the synthesis and regulation of the terpenoid compounds at the molecular level and lay a foundation for studying the regulatory functions of CnFPPS in terpenoid biosynthetic pathway in *C. nobile*.

Keywords: expression profile, farnesyl diphosphate synthase, Roman chamomile, sequence analysis

Introduction

Chamaemelum nobile, commonly referred to as “Roman chamomile”, a perennial herb belonging to the Asteraceae family native to Western Europe (Xiao, 2003), is a widely used medicinal plant due to its antioxidant and anti-inflammatory effects (Ma *et al.*, 2007). *C. nobile* is utilized to treat indigestion, appetite loss, nausea, vomiting, and other symptoms because it alleviates spasm and elicits anti-inflammatory and sedative effects (Srivastava *et al.*, 2010). *C. nobile* is rich in complex essential oils, and more than 100 kinds of components of essential oils have been identified from this plant (Fauconnier *et al.*, 2010). The active chemical constituents of the essential oils include terpenoids, flavonoids, esters, etc.

Among these chemical constituents, terpenoids are the main active components (Farhoudi, 2013). Therefore, we can significantly improve the quality and therapeutic value of *C. nobile* by increasing the content of terpenoids.

Terpenoids or isoprenoids, the most abundant secondary metabolite in terms of chemical structure, are a class of compounds composed of isoprene units (Yue *et al.*, 2011). Terpenoids are divided into the following groups based on the number of isoprene units: monoterpene (C₁₀), sesquiterpene (C₁₅), diterpene (C₂₀), and triterpene (C₃₀). More than 25,000 types of terpenoids have been identified in plants (Chang *et al.*, 2015). Many terpenoids are utilized for economic and chemical applications, including flavors, pigments, waxes, rubbers, and pharmaceuticals such as vitamins, taxol and artemisinin (Liu *et al.*, 2006). Terpenoids can also play an important

role in the essential biological processes of plants such as growth, development, reproduction, and adaptation to environmental challenges (Yu *et al.*, 2009). In higher plants, terpenoids are synthesized by the mevalonate (MVA) pathway in the cytoplasm and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway in plastids (Vranová *et al.*, 2012). The synthetic precursor of all terpenoid substances is isopentenyl pyrophosphate (IPP) (Enfissi *et al.*, 2005). As an important intermediate in the downstream of MVA pathway, farnesyl pyrophosphate (FPP) is generated by combining dimethylallyl pyrophosphate (DMAPP) with IPP and various types of terpenoids are consequently produced by the action of different terpene synthases and further by the modification such as glycosylation and methylation (He *et al.*, 2011; Ma *et al.*, 2015).

Farnesyl diphosphate synthase (FPPS), a key rate-limiting enzyme in the biosynthesis pathway of terpenoids, catalyzes 1, 4-head-to-tail condensation reaction of DMAPP and IPP into FPP (Lange *et al.*, 2000; Cornish, 1993). Many important secondary metabolites, such as terpene alcohols, sterols, ubiquinones, and carotenoids, are synthesized from FPP. Under the action of synthetic enzymes, FPP also generates various sesquiterpene compounds, such as (-)- α -bisabolone oxide A and α -bisabolol (Delourme *et al.*, 1994; Chappell *et al.*, 1995), which are essential for plants and humans. For example, phytoalexins help regulate plant growth, development, and immunity. Sesquiterpene compounds, such as α -bisabolol, induce anti-inflammatory, antibacterial, antispasmodic, sedative, analgesic, anticorrosion, antioxidant, and anticancer effects, and these compounds are important additives in several products, such as cosmetics and medicine (Sharkey *et al.*, 2013). FPPS is one of the most widely examined isopentenyl transferase (Ohto *et al.*, 1999). With the significance of this enzyme in isoprenoid metabolism, genes encoding FPPSs have been isolated and extensively characterized from many plants, including *Artemisia annua* (Matsushita *et al.*, 1996), peppermint (Lange *et al.*, 2000), cotton (Liu *et al.*, 1997), white lupine (Attucci *et al.*, 1995), and *Arabidopsis thaliana* (Delourme *et al.*, 1994).

Although *C. nobile* is an important medicinal plant, enzymes or genes involved in the biosynthetic pathway of its main medicinal constituents sesquiterpene compounds have been rarely identified. To investigate the biosynthesis pathway of terpenoids in *C. nobile*, we should identify and characterize each gene involved in this pathway.

In this study, we report the cloning and characterization of a full-length cDNA of FPPS (*CnFPPS*) from *C. nobile*. We also examined the expression pattern in different tissues in *C. nobile*.

Materials and Methods

Plant material

C. nobile was grown at 25/18 °C in a growth chamber (16 h/8 h light/dark). The leaves, stems, roots, and flowers of *C. nobile* were collected from the Botanical Garden at Yangtze University and immediately placed in a -80 °C until use. Primer synthesis and DNA sequencing were performed by Shanghai Sangon Biotechnology Company, China.

Cloning of *CnFPPS*

Total RNA was isolated from frozen plant tissues of *C. nobile* using the TaKaRaMiniBEST Plant RNA Extraction kit. The first-strand cDNA was synthesized using PrimeScript™ First-Strand cDNA Synthesis Kit according to manufacturer's instructions. The specific primers, namely, G1 and G2, (Table 1) as given below were designed based on the FPPS unigene sequence of *C. nobile* transcriptome data (GenBank accession number SRR4021832). *CnFPPS* cDNA was amplified under the following conditions: 94 °C for 4 min; 30 cycles of amplification at 94 °C for 30 s, 61 °C for 30 s, and 72 °C for 90 s; and a final extension at 72 °C for 10 min. The amplified products were detected through 1% gel electrophoresis and purified using agarose gel DNA purification Kit Ver. 4.0. The purified product was cloned into the pMD18-T vector and then transformed into *Escherichia coli* DH5 α . Positive clones were selected and sent to Shanghai Sangon Biotechnology Company for sequencing.

Bioinformatic analysis of *CnFPPS*

Gene sequencing was performed on the <http://www.ncbi.com> (NCBI) website using the BLAST tool to determine the similarity of the *CnFPPS* nucleotide sequence with other plant FPPS nucleotide sequences. The open reading frame (ORF) of the *CnFPPS* gene was predicted, and *CnFPPS* protein sequence was compared with other homology of plant FPPS protein sequences using Vector NTI 11.5. The calculated pI and molecular weight of the *CnFPPS* protein were computed using the software of Compute pI/Mw Tool at http://web.expasy.org/compute_pi/. Multiple sequence alignments were performed using the software Vector NTI 11.5 program. Phylogenetic analysis of *CnFPPS* protein sequence and other FPPS protein sequences from other plants was aligned with CLUSTALX 2, and subsequently, a phylogenetic tree was constructed through the neighbor-joining (NJ) method with MEGA 6 software.

Quantitative Real-time Polymerase Chain Reaction (qRT-PCR) analysis

Total RNA was extracted from roots, stems, leaves and flowers of *C. nobile* and reverse transcribed into cDNA. The

Table 1. Primer sequences in this study

Primer	Sequences (5'-3')
Upstream primer of <i>CnFPPS</i> gene (G1)	AAAGAAGCGATACATCACTGAC
Downstream primer of <i>CnFPPS</i> gene (G2)	CCACTCGCAAGACTCAAATCAG
Upstream primer of qRT-PCR (R1)	GTTGCCCTCTGCGTGTATGAGACTC
Downstream primer of qRT-PCR (R2)	GATTTTCTTTTCATCCGCTCTTGG
Upstream primer of 18S RNA (S1)	AACGAGCGTCGAGTGGATTAA
Downstream primer of 18S RNA (S2)	CCCATCGAAGGACTCCTATT

quantitative real-time polymerase chain reaction (qRT-PCR) primers (Table 1) of the *CnFPPS* gene were designed according to the real-time quantitative PCR kit of AceQ[®] qPCR SYBR[®] Green Master Mix (without ROX). The reference gene for the RT-PCR is *18S* RNA, and the upstream and downstream primers were S1 and S2, respectively (Table 1). qRT-PCR was performed on a Bio-Rad CFX Fluorescence Quantitative PCR instrument using the Vazyme AceQ[®] qPCR SYBR[®] Green Master Mix (Without ROX) kit instructions. The solution curve was added under the following conditions: 40 cycles of at 95 °C for 30 s, 95 °C for 5 s, and 60 °C for 30 s. qRT-PCR data were technically replicated with error bars, representing mean \pm SD (n = 3). The relative expression fold of each sample was calculated by its C_t value normalized to the C_t-value of reference gene using the 2^{- $\Delta\Delta$ C_t} method (Schmittgen et al., 2008).

Results

Cloning and sequence analysis of *CnFPPS*

A full-length cDNA of *CnFPPS* was 1239 bp long and contained a 1029 bp ORF encoding 342 amino acid proteins (Fig. 1). The results of BLASTN analysis at NCBI site showed that the cDNA sequence of *CnFPPS* had a high similarity to those of other *FPPS* genes. Collectively, these results indicated that the gene we cloned is the cDNA sequence of *CnFPPS*. Therefore, this gene was designated as *CnFPPS* (GenBank Accession No. KY432438).

Characterization of the deduced *CnFPPS* protein

The theoretical molecular weight and pI of the deduced *CnFPPS* protein were 39.38 KDa and 5.59, respectively.

```

1      AAAGAAGCGATACATCACTGACTTAAAAAAAAAGCGTGTTTTATTCCCAAAAACACAGCGGCATATTCACCT
73     TACAAATACCTCAACACATCACACACACTCAAACACACACAGTTTGTACAAACAGAATCTGTTATTCTGTTA
145    GCTAATTTGTATTTGAAAATGAGTACCGATCTGAAATCTAAGTTTTTAAAAGTGATGACACGCTTAAATCG
49     M S T D L K S K F L K V Y D T L K S
217    GAGCTAATTAACGATCCTGCTTTTGAATTTGATGATGATTCTCGTCGATGGGTTGAGAAGATGCTTGACTAC
73     E L I N D P A F E F D D D S R R W V E K M L D Y
289    AATGTACCTGGAGAAAAGCTAAACCGGGGACTATCTGTTGTCGACAGTTATCAGCTGCTTAAAGGAGGAGAA
97     N V P G G K L N R G L S V V D S Y Q L L K G G E
361    TTGACTGAAGAAGAGATTTTCTTTCATCCGCTCTTGGTTGGTGCATTGAATGGCTTCAAGCATACTTTCCT
121    L T E E E I F L S S A L G W C I E W L Q A Y F P
433    GTGCTTGATGATATCATGGATGAGTCTCATAACGACAGAGGGCAACCCTGTTGGTTTAGATTACCAAAGGTT
145    V L D D I M D E S H T R R G Q P C W F R L P K V
505    GGTATGATTGCTGCAAATGATGGAATCTTCTTCGCAACCATGTCCCGAGAATCTTAAAGAATCATTTCGGA
169    G M I A A N D G I L L R N H V P R I L K N H F R
577    GGAAAGCCTTACTATGTGGATCTTGTGGACCTGTCAACGAGGTTGAATCCAACAGCCTCGGGACAAATG
193    G K P Y Y V D L V D L F N E V E F Q T A S G Q M
649    ATTGATTTGATCACTACACTTGTGGAGAGAAAGATCTCTCAAAGTATTCAATTGTCTGTTCCACCGCCGATT
217    I D L I T T L V G E K D L S K Y S L S V H R R I
721    GTTCAATACAAAACAGCTTACTACTCATTTTACCTTCCAGTTGCCTGTGCACTCCTTATGTTTGGAGAGGAT
241    V Q Y K T A Y Y S F Y L P V A C A L L M F G E D
793    CTTGACAAGCATGTTGAAGTGAAGAATGTACTCGTTGAAATGGGTACCTATTTTCAAGTTCAGGACGATTAT
265    L D K H V E V K N V L V E M G T Y F Q V Q D D Y
865    CTAGACTGTTTTGGTACTCCCGAGGTGATTGGAAGATTGGAACCGATATTGAAGACTTTAAGTGCTCTTGG
289    L D C F G T P E V I G K I G T D I E D F K C S W
937    TTAGTTGTCAAAGCATTGAACTCGCTAACGAGGAACAAAACAAAATTCCTACATGAGAACTATGGGAAAAAG
313    L V V K A L E L A N E E Q T K F L H E N Y G K K
1009   GACCCCGCTTCCGTTGCAAAAGTGAAGGAAGTATACCACACTCTCAATCTTCAGGCTGTATTTGAAGATTAT
337    D P A S V A K V K E L Y H T L N L Q A V F E D Y
1081   GAGGCCACAAGCTACAAGAAGCTGATTACATCAATTGAAAATCACCCAAGCAAAGCAGTCCAAGCAGTGTG
361    E A T S Y K K L I T S I E N H P S K A V Q A V L
1153   AAATCTTCTTGGGTAAAACTACAAGAGGCAAAAAGTAAAGTAGATTTATTGCAGCAAATGATTTCTGATTTT
385    K S F L G K I Y K R Q K *
1225   GAGTCTTGCGAGTGG

```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of *CnFPPS*. The initial codon and the stop codon are highlighted in red square box; the primer sequences are underlined

Table 2. Nucleotide sequence of *CnFPPS* similarity to the FPPSs of other plant species

Species	GenBank Accession No.	Identity (%)	E-value
<i>Achillea asiatica</i>	JX424551.1	93%	0.0
<i>Tanacetum coccineum</i>	JX424559.1	95%	0.0
<i>Leucanthemum vulgare</i>	JX424557.1	94%	0.0
<i>Matricaria chamomilla</i>	KJ130321.1	96%	0.0
<i>Artemisia annua</i>	U36376.1	92%	0.0
<i>Artemisia tridentata</i>	AY308477.1	93%	0.0
<i>Aster ageratoides</i>	JX424562.1	88%	0.0
<i>Taraxacum kok-saghyz</i>	KJ558350.1	87%	0.0
<i>Leibnitzia anandria</i>	JX424569.1	86%	0.0

TcFPPS	(1)	<u>MS</u> <u>IDLKSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
McFPPS	(1)	<u>MGG</u> <u>SDLKSKFMDVYK</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVDKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
AcaFPPS	(1)	<u>MT</u> <u>IDLKSKFLQVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
LvFPPS	(1)	<u>MS</u> <u>IDLKSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
AraFPPS	(1)	<u>MS</u> <u>TDI</u> <u>KSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
TkFPPS	(1)	<u>MS</u> <u>TDI</u> <u>KSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
AsaFPPS	(1)	<u>MS</u> <u>TDL</u> <u>KSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
CnFPPS	(1)	<u>MS</u> <u>TDL</u> <u>KSKFLK</u> <u>QVYD</u> <u>TLKSEL</u> <u>INDPAFE</u> <u>DDDSRQ</u> <u>WVEKMLDYNVPGGKLN</u> <u>RGLSVVDSYQLL</u> <u>KGGE</u> <u>LDDE</u> <u>IFLSSAL</u>
Consensus	(1)	<u>M</u> <u>S</u> <u>T</u> <u>D</u> <u>L</u> <u>K</u> <u>S</u> <u>K</u> <u>F</u> <u>L</u> <u>K</u> <u>Q</u> <u>V</u> <u>Y</u> <u>D</u> <u>T</u> <u>L</u> <u>K</u> <u>S</u> <u>E</u> <u>L</u> <u>I</u> <u>N</u> <u>D</u> <u>P</u> <u>A</u> <u>F</u> <u>E</u> <u>D</u> <u>D</u> <u>S</u> <u>R</u> <u>Q</u> <u>W</u> <u>V</u> <u>E</u> <u>K</u> <u>M</u> <u>L</u> <u>D</u> <u>Y</u> <u>N</u> <u>V</u> <u>P</u> <u>G</u> <u>G</u> <u>K</u> <u>L</u> <u>N</u> <u>R</u> <u>G</u> <u>L</u> <u>S</u> <u>V</u> <u>V</u> <u>D</u> <u>S</u> <u>Y</u> <u>Q</u> <u>L</u> <u>L</u> <u>K</u> <u>G</u> <u>G</u> <u>E</u> <u>L</u> <u>D</u> <u>D</u> <u>E</u> <u>I</u> <u>F</u> <u>L</u> <u>S</u> <u>S</u> <u>A</u> <u>L</u>

TcFPPS	(80)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
McFPPS	(81)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
AcaFPPS	(79)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
LvFPPS	(80)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
AraFPPS	(79)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
TkFPPS	(79)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
AsaFPPS	(79)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
CnFPPS	(79)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>
Consensus	(81)	<u>GWC</u> <u>IEWLQAYFLVI</u> <u>DDIMDES</u> <u>HTRRGQPCWFR</u> <u>L</u> <u>PKVGMIA</u> <u>ANDGI</u> <u>LRNHVPRI</u> <u>LK</u> <u>KHFRGKPY</u> <u>Y</u> <u>DLV</u> <u>DLFNEVEF</u> <u>QTA</u>

TcFPPS	(160)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
McFPPS	(161)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
AcaFPPS	(159)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
LvFPPS	(160)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
AraFPPS	(159)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
TkFPPS	(159)	<u>SGQ</u> <u>MTDI</u> <u>ITTLVGEK</u> <u>DI</u> <u>SKYSI</u> <u>SVHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
AsaFPPS	(159)	<u>SGQ</u> <u>MTDI</u> <u>ITTLVGEK</u> <u>DI</u> <u>SKYSI</u> <u>SVHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
CnFPPS	(159)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>
Consensus	(161)	<u>SGQ</u> <u>MIDLITTLVGEK</u> <u>DLSKYSLS</u> <u>VHRRIVQYK</u> <u>TAYYSFY</u> <u>LPVACAL</u> <u>LMFGEDL</u> <u>DKHVEVK</u> <u>NVL</u> <u>VEMGTYFQ</u> <u>VGDDY</u> <u>LD</u> <u>DCF</u>

TcFPPS	(240)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
McFPPS	(241)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
AcaFPPS	(239)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
LvFPPS	(240)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
AraFPPS	(239)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
TkFPPS	(239)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
AsaFPPS	(239)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
CnFPPS	(239)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>
Consensus	(241)	<u>GAP</u> <u>EVIGKIGTD</u> <u>I</u> <u>EDFKCSW</u> <u>L</u> <u>VKALELANE</u> <u>EQK</u> <u>FLHENY</u> <u>GKKDP</u> <u>PASVAKV</u> <u>KE</u> <u>LYHT</u> <u>LN</u> <u>LQAV</u> <u>F</u> <u>DYE</u> <u>AT</u> <u>SYK</u> <u>KL</u> <u>ITS</u> <u>I</u>

TcFPPS	(320)	<u>FN</u> <u>RPSKAVQAVI</u> <u>KSFI</u> <u>GKTYK</u> <u>RQK</u>
McFPPS	(321)	<u>ES</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>
AcaFPPS	(319)	<u>EA</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>
LvFPPS	(320)	<u>EN</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>
AraFPPS	(319)	<u>EN</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>
TkFPPS	(319)	<u>FN</u> <u>HPSKAVQAVI</u> <u>KSFI</u> <u>GKTYK</u> <u>RQK</u>
AsaFPPS	(319)	<u>FN</u> <u>HPSKAVQAVI</u> <u>KSFI</u> <u>GKTYK</u> <u>RQK</u>
CnFPPS	(319)	<u>EN</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>
Consensus	(321)	<u>EN</u> <u>HPSKAVQAVL</u> <u>KSFL</u> <u>GKIYK</u> <u>RQK</u>

Fig. 2. Similarity analysis of *CnFPPS*-coding protein and other known FPPS proteins. TcFPPS, *T. coccineum*; McFPPS, *M. chamomilla*; AcaFPPS, *A. asiatica*; LvFPPS, *L. vulgare*; AraFPPS, *A. annua*; TkFPPS, *T. kok-saghyz*; AsaFPPS, *A. ageratoides*; CnFPPS, *C. nobile*. White on black portion represents exactly the same, white on grey represents a conservative district. Five conserved domains are underlined, which are the substrate binding pocket, aspartate-rich regions 1, catalytic site, substrate-Mg²⁺ binding site and aspartate-rich regions 2. Two aspartate-rich regions highlighted in square box

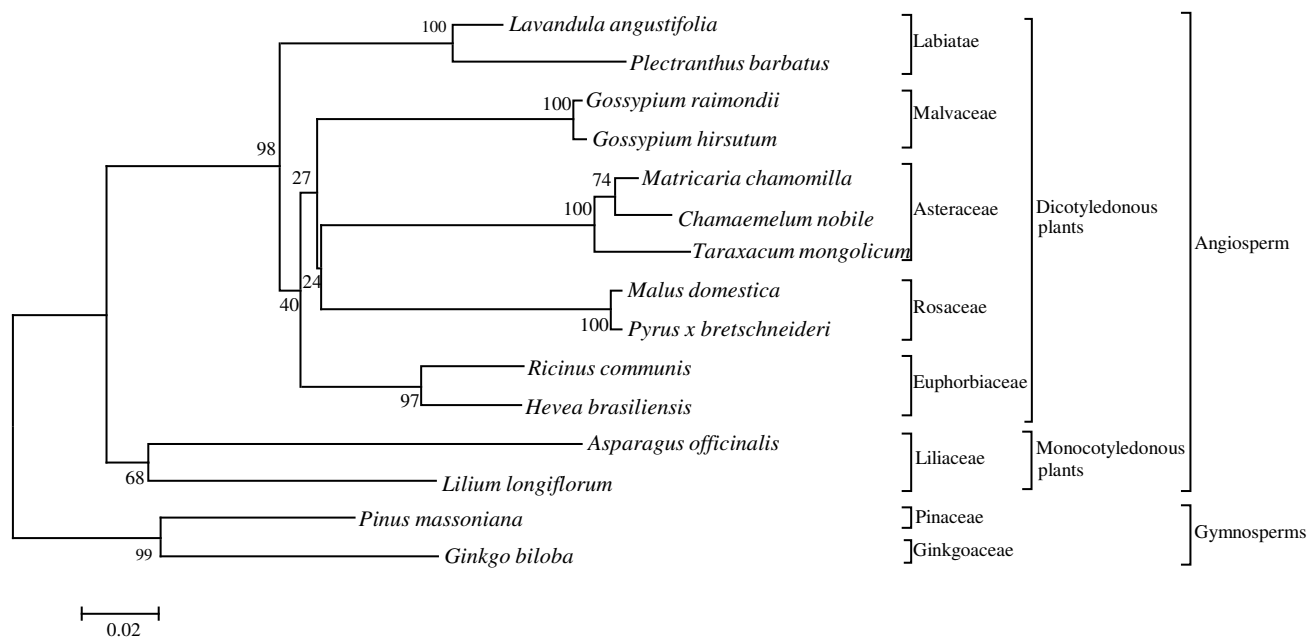


Fig. 3. Phylogenetic tree of genes in plant FPPS family

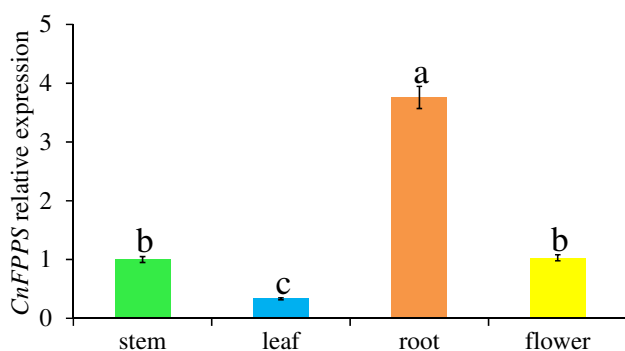


Fig. 4. Expression pattern of *CnFPPS* among different tissues of *C. nobile*

The nucleotide sequence of *CnFPPS* had a similarity of more than 85% to those with other *FPPS* genes (Table 2), indicating that the gene we cloned is a member of the *FPPS* gene family. Multiple sequence alignments revealed that *CnFPPS* had high homologous to *FPPS*s from other species. The *CnFPPS* protein showed 98%, 98%, 97%, 97%, 95%, 94%, and 92%, similarity to the counterparts of *Tanacetum coccineum*, *M. chamomilla*, *Achillea asiatica*, *A. annua*, *Leucanthemum vulgare*, *Taraxacum koksaghyz*, and *Aster ageratoides*, respectively. In addition, *CnFPPS* contained five conservative domains, namely, a substrate binding pocket, a substrate-Mg²⁺ binding site, a catalytic site, and aspartate-rich regions 1 and 2 (Fig. 2).

Analysis of the molecular evolution of *CnFPPS*

To investigate the evolutionary relationships in *CnFPPS* and *FPPS* proteins from other species, we selected the typical *FPPS* proteins from the GenBank. We constructed a phylogenetic tree using the software MEGA 6.0 with NJ

method. As shown in Fig. 3, *FPPS* phylogenetic tree is divided into two broad clades: one is angiosperm, including Labiatae, Euphorbiaceae, Malvaceae, Rosaceae and Asteraceae of Dicotyledoneae and Liliaceae of Monocotyledoneae; the other is gymnosperm, including Ginkgoaceae and Pinaceae. The *CnFPPS* protein has the closest relationship to *McFPPS* of *M. chamomilla* and *TmFPPS* of *Taraxacu mmongolicum*. They both belong to the Asteraceae family, indicating that the genetic relationship between the *FPPS*s gene of the Asteraceae family was close. Meanwhile, Asteraceae *FPPS* has the furthest relationship with Liliaceae, Ginkgoaceae, and Pinaceae. Liliaceae belongs to monocotyledonous plants, whereas Ginkgoaceae and Pinaceae belong to gymnosperms.

Expression analysis of *FPPS* in different tissues of *C. nobile*

qRT-PCR analysis showed that the *CnFPPS* gene was expressed in the roots, stems, leaves and flowers of *C. nobile*. However, *CnFPPS* gene expression was variuous in different tissues (Fig. 4), with the highest expression level in roots, followed by flowers and stems, and the lowest expression level in leaves.

Discussion

In recent years, remarkable progress has been made in the molecular regulation of sesquiterpenoid biosynthesis in plants (Yu et al., 2009; Degenhardt et al., 2009). *FPPS* is a key enzyme in the sesquiterpenoid biosynthetic pathway, and its activity is closely related to the accumulation of subsequent products. It has been studied to increase the content of active ingredients in plants by overexpressing the *FPPS* gene. For example, in *A. annua*, overexpressed the *FPPS* gene of *Asian cotton* and *A. annua*, could improve content of artemisinin (Han et al., 2006; Banyai et al., 2010). The amount of sterol and triterpene substances was increased with the increased expression of

ginseng *FPPS* gene in *Centella asiatica* (Kim et al., 2010). Thence, with the isolation and identification of secondary metabolite synthesis enzymes in the metabolic pathways of plants, the use of genetic engineering to alter the expression of enzymes and thus change the yield and species of secondary metabolism has become a new method to improve plant quality (Ren et al., 2005). Therefore, the overexpression of the *CnFPPS* gene in *C. nobile* is hypothesized to increase the sesquiterpene content of *C. nobile*.

In this study, multiple alignments showed that the deduced *CnFPPS* sequence exhibited high similarity to *FPPS* proteins from other plants. The *FPPS* gene of *C. nobile* may play similar biological functions as that of *FPPS* of other plants and participate in terpenoid biosynthesis in *C. nobile*. The phylogenetic tree indicated that *CnFPPS* has a distinct and long-standing relationship with the *FPPSs* from other Asteraceae plants. Moreover, protein motif analysis showed that the *CnFPPS* contained five conserved domains in the *FPPS* protein family, namely, substrate binding pocket, substrate-Mg²⁺ binding site, catalytic site, and aspartate-rich regions 1 and 2. Almost all *TPSs* have been reported to contain a typical conserved sequence of the sesquiterpene synthase gene family, that is, the aspartic acid-enriched motif, which is thought to have an important effect of binding to metal ions and guide substrate catalysis (Christianson et al., 2006). And the motif is thought to have an important effect of binding to metal ions and guide substrate catalysis (Christianson et al., 2006), suggesting that *CnFPPS* plays an important role in the production of sesquiterpenoids in *C. nobile*.

Several studies showed that the expression pattern of *FPPS* in plant tissues significantly varies across different plants. For example, *FPPS* is strongly expressed in leaf and root, moderately expressed in stem, and weakly expressed in the stem of *G. biloba* (Wang et al., 2004). *FPPS* genes are commonly found in tomato plants and are regulated during development (Gaffe et al., 2000). *FPPS* was mainly expressed in the tubular flowers, moderately expressed in the ligulate flowers and leaves, and least expressed in the root of *M. chamomilla* (Su et al., 2015). The expression pattern of *CnFPPS* in *C. nobile* revealed that the gene was expressed in all tissues but is expressed at a high level in the root, followed by that in flower and stem, and expression level was lowest in leaf. *C. nobile* is rich in active components and is among the most popular herbs since ancient times. The flowers of chamomile plants are commonly used for medicinal purposes. A recent chemical analysis of chamomile has shown that the content of sesquiterpene in flowers is highest and chamazulene is the most abundant sesquiterpene compound (Farhoudi et al., 2013). Given the high number of members of the *FPPS* gene family, the *CnFPPS* cloned in this study is only a part of its family and does not represent the function of the entire *FPPS* family in *C. nobile*. To better analyze the function of *FPPS* gene in *C. nobile* and determine whether the content of terpene compounds in *C. nobile* can be improved by transgenic technology of the gene obtained by cloning, further studies about gene cloning and functional analysis of other members of *FPPS* gene family of *C. nobile* are needed. Such studies will lay the foundation for studying the molecular mechanism of the biosynthesis of terrestrial compounds of *C. nobile*.

Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 31400603).

References

- Attucci S, Aitken SM, Ibrahim RK, Gulick PJ (1995). A cDNA encoding farnesyl pyrophosphate synthase in white lupin. *Plant Physiology* 108(2):835-836.
- Banyai W, Kirdmanee C, Mii M, Supaibulwatana K (2010). Overexpression of farnesyl pyrophosphate synthase (*FPS*) gene affected artemisinin content and growth of *Artemisia annua* L. *Plant Cell, Tissue and Organ Culture* 103(2):255-265.
- Chang J, Ning Y, Xu F, Cheng S, Li X (2015). Research advance of 3-hydroxy-3-methylglutaryl-coenzyme a synthase in plant isoprenoid biosynthesis. *Journal of Animal & Plant Sciences* 25(5):1441-1450.
- Chappell J (1995). Biochemistry and molecular biology of the isoprenoid biosynthetic pathway in plants. *Annual Review of Plant Biology* 46(1):521-547.
- Christianson DW (2006). Structural biology and chemistry of the terpenoid cyclases. *Chemical Reviews* 106(8):3412-3442.
- Cornish K (1993). The separate roles of plant cis and trans prenyl transferases in cis-1,4-polyisoprene biosynthesis. *European Journal of Biochemistry* 218(1):267-271.
- Degenhardt J, Köllner TG, Gershenzon J (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70(15-16):1621-1637.
- Delourme D, Lacroute F, Karst F (1994). Cloning of an *Arabidopsis thaliana* cDNA coding for farnesyl diphosphate synthase by functional complementation in yeast. *Plant Molecular Biology* 26(6):1867-1873.
- Enfissi E, Fraser PD, Lois LM, Boronat A, Schuch W, Bramley PM (2005). Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnology Journal* 3(1):17-27.
- Farhoudi R (2013). Chemical constituents and antioxidant properties of *Matricaria recutita* and *Chamaemelum nobile* essential oil growing wild in the south west of Iran. *Journal of Essential Oil Bearing Plants* 16(4):531-537.
- Fauconnier ML, Jaziri M, Homes J, Shimomura K, Marlier M (1996). *Anthemis nobilis* L. (Roman chamomile): *In vitro* culture, micropropagation, and the production of essential oils. *Medicinal & Aromatic Plants IX* pp 16-37.
- Gaffe J, Bru JP, Causse M, Vidal A, Stamitti-Bert L, Carde JP, Gallusci P (2000). LEFSP1, a tomato farnesyl pyrophosphate gene highly expressed during early fruit development. *Plant Physiology* 123(4):1351-1362.
- Han JL, Liu BY, Ye HC, Wang H, Li ZQ, Li GF (2006). Effects of overexpression of the endogenous farnesyl diphosphate synthase on the artemisinin content in *Artemisia annua* L. *Journal of Integrative Plant Biology* 48(4):482-487.
- He YF, Gao W, Liu TS, Li WY, Huang LQ (2011). Research advances of diterpene synthase. *Acta Pharmaceutica Sinica* (9):1019-1025.

- Kim OT, Kim SH, Ohyama K, Muranaka T, Choi YE, Lee HY, Hwang B (2010). Upregulation of phytosterol and triterpene biosynthesis in *Centella asiatica* hairy roots overexpressed ginseng farnesyl diphosphate synthase. *Plant Cell Reports* 29(4):403-411.
- Lange BM, Rujan T, Martin W, Croteau R (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proceedings of the National Academy of Sciences USA* 97(24):13172-13177.
- Lange BM, Wildung MR, Stauber EJ, Sanchez C, Pouchnik D, Croteau R (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proceedings of the National Academy of Sciences USA* 97(6):2934-2939.
- Liu C, Meng Y, Hou S, Chen X (1997). Cloning and sequencing of a cDNA encoding farnesyl pyrophosphate synthase from *Gossypium arboreum* and its expression pattern in the developing seeds of *Gossypium hirsutum* cv. 'Sumian-6'. *Acta Botanica Sinica* 40(8):703-710.
- Liu Z, Li X, Simoneau AR, Jafari M, Zi X (2006). Isoprenoid biosynthesis in plants: pathways, genes, regulation and metabolic engineering. *Journal of Biological Sciences* 6(1):371-374.
- Ma CM, Winsor L, Daneshtalab M (2007). Quantification of spiroether isomers and hemiarin of different parts of *Matricaria matricarioides* and flowers of *Chamaemelum nobile*. *Phytochemical Analysis* 18(1):42-49.
- Ma ZZ, Pang XQ, Sheng R, Pei XL, Wang QY, Xie T, Yin XP (2015). Research advances of key enzymes in the biosynthesis pathways of isoprenoids. *Journal of Hangzhou Normal University (Natural Science Edition)* 14(6):608-615.
- Matsushita Y, Kang W, Charlwood BV (1996). Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from *Artemisia annua*. *Gene* 172(2):207-209.
- Ohto C, Ishida C, Nakane H, Muramatsu M, Nishino T, Obata S (1999). A thermophilic cyanobacterium *Synechococcus elongatus* has three different Class I prenyltransferase genes. *Plant Molecular Biology* 40(2):307-321.
- Ren Y, Lu G, Cao JS, Li YJ (2005). Improving flavor by metabolic engineering of the terpenoid pathway in crop plants. *Chinese Journal of Cell Biology* 27(3):319-324.
- Schmittgen TD, Livak KJ (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* 3(6):1101-1108.
- Sharkey TD, Gray DW, Pell HK, Breneman SR, Topper L (2013). Isoprene synthase genes form a monophyletic clade of acyclic terpene synthases in the TPS-B terpene synthase family. *Evolution* 67(4):1026-1040.
- Srivastava JK, Shankar E, Gupta S (2010). Chamomile: A herbal medicine of the past with bright future. *Molecular Medicine Reports* 3(6):895-901.
- Su SS, Zhang HM, Liu XY, Pan GF, Ling SP, Zhang XS, Yang XM, Tai YL, Yuan Y (2015). Cloning and characterization of a farnesyl pyrophosphate synthase from *Matricaria recutita* L. and its upregulation by methyl jasmonate. *Genetics & Molecular Research* 14(1):349-361.
- Vranová E, Coman D, Gruitsem W (2012). Structure and dynamics of the isoprenoid pathway network. *Molecular Plant* 5(2):318-333.
- Wang P, Liao Z, Guo L, Li W, Chen M, Pi Y, Tang K (2004). Cloning and functional analysis of a cDNA encoding *Ginkgo biloba* farnesyl diphosphate synthase. *Molecules & Cells* 18(2):150-156.
- Xiao LYD (2003). Antipruritic effects of *Matricaria* extracts. *International Journal of Traditional Chinese Medicine* 25(3):173-174.
- Yu F, Utsumi R (2009). Diversity, regulation, and genetic manipulation of plant mono- and sesquiterpenoid biosynthesis. *Cellular & Molecular Life Sciences* 66(18):3043-3052.
- Yue YC, Fan YP (2011). The terpene synthases and regulation of terpene metabolism in plants. *Acta Horticulturae Sinica* 38(2):379-388.