An extremely promiscuous terpenoid synthase from the Lamiaceae plant *Colquhounia coccinea* var. *mollis* catalyzes the formation of sester-/di-/sesqui-/mono-terpenoids

De-Sen Li¹,4, Juan Hua¹,³, Shi-Hong Luo¹,³, Yan-Chun Liu¹, Yue-Gui Chen¹,4, Yi Ling¹, Kai Guo², Yan Liu¹,²,* and Sheng-Hong Li¹,²,*

¹State Key Laboratory of Phytochemistry and Plant Resources in West China and Yunnan Key Laboratory of Natural Medicinal Chemistry, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, P. R. China
²State Key Laboratory of Southwestern Chinese Medicine Resources and Innovative Institute of Chinese Medicine and Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, P. R. China
³College of Bioscience and Biotechnology, Shenyang Agricultural University, Shenyang 110866, P. R. China
⁴University of Chinese Academy of Sciences, Beijing 100049, P. R. China

*Correspondence: Yan Liu (liuyan@mail.kib.ac.cn), Sheng-Hong Li (shli@mail.kib.ac.cn)
https://doi.org/10.1016/j.xplc.2021.100233

ABSTRACT

Terpenoids are the largest class of natural products with complex structures and extensive bioactivities; their scaffolds are generated by diverse terpenoid synthases (TPSs) from a limited number of isoprenoid diphosphate precursors. Promiscuous TPSs play important roles in the evolution of terpenoid chemodiversity, but they remain largely unappreciated. Here, an extremely promiscuous terpenoid synthase (CcTPS1) of the TPS-b subfamily was cloned and functionally characterized from a leaf-specific transcriptome of the Lamiaceae plant *Colquhounia coccinea* var. *mollis*. CcTPS1 is the first sester-/di-/sesqui-/mono-TPS identified from the plant kingdom, accepting C₂₅/C₂₀/C₁₅/C₁₀ diphosphate substrates to generate a panel of sester-/di-/sesqui-/mono-terpenoids. Engineered *Escherichia coli* expressing CcTPS1 produced three previously unreported terpenoids (two sesterterpenoids and a diterpenoid) with rare cyclohexane-containing skeletons, along with four sesquiterpenoids and one monoterpenoid. Their structures were elucidated by extensive nuclear magnetic resonance spectroscopy. *Nicotiana benthamiana* transiently expressing CcTPS1 also produced the diterpenoid and sesquiterpenoids, demonstrating the enzyme’s promiscuity in planta. Its highly leaf-specific expression pattern combined with detectable terpenoid products in leaves of *C. coccinea* var. *mollis* and *N. benthamiana* expressing CcTPS1 suggested that CcTPS1 was mainly responsible for diterpenoid and sesquiterpenoid biosynthesis in plants. CcTPS1 expression and the terpenoid products could be induced by methyl jasmonate, suggesting their possible role in plant–environment interaction. CcTPS1 was localized to the cytosol and may differ from mono-TPSs in subcellular compartmentalization and substrate tolerance. These findings will greatly aid our understanding of plant TPS evolution and terpenoid chemodiversity; they also highlight the enormous potential of transcriptome mining and heterologous expression for the exploration of unique enzymes and natural products hidden in plants.

Key words: terpenoid diversity, terpenoid synthase, promiscuous enzyme, sester-/di-/sesqui-/mono-terpenoids, *Colquhounia coccinea* var. *mollis*

Li D.-S., Hua J., Luo S.-H., Liu Y.-C., Chen Y.-G., Ling Y., Guo K., Liu Y., and Li S.-H. (2021). An extremely promiscuous terpenoid synthase from the Lamiaceae plant *Colquhounia coccinea* var. *mollis* catalyzes the formation of sester-/di-/sesqui-/mono-terpenoids. Plant Comm. 2, 100233.
INTRODUCTION

Plants biosynthesize large numbers of structurally diverse special- ized metabolites in response to varying ecological environments. The increase in astounding chemodiversity is thought to have arisen through gene or genome duplication followed by sub- or neofunctionalization, including broadening of substrate selection and flattening of catalytic landscapes through divergent evolution (Weng et al., 2012). Consequently, promiscuous enzymes involved in the biosynthesis of specialized metabolites occur widely in plants and are considered to play important roles in the molecular evolution of enzymes with higher fidelity and activity (Yoshikuni et al., 2006). Numerous biochemical analyses have shown that alteration of a small number of amino acids in active or binding sites (plasticity residues) has a significant impact on enzyme activity, as well as substrate and product specificities, thus enabling the natural evolution of new enzymes with minimal investment (Yoshikuni et al., 2006; Xu et al., 2007; Morrone et al., 2008). Accordingly, exploitation of promiscuous enzymes and their underlying evolutionary mechanisms can contribute to the discovery of novel natural products and also to the rational design of efficient biocatalysts.

Terpenoids are the largest family of natural products, with more than 80 000 compounds discovered to date (Christianson, 2017). The vast majority of terpenoids are identified from plants and fulfill significant physiological roles (e.g. as hormones, photosynthetic pigments, and electron carriers) and ecological functions (including chemical defense, attraction, and communication) (Gershenzon and Dudareva, 2007; Pichersky and Raguso, 2018). In addition, plant terpenoids exhibit promising pharmacological activities and commercial values and have therefore been extensively used as pharmaceuticals (e.g. artemisinin and taxol), flavors (e.g. camphor and menthol), or pesticides (e.g. azadirachtin and pyrethrins). For decades, the complex structures and diverse bioactivities of plant terpenoids have inspired chemists and biologists to engage in the discovery of novel compounds in the huge terpenoid family. In particular, plant sesterterpenoids derived from geranylfarnesyl diphosphate (GFDP, C25) are a relatively unexplored subgroup of terpenoids, with only about 190 members characterized to date (Guo et al., 2021). Plant sesterterpenoids are widely distributed in lichens, ferns, and angiosperms, including the three important food crops rice (Oryza sativa) (Yang et al., 2014), wheat (Triticum aestivum) (Akihisa et al., 1999), and potato (Solanum tuberosum) (Toyoda et al., 1969). They show high structural diversity and remarkable biological functions involving anti-inflammatory, cytotoxic, antimicrobial, and antifeedant activities (Liu et al., 2016; Guo et al., 2021). However, the chemical diversity and biosynthesis of this unique class of specialized metabolites in plants remain largely unexplored, with the exception of the recent functional characterization of a few GFDP synthases and sester-TPSs from Brassicaceae and Lamiaceae plants (Nagel et al., 2015; Liu et al., 2016; West et al., 2016; Huang et al., 2017, 2018; Shao et al., 2017; Chen et al., 2019, 2020).

Diverse terpenoids are constructed from the universal C5 precursors isopentenyl diphosphate (IDP) and dimethylallyl diphosphate (DMADP), which are assembled by various prenyltransferases (PTs) into polyisoprenyl diphosphates with different chain lengths, such as geranyl diphosphate (GDP, C10), farnesyl diphosphate (FDP, C15), geranylgeranyl diphosphate (GGDP, C20), and GFDP. Subsequently, terpenoid synthases (TPSs) catalyze the conversion of linear and achiral polyisoprenyl diphosphates into diverse hydrocarbon and oxygenated scaffolds, which are further decorated by tailoring enzymes such as oxygenases and glycosyltransferases. The scaffold formation reactions catalyzed by TPSs are achieved via cationic intermediates generated by metal-triggered ionization (class I) or protonation-mediated cyclization (class II). These are considered to be the most complex chemical reactions in biology and play crucial roles in determining the diversity and complexity of terpenoids (Christianson, 2017). Recent researches have revealed that promiscuous TPSs that accept polyisoprenyl diphosphates with different chain lengths (substrate promiscuity) or convert a single substrate into diverse terpenoid skeletons (catalytic promiscuity) are prevalent within the plant kingdom (Bohlmann and Keeling, 2008; Degenhardt et al., 2009; Pazouki and Niinemets, 2016; Schreper et al., 2016). The spatial proportions of the active site pocket are likely to contribute to substrate promiscuity of TPSs, whereas catalytic promiscuity is probably dependent on conformational flexibility of the active center, as well as substrate and carbocyclic intermediates stabilized by active sites (Degenhardt et al., 2009; Pazouki and Niinemets, 2016). However, promiscuous TPSs that contribute to terpenoid diversity in plants have not been well appreciated. Most multi-substrate TPSs have been shown to convert two substrates (C10/C5, C15/C10, C20/C15 prenyl diphosphates) into the corresponding terpenoids, and a few of them also use C20/C15/C10 substrates (Pazouki and Niinemets, 2016; Johnson et al., 2019; Dhandapani et al., 2020). However, TPSs that accept four prenyl diphosphate substrates (C20/C20/ C15/C10) have not been reported to date, probably because commercial GFDP is not available.

Phylogenetically, plant TPSs can be classified into seven subfamilies designated TPS-a to TPS-g (Bohlmann et al., 1998; Chen et al., 2011). The TPS-b subfamily is dominated by monofunctional mono-TPSs from angiosperms, but a few sesqui- and di-TPSs and promiscuous TPSs have also been reported in this subfamily. For instance, EtTPS31 from Eremophila lucida and TwTPS27 from Tripterygium wilfordii were found to be class I di-TPSs, producing a macrocyclic diterpenoid 5-hydroxyxiscisane from GGDP (Gercke et al., 2020) and a miltiradiene from (+)-copalyl diphosphate (Hansen et al., 2017), respectively. MdfARs from apple (Malus domestica) and AtTPS03 from Arabidopsis thaliana are promiscuous TPSs, both utilizing FDP and GDP to generate multiple sesquiterpenoids and monoterpenoids, respectively (Green et al., 2007; Huang et al., 2010). Interestingly, some sesqui-TPSs and multi-substrate TPSs in the TPS-b subfamily, such as LaBerS and ObZis (Landmann et al., 2007; Davidovich-Rikanati et al., 2008), lack the N-terminal transit peptide and localize to the cytosol, implying that these enzymes may have evolved from mono-TPSs by losing the plastid signal peptide and broadening their substrate spectra.

Colquhounia coccinea var. mollis Prain is a Chinese-Himalayan species from the family Lamiaceae with great pharmaceutical and ornamental value (Li et al., 1994). Our previous phytochemical study indicated that this plant is particularly rich
in diverse terpenoids, especially sesterterpenoids. Defensive and immunosuppressive sesterterpenoids and norsesterterpenoids have been identified from *C. coccinea* var. *mollis* (Li et al., 2013; Jing et al., 2021); their 5/6/5 core structure is closely related to, but distinct from, leucosceptroids discovered in another Lamiaceae plant, *Leucosceptrum canum* (Luo et al., 2010). A geranylfarnesyl diphosphate synthase (GFDPS) that synthesizes the C25 prenyl diphosphate precursor of sesterterpenoids (Liu et al., 2016) and a cedrol synthase were functionally characterized from the glandular trichomes of *L. canum* by transcriptome mining (Luo et al., 2019). However, the biosynthesis of terpenoids, especially sesterterpenoids in *C. coccinea* var. *mollis*, remains unexploited. Here, we describe the molecular cloning and functional identification of an extremely promiscuous TPS (*CcTPS1*) from the TPS-b subfamily in *C. coccinea* var. *mollis*. *CcTPS1* can accept C25/C20/C15/C10 linear diphosphate substrates to generate a panel of sester-/di-/sesqui-/mono-terpenoids with cyclohexane-containing skeletons, including three previously undescribed terpenoids (two sesterterpenoids and a diterpenoid), along with four sesquiterpenoids and one monoterpenoid. Phylogenetic analysis in combination with a subcellular localization experiment and site-directed mutagenesis suggested that *CcTPS1* may have diverged from mono-TPSs via loss of the plastidial transit peptide and acquisition of the ability to accommodate the larger FDP, GGDP, and GFDP substrates.

**RESULTS**

**Terpenoid profiles and putative TPS annotation in *C. coccinea* var. *mollis***

Our previous phytochemical investigations led to the discovery of abundant sesterterpenoids from the peltate glandular trichomes and aerial parts of *C. coccinea* var. *mollis* (Li et al., 2013; Jing et al., 2021), indicating that sesterterpenoids are probably the major non-volatile chemical constituents of this plant. Here, gas chromatography–tandem mass spectrometry (GC–MS) was used to investigate the volatile terpenoid diversity of *C. coccinea* var. *mollis*; n-hexane extracts of fresh leaves, flowers, and epidermis-free tender stems were analyzed. As a result, 24 terpenoids, including 20 sesquiterpenoids and 4 diterpenoids tentatively identified based on their mass spectra, were detected in the leaves and flowers (Figure 1 and Supplemental Table 1). Intriguingly, the leaf extract showed a much higher diversity of terpenoids, especially sesquiterpenoids and diterpenoids, whereas rather limited numbers of metabolites were detected in the epidermis-free tender stems. In addition, an undescribed diterpenoid with a characteristic molecular ion peak at 272 in its mass spectrum was detected in the leaf extract. It was ultimately identified as compound 3 through comparison of its mass spectrum and retention time with those of purified compounds as described in the “Structural elucidation of terpenoids produced by CcTPS1” section.

Because of the significant difference in terpenoid profiles between leaves and epidermis-free tender stems, we generated a tissue-specific transcriptome library of *C. coccinea* var. *mollis* using the Illumina HiSeq 2000 sequencing platform. As a result, 60 674 unigenes were obtained, 31 775 of which were annotated by 4 public databases, including the NCBI non-redundant protein database (Nr), Swiss-Prot, Cluster of Orthologous Groups (COG), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Supplemental Figure 1 and Supplemental Table 2). Unigenes encoding all known enzymes in both the mevalonic acid (MVA) pathway and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway responsible for IDP and DMADP biosynthesis were discovered in the transcriptome library (Supplemental Table 3). In addition, 9 putative trans-PT and 6 putative cis-PT unigenes probably involved in the biosynthesis of polypropenyl diphosphates were found (Supplemental Table 3).
Homology-based searches against a set of characterized plant TPS sequences suggested 46 unigenes that were predicted to encode TPSs in the transcriptomes of leaves and epidermis-free tender stems. In total, 29 putative TPS unigenes were predicted to be sesqui-/mono-TPSs, 3 to be geranylilalool synthases, 9 to be copalyl diphosphate synthases, and 3 to be kaurane synthases. In addition, 4 unigenes were probably involved in triterpenoid biosynthesis. The expression levels of 42 putative TPS unigenes were higher than 0.5 reads per kilobase per million mapped reads (RPKM) in at least one sequenced tissue, and 13 of these unigenes had expression levels ranging from 10 to 146 RPKM. These putative TPSs may be responsible for the terpenoid chemodiversity in *C. coccinea* var. *mollis*. Consistent with the highly leaf-specific terpenoid profiles detected using GC–MS, 23 putative TPSs showed higher expression levels in leaves than in epidermis-free tender stems (Supplemental Table 4).

**Identification of a unique TPS from leaf-specific transcriptome**

The unigene no. 0039653 was highly expressed in the leaf transcriptome of *C. coccinea* var. *mollis* with an RPKM value of 28.2, but it was undetectable in the transcriptome of epidermis-free tender stems. Its expression level in different tissues of *C. coccinea* var. *mollis* was validated using quantitative real-time PCR (real time qPCR). As shown in Supplemental Figure 2, the candidate gene was expressed predominantly in leaves but at a low level in epidermis-free tender stems and flowers, consistent with the transcriptome results. In addition, the candidate unigene harbored the full N-terminal coding sequence and was predicted to be a mono-TPS without a plastid signal peptide (http://www.cbs.dtu.dk/services/TargetP-2.0/; Armenteros et al., 2019), unlike classical mono-TPSs that are plastid localized. The switching of subcellular compartmentalization may be accompanied by a broadening of substrate tolerance (Aharoni et al., 2004). Maximum-likelihood phylogenetic analysis was performed using amino acid sequences of plant TPSs obtained from the NCBI protein database, and seven subfamilies were defined (Figure 2). CcTPS1 clustered in the TPS-b subfamily, which consisted mainly of angiosperm monoterpenoid synthases but also included a few promiscuous TPSs and sesqui- and di-TPSs.

**Functional characterization of CcTPS1 as a promiscuous sester-/di-/sesqui-/mono-TPS in *E. coli***

For functional identification, the full-length ORF of CcTPS1 was cloned into the pCold-TF vector. A promiscuous GFDPs from the archaeon *Methanosarcina mazei* (MmGFDPs) capable of catalyzing the production of GFDP, GGDP, FDP, and GDP was used to provide diverse polyprenyl diphosphate precursors in the *E. coli* system (Ogawa et al., 2010). CcTPS1 was co-expressed with MmGFDPs in engineered *E. coli* harboring seven genes of the MVA pathway and an FDP synthase gene (Peralta-Yahya et al., 2011). GC–MS analysis of the extract from the resultant *E. coli* revealed multiple terpenoid products, including two sesterterpenoids, one diterpenoid, four sesquiterpenoids, and one monoterpentoid based on their characteristic molecular ion peaks at *m/z* 340 [M]+, 272 [M]+, 204 [M]+, and 154 [M]+, respectively, in the mass spectra (Figure 3).

The functions of CcTPS1 were further confirmed by *in vitro* enzymatic assays using purified recombinant protein incubated with different individual prenyl diphosphate substrates. GC–MS analysis indicated that CcTPS1 recombinant protein could catalyze GGDP, FDP, and GDP to generate a diterpenoid, four sesquiterpenoids, and a monoterpentoid, respectively (Supplemental Figures 4A and 5). Clearly, CcTPS1 showed extreme substrate and catalytic promiscuity, although *in vitro* enzymatic activity of
sesterterpenoid synthase was not measured because GFDP is not commercially available.

Structural elucidation of terpenoids produced by CcTPS1

To determine the structure of these enzyme products, 150 l of E. coli culture co-expressing CcTPS1, MmGFDPS, FDPS, and seven genes in MVA pathway was prepared and extracted. Afterward, a mixture of petroleum ether (PE) and ethyl acetate was used as a gradient eluent in repeated silica gel column chromatography to obtain compounds 1–7, including two novel sesterterpenoids (1 and 2), one new diterpenoid (3), four known sesquiterpenoids (4–7), and one known monoterpoid (8) (Figure 4A).

Compound 1 was obtained as a colorless oil, and high resolution electron ionization mass spectroscopy (HR-EI–MS) (m/z 340.3134 [M]+, calcd. 340.3130) revealed a molecular formula of \( \text{C}_{25}\text{H}_{40} \) with six degrees of unsaturation (Supplemental Figure 6). In the \(^1\)H NMR spectrum, six tertiary methyls at \( \delta_H 1.68 (3\ H, s) \), 1.64 (3H, s), 1.63 (3H, s), 1.61 (3H, s), and 1.60 (6H, overlap, \( \text{CH}_3 \times 2 \)) were clearly shown. Five olefinic signals at \( \delta_H 5.39 (s) \), 5.14, and 5.11 (3H, overlap) were attributable to five trisubstituted double bonds. In addition, a two-proton triplet was present at \( \delta_H 2.70 (2\ H, J = 7.0 \text{ Hz}) \), resonating from a methylene. Other signals were also displayed in the relatively high-field region between \( \delta_H 2.07 \) and \( \delta_H 1.47 \) and mostly overlapped, corresponding to either methine or methylene signals. The \(^{13}\)C NMR and DEPT spectra displayed 25 carbon resonances (see Supplemental Table 5), including six methyls (\( \delta_C 25.7, 23.5, 17.1, 16.1, 16.0, \) and 14.2), eight methylenes (\( \delta_C 39.7, 39.6, 30.8, 30.7, 27.9, 26.8, 26.7, \) and 26.6), six methines including five olefinic ones (\( \delta_C 124.4, 124.2, 123.4, 121.8, \) and 120.9), and five quaternary carbons (\( \delta_C 139.3, 134.9, 134.8, 133.7, \) and 131.3). The five double bonds accounted for five degrees of unsaturation, suggesting the existence of a monocycle in 1, indicating that compound 1 is likely to be a polyene-containing monocylic sesterterpenoid hydrocarbon. The 1D NMR spectra of 1 resembled those of somaliensene B (Yang et al., 2018),
and the only difference between them was that a vicinal-disubstituted double bond in somaliensene B was replaced by a trisubstituted double bond in 1. In the heteronuclear multiple bond correlation (HMBC) spectrum (Figure 4B) of 1, the $^{1}$H-$^{13}$C correlations of the protons at $\delta_{H}1.61$ (CH$_3$-24) with C-6 ($\delta_{C}42.8$), C-7 ($\delta_{C}139.3$), and C-8 ($\delta_{C}121.8$) indicated that a tertiary methyl was attached to C-7. The absolute configuration of 1 was identified as 6$^{R}$ by comparison of its optical rotation $\alpha_{D}^{24}+12.86$ (C$\chi$C l, CHCl$_3$) with those of (+)-axinyssene ($\alpha_{D}^{24}+34.6$ (C$\chi$C l, CHCl$_3$) (Kodama et al., 2003) and (+)-$\alpha$-bisabolene ($\alpha_{D}^{24}+54.3$ (C$\chi$C l, CHCl$_3$) (Fran et al., 1979), a sesquiterpenoid also isolated in this study ($\alpha_{D}^{24}+51.0$ (C$\chi$C l, CHCl$_3$)). Therefore, the structure of 1 was determined, and it was named (+)-$\alpha$-geranylbisabolene.

Compound 2 was isolated as a colorless oil, and its molecular formula was also determined to be C$_{25}$H$_{40}$ with six degrees of unsaturation based on HR-EI–MS ($m/z$ 340.3128 [M]+, calcd. 340.3130) and $^{13}$C NMR spectroscopic data (Supplemental Figure 7). The $^{1}$H and $^{13}$C NMR spectra of 2 revealed the presence of four trisubstituted double bonds ($\delta_{H}$ 5.40 [H-10], 5.14 [H-19], 5.11 [2H, H-14, and H-18]; $\delta_{C}$ 135.1 [C-15], 134.9 [C-11], 133.7 [C-3], 131.3 [C-19], 124.4 [C-18], 124.2 [C-14], 124.2 [C-10], 120.8 [C-2]), a terminal disubstituted double bond ($\delta_{H}$ 4.75 [2H, H-24]; $\delta_{C}$ 154.3 [C-7] and 107.2 [C-24]), and five vinylic methyls ($\delta_{H}$ 1.68 [3H, CH$_3$-21], 1.65 [3H, CH$_3$-25], 1.60 [9H, CH$_3$-20, 22 and 23]; $\delta_{C}$ 25.7 [C-21], 23.5 [C-25], 17.7 [C-20], 16.0 [C-22 and C-23]). These data suggested that 2 was also a monocyclic sesterterpenoid. Further comparison of the NMR data of 2 with those of somaliensene B clearly indicated the same planar structure for the two compounds (Yang et al., 2018). However, the optical rotation of 2 ($\alpha_{D}^{24}+5.3$ (C$\chi$C l, CHCl$_3$) was opposite to that of somaliensene B ($\alpha_{D}^{24}+93.0$ (C$\chi$C l, CHCl$_3$) suggesting that the absolute configuration of 2 is 6$^{R}$, as in 1. Accordingly, the structure of 2 was deduced to be (+)-somaliensene B.

Compound 3 was isolated as a colorless oil. Its molecular formula was determined to be C$_{20}$H$_{32}$ (5 degrees of unsaturation) based on HR-EI–MS ($m/z$ 272.2506 [M]$^{*}$, calcd. 272.2504) and $^{13}$C NMR spectroscopic data (Supplemental Figure 8). In the $^{1}$H NMR spectrum, five tertiary methyl resonances at $\delta_{H}$ 1.68, 1.64, 1.62, 1.61, and 1.60 (each 3H, s) were clearly shown. The $^{13}$C NMR and DEPT spectra exhibited 20 skeletal carbon signals: 5 methyls, 6 methylenes, 5 methines, and 4 quaternary carbons. Compound 3 was deduced to be a monocyclic diterpenoid to meet the requirement of unsaturation degrees on the basis of the analysis above. These data combined with the heteronuclear singular quantum correlation (HSQC) experiment showed that it was similar to (+)-$\alpha$-bisabolene (5), except for the presence of an extra trisubstituted double bond ($\delta_{H}$ 5.10 [H-14]; $\delta_{C}$ 124.3 [C-14], 131.3 [C-15], two methylenes ($\delta_{H}$ 2.06 [H$_2$-13], 1.98 [H$_2$-12]; $\delta_{C}$ 39.7 [C-12], 26.7 [C-13]), and one vinylic methyl group ($\delta_{H}$ 1.62 [3H, s, CH$_3$-18]; $\delta_{C}$ 16.1[C-18]). The key $^{1}$H-$^{1}$H correlation spectroscopy (COSY) correlations of H-9 ($\delta_{H} 2.70$) with H-8/H-10 and H-13 ($\delta_{H} 2.06$) with H-12/H-14, combined with the HMBC...
A promiscuous terpenoid synthase from a Lamiaceae plant

The structures of compounds was determined to be (+)-optical rotation (\[\alpha\]D +43.1 °, CHCl₃) with that of (+)-\(\alpha\)-bisabolene (\[\alpha\]D +54.3 °, CHCl₃) (Fran et al., 1979). Accordingly, compound 3 was determined to be (+)-\(\alpha\)-prenylibisabolene.

The structures of compounds 4–7 were straightforwardly identified as (+)-\(\alpha\)-bisabolol (Ying et al., 2010), (+)-\(\alpha\)-bisabolene (Fran et al., 1979; Sullivan et al., 1986), (+)-\(\beta\)-bisabolene (Scheffrahn et al., 1983), and \(\alpha\)-trans-bergamotene (Snider and Beal, 1988), respectively, through comparison of their NMR spectroscopic data (Supplemental Figures 9–12) and optical rotations (4: [\(\alpha\]D +20.1 °, CHCl₃; 5: [\(\alpha\]D +51.0 °, CHCl₃; 6: [\(\alpha\]D +66.3 °, CHCl₃; 7: [\(\alpha\]D +10.8 °, CHCl₃] with those reported in the literature. Compound 8 was confirmed to be \(\alpha\)-terpineol by comparison of its mass spectrum and retention time with those of an authentic sample.

Subcellular localization of CcTPS1 in the cytosol

In general, sester-, di-, sesqui-, and mono-terpenoids are synthesized in distinct compartments. For instance, mono-, di-, and sester-terpenoids are synthesized in the chloroplasts and sesquiterpenoids in the cytosol. Although CcTPS1 was predicted to be a cytosol-localized protein without a transit peptide using TargetP software, it was important to verify its subcellular localization experimentally in light of the multiple biochemical functions described above. The full-length cDNA of CcTPS1 was fused in-frame to green fluorescent protein (GFP) under the control of the cauliflower mosaic virus 35S promoter, then transiently expressed in Nicotiana benthamiana. Confocal laser scanning microscopy analysis showed a typical cytosol-localized pattern of CcTPS1::GFP, resembling that of the empty GFP vector (Figure 5A). These findings indicated that CcTPS1 was localized to the cytosol.

Functional confirmation of CcTPS1 as a promiscuous TPS that produces a diterpenoid as its major product in planta

Among the enzymatic products of CcTPS1 in E. coli, only diterpenoid 3 and sesquiterpenoid 7 were detectable in the leaf extract of mature C. coccinea var. mollis (Figure 1 and Supplemental Table 1). However, the cytosolic localization of CcTPS1 was inconsistent with this result, as diterpenoids are primarily biosynthesized in the chloroplasts. Therefore, the biochemical function of CcTPS1 was further investigated in planta. The A. thaliana GFDPS2 (AtGFDPS2) gene was used to provide linear prenyl diphosphate precursors in the plant system, as it produces predominantly GDP and minor amounts of GGDP and GDP (Nagel et al., 2015; Wang et al., 2016). CcTPS1 and AtGFDPS2 genes were individually constructed into the expression vector pEAQ-HT, which was designed for high-level transient gene expression in plants (Sainsbury et al., 2009). The resulting plasmids were transformed into Agrobacterium tumefaciens and were then mixed and co-infiltrated into the leaves of N. benthamiana. The co-expression of CcTPS1 and AtGFDPS2 genes in N. benthamiana resulted in the occurrence of multiple specific peaks that were identified as compounds 3–7 by comparison of their retention times and mass spectra with those of the purified compounds described above (Figure 5B). The results confirmed that CcTPS1 was a promiscuous TPS in planta. Unfortunately, sesterterpenoids 1 and 2 and monoterpenoid 8 were not detected in the transgenic N. benthamiana, probably owing to the low sester-/-mono-TPS activities of CcTPS1 in planta. Intriguingly, diterpenoid 3 was the major product of CcTPS1 expressed in N. benthamiana, unlike E. coli in which sesquerterpenoids were its major products. Differences between E. coli and N. benthamiana systems in terms of physiological

Figure 5. Subcellular localization and product profile analysis of CcTPS1 using N. benthamiana.

(A) GFP fusion proteins in N. benthamiana leaf epidermal cells visualized by laser confocal microscopy.

(B) GC–MS analysis of the extracts of N. benthamiana leaves harboring pEAQ-HT/CcTPS1 and pEAQ-HT/AtGFDPS2 (a), pEAQ-HT/AtGFDPS2 (b), and empty vector pEAQ-HT (c).

The co-expression of CcTPS1 and AtGFDPS2 genes in N. benthamiana mixed and co-infiltrated into the leaves of A. thaliana. The resulting plasmids were transformed into Agrobacterium tumefaciens and were then mixed and co-infiltrated into the leaves of N. benthamiana. The co-expression of CcTPS1 and AtGFDPS2 genes in N. benthamiana resulted in the occurrence of multiple specific peaks that were identified as compounds 3–7 by comparison of their retention times and mass spectra with those of the purified compounds described above (Figure 5B). The results confirmed that CcTPS1 was a promiscuous TPS in planta. Unfortunately, sesterterpenoids 1 and 2 and monoterpenoid 8 were not detected in the transgenic N. benthamiana, probably owing to the low sester-/-mono-TPS activities of CcTPS1 in planta. Intriguingly, diterpenoid 3 was the major product of CcTPS1 expressed in N. benthamiana, unlike E. coli in which sesquerterpenoids were its major products. Differences between E. coli and N. benthamiana systems in terms of physiological
Plant Communications

content and available enzymatic partners are likely to have caused the differences in the observed CcTPS1 product profiles.

Identification of a critical role for Trp262 in the catalytic activity of CcTPS1

Recent research has suggested that an amino acid with a small side chain in Arabidopsis thaliana sester-TPSs (Gly328 in AtTPS18 and AtTPS25, Gly325 in AtTPS19, Pro328 in AtTPS30) plays a critical role in the accommodation of larger prenyl diphosphate substrates. Mutation to aromatic tryptophan (Trp) resulted in the loss of sester-TPS activity and the acquisition of mono-, sesqui-, or di-TPS activities (Chen et al., 2019). Interestingly, a conserved aromatic Trp residue was found at the corresponding position of CcTPS1, as well as mono- and sesqui-TPSs. Therefore, site-directed mutagenesis of the corresponding Trp262 in CcTPS1 was performed to explore the underlying mechanism of its broadened substrate tolerance. First, the Trp residue was mutated to Gly or Pro with a small side chain, and the obtained variants CcTPS1W262G and CcTPS1W262P were individually co-expressed with MmGFDPs in E. coli that harbored all MVA pathway genes and the FDP synthase gene. Similar levels of soluble recombinant proteins were produced by the wild-type and mutated genes (Supplemental Figure 4B). However, both mutated proteins lost their TPS activity completely, as revealed by undetectable product peaks in their GC–MS chromatograms. The hydrophobic tryptophan was then changed to hydrophilic tyrosine, which also contains an aromatic side chain. The resulting CcTPS1W262Y variant could produce di- and sesqui-terpenoids, albeit with decreased enzymatic activity, but its sester- and mono-TPS activities were abolished (Supplemental Figure 13). These findings suggested that the presence of an aromatic and hydrophilic amino acid at position 262 is critical for the catalytic activity of CcTPS1. It is probably involved in stabilizing the carboxylation intermediates in the catalytic pocket, distinct from the mechanism of previously reported sester-TPSs in A. thaliana.

CcTPS1 expression and terpenoid accumulation in C. coccinea var. mollis induced by methyl jasmonate

Terpenoids have been reported to perform important biological functions, especially in defense against biotic and abiotic stresses (Gershenzon and Dudareva, 2007). We therefore examined CcTPS1 expression and terpenoid accumulation in C. coccinea var. mollis treated with the defense hormone methyl jasmonate (MeJA) using real time qPCR and GC–MS analyses. CcTPS1 transcripts were upregulated in 3- to 4-month-old plants at 3 h after MeJA treatment (Supplemental Figure 14), and both compounds 3 and 7 were detected in the plants after MeJA treatment for 12 h (Supplemental Figure 15). Strangely, none of these terpenoid products were detectable in the control plants (Supplemental Figure 15). The antifeedant activity of diterpenoid and sesquiterpenoid products 3–7 against the generalist plant-feeding insect Helicoverpa armigera was also tested as described previously (Li et al., 2013), but none of the compounds were active. These findings suggested that CcTPS1 and its products may be involved in plant–environment interactions, presumably in plant response to abiotic stresses.

DISCUSSION

Promiscuous TPSs have been considered to play important roles in the evolution of terpenoid chemodiversity, contributing to plant diversification with tremendous adaptability and also benefitting human beings because of the pharmaceutical and economic value of terpenoids (Pazouki and Ninemets, 2016; Consortium, 2018; Pichersky and Raguso, 2018). In this study, CcTPS1, with extreme substrate and catalytic promiscuity, was functionally characterized from the Lamiaceae plant C. coccinea var. mollis. Intriguingly, CcTPS1 is the first sester-/-di/-sesqui/-mono-TPS identified from the plant kingdom that can act on four prenyl diphosphate precursors (GFDP, GGDP, FDP, and GDP) to catalyze the formation of a variety of terpenoids. These include two previously undescribed sesterterpenoids (1 and 2) and a previously undescribed diterpenoid (3), as well as four sesquiterpenoids (4–7) and one monoterpenoid (8) that have been reported previously in many organisms. The extreme promiscuity of CcTPS1 may result in rapid changes in terpenoid profiles of C. coccinea var. mollis, presumably contributing to its adaptability in response to environmental changes. Because multi-substrate TPSs are considered to be common in the plant kingdom, more plant sester-/-di/-sesqui/-mono-TPSs should be characterized with the availability of GFDP using metabolic engineering approaches.

Sesterterpenoids 1 and 2 feature an alotane carbon framework that has not previously been isolated from any plants. Naturally occurring allotane sesterterpenoids have been reported from only a handful of marine sponges and also as the enzymatic products of a bacterial sesterterpenoid cyclase from the UbiA superfamily. Some serve as potent activators of cAMP cell signaling or exhibit cytotoxic activity (Yang et al., 2018; Li and Gustafson, 2021). In addition, compound 3 is a new member of the prenylribosabalone-type diterpenoids, a relatively rare subclass of diterpenoids previously discovered in a few plants of the Asteraceae, Lamiaceae, and Euphorbiaceae families and in marine organisms. They have been reported to have cytotoxic, antimalarial, and antiviral activities (Rucker et al., 1992; Kodama et al., 2003; Tesso and König, 2004; Gu et al., 2014). However, TPSs for constructing the scaffolds of allotane sesterterpenoids and prenylribosabalone diterpenoids have not been reported in plants to date. Therefore, characterization of the promiscuous CcTPS1 not only enriches the chemical diversity of plant sesterterpenoids and diterpenoids but also provides a material basis for further in-depth investigation of their biological activities.

The formation mechanism of compounds 1 and 2 is clearly different from those of reported plant sester-TPSs using either the type A (C1–IV–V, cyclization between C1-cation, C14–C15 olefin [double bond IV], and C18–C19 olefin [double bond V] of GFDP) or type B (C1–III–IV, cyclization between C1-cation, C10–C11 olefin [double bond III], and C12–C15 olefin [double bond IV] of GFDP) cyclization route (Huang et al., 2017; Minami et al., 2018; Chen et al., 2020, 2021) (Figure 6). According to the catalytic mechanism of the monoterpenoid and sesquiterpenoid products of CcTPS1 delineated in previous enzymological studies, the linear prenyl diphosphate precursors probably undergo an ionization–isomerization–reionization process to form cation A with different chain lengths. This enables electrophilic attack by C1 on the C6–C7 olefin (double bond II), resulting in the formation of the key carbocationic intermediate B containing a six-membered carbon ring (Figure 6). Carbocation B is terminated by loss of a proton to generate 2/6.
and 1/3/5 or by addition of water to yield 4/8. Alternatively, intermediate B may undergo sequential 2,7-cyclization and proton elimination to yield 7. Notably, the sesquiterpenoid products obtained from cation B via termination by proton elimination or water addition, as well as consecutive cyclization reactions followed by deprotonation, were found, demonstrating the catalytic promiscuity of CcTPS1. Similarly, corresponding sester-/di-/mono-terpenoid products may also exist, probably at trace amounts below the detection limit. Considering that bisabolene, terpineol, and their biosynthetic TPSs occur widely in plants, re-exploration of the substrate and product specificities of those TPSs and their underlying mechanisms is an interesting topic for future research. Such studies will not only facilitate the discovery of new TPSs and natural sester-/di-terpenoids with rare cyclohexane-containing skeletons but also enrich our understanding of the evolution of promiscuous TPSs.

CcTPS1 is one of the most highly expressed leaf-specific TPSs in C. coccinea var. mollis according to transcriptome analysis and real time qPCR results. The detection of diterpenoid 3 and sesquiterpenoid 7 in the leaves of C. coccinea var. mollis and N. benthamiana expressing CcTPS1 suggested that CcTPS1 may be responsible for the biosynthesis of diterpenoids and sesquiterpenoids in planta. Although diterpenoids have generally been shown to be biosynthesized in the plastids where GGDP is produced, cytosolic di-TPSs have also been reported over the last decade; these include geranylinalool synthases in A. thaliana and tomato (Herde et al., 2008; Zhou and Pichersky, 2020). CcTPS1 is an additional example of these rule-breaking TPSs, suggesting that GGDP is probably present in the cytosols of C. coccinea var. mollis and N. benthamiana because of GGDP transport from the plastids or in situ GGDP biosynthesis by cytosolic GGDP synthases. The induced CcTPS1 expression and the production of diterpenoid 3 and sesquiterpenoid 7 in MeJA-treated C. coccinea var. mollis suggested that they may participate in plant–environment interactions. Because none of these terpenoid products showed obvious antifeedant activity, we speculated that CcTPS1 and its products play a role in plant responses to abiotic rather than biotic stresses. However, the other six enzymatic products of CcTPS1 were neither isolated nor detected in C. coccinea var. mollis, probably because these products or their oxidized derivatives were below the detection limit or because CcTPS1 did not participate in the biosynthesis of these terpenoids in the plant. Alternatively, the biosynthesis and accumulation of these terpenoid products may be restricted to only a subset of specific growth and developmental stages, or

Figure 6. The proposed cyclization mechanism of sesterterpenoids 1 and 2 biosynthesized by CcTPS1 is distinct from that of recently reported plant sester-TPSs.

(A) Proposed cyclization scheme of compounds 1–8 produced by CcTPS1.

(B) Illustration of sesterterpenoid cyclization mechanisms catalyzed by previously reported plant sester-TPSs.
they may be tightly deployed to mediate interactions with certain environmental cues. CcTPS1 belongs to the TPS-b subfamily, which is thought to have arisen very early in the evolution of angiosperms and consists mainly of angiosperm mono-TPSs that typically contain plastid targeting signal peptides (Sharkey et al., 2013). In addition, the classical RR(X)4W motif in monoterpenoid synthases, probably required for the isomerization of GDP to form a linalyl cation, is highly conserved in CcTPS1 (Tholl and Lee, 2011). It has been speculated that CcTPS1 diverged from mono-TPS progenitors through loss of the plastidial transit peptide and acquisition of the ability to accommodate the larger FDP, GGD, and GFD substrates, sharing similar evolutionary routes with the TPS-b sesqui-TPS enzymes (Green et al., 2011). This hypothesis highlights the change in substrate specificity and subcellular compartmentalization during TPS evolution. Coincidentally, the recently reported Brassicaceae sester-TPSs in the TPS-a subfamily were suggested to have evolved from di-/sesqui-/mono-TPSs by duplication and neo-functionalization under positive selection, and a single residue played a crucial role in the acquisition of their sester-TPS activity (Chen et al., 2019). Therefore, the evolutionary routes for acquisition of sester-TPS activity may be distinct for CcTPS1 and Brassicaceae sester-TPSs. We also analyzed the impact of the CcTPS1 Trp262 residue, corresponding to the key residue of Brassicaceae sester-TPSs. Trp262 was found to play a critical role in the catalytic activity of CcTPS1, but it appeared not to influence substrate selection, probably owing to the distinct plasticity residues that modulate substrate specificity in different TPS subfamilies. Alternatively, the broad substrate selection of CcTPS1 may be a primordial trait, consistent with an inherent enzyme promiscuity and the diffuse spread of promiscuous terpenoid synthases across different TPS subfamilies (Johnson et al., 2019). Investigation of additional promiscuous TPSs, especially from the TPS-b subfamily, will provide further evidence to enable a better understanding of TPS evolution.

In summary, we cloned and functionally characterized a sester-/di-/sesqui-/mono-TPS from C. coccinea var. mollis. It produced a panel of eight terpenoids, including two new sesterterpenoids and a new diterpenoid with rare cyclohexane-containing skeletons. CcTPS1 tended to the TPS-b subfamily and was localized to the cytosol; we speculated that it diverged from mono-TPSs via changes in subcellular compartmentalization and substrate tolerance. CcTPS1 expression and terpenoid products were induced by the defense hormone MeJA, which suggested that they may have ecological functions in plant–environment interactions. This work enriches our knowledge of plant promiscuous TPSs and terpenoid diversity, benefitting our understanding of the evolution of plant chemodiversity and the development of enzyme design methods for efficient biocatalysts.

**MATERIALS AND METHODS**

**General procedures**

GC–MS analyses were performed using an Agilent GC (7890A)-MSD (5975C) instrument with an HP-5MS quartz capillary column (30 mm × 250 µm i.d., 0.25 µm film thickness). High purity helium was used as the carrier gas. For the mass spectral detector, ion source, transfer line, and quadrupole temperatures were set at 230 °C, 250 °C, and 150 °C, respectively, with electronic ionization (EI) mode at 70 eV and a scan range of m/z 40–450. Column chromatography (CC) was carried out on silica gel (200–300 mesh, Qingdao Marine Chemical Factory, China). Both analytical and preparative thin-layer chromatography (TLC and PTLC) were performed on silica gel plates (GF254, 10–40 µm, Qingdao Marine Chemical Factory). Spots on TLC plates were visualized under UV light or by heating after spraying with 5% H2SO4 in EtOH (v/v). NMR experiments were carried out on a Bruker AV-800 spectrometer with tetramethylsilane (TMS) as an internal standard. Mass spectra were obtained on a Waters AutoSpec Premier P776 spectrometer. Optical rotations were obtained on a Jasco P-1020 spectropolarimeter. GDP, GFD, and GGD were purchased from Sigma-Aldrich, and the authentic α-terpineol sample was purchased from Aladdin (Shanghai Aladdin Bio-Chem Technology). The n-hexane used for extraction was analytical grade, and PE, acetone, and ethyl acetate (EtOAc) were distilled before use.

**Plants, microbes, and vector materials**

Mature plants of C. coccinea var. mollis were grown in the Botanical Garden of Kunming Institute of Botany, Chinese Academy of Sciences (CAS). Three- to four-month-old plants for MeJA treatment were propagated from seeds collected from the above mature plants and were grown in a constant greenhouse at 22°C with a 16-h-light/8-h-dark regime. N. benthamiana seeds were kindly donated by Prof. Jinsong Wu at Kunming Institute of Botany, CAS, and were grown in the greenhouse for 5–8 weeks. The pBBA5c-MevT-MBIS plasmid used for identification of enzyme function and isolation of products was kindly provided by Prof. Tao Liu at Tianjin Institute of Industrial Biotechnology, CAS. The pEQA-HT vector used for transient expression of the gene of interest in N. benthamiana was kindly donated by Dr. Qing Zhao at Shanghai Chenshan Plant Science Research Center, CAS. Agrobacterium tumefaciens LBA4404 used for infiltration of N. benthamiana was kindly provided by Prof. Jianqiang Wu at Kunming Institute of Botany, CAS. All plasmids constructed in this study were confirmed by DNA sequencing.

**GC–MS analysis of terpenoid constituents of C. coccinea var. mollis**

For analysis of terpenoid profiles of C. coccinea var. mollis, 3 g of fresh leaves, flowers, and epidermis-free tender stems (epidermis peeled off) were immediately frozen in liquid nitrogen and ground into powder. Extraction was performed using 20 ml acetone in an ultrasonic bath for 30 min. Subsequently, the extracts were filtered and concentrated using a rotary evaporator, then extracted with n-hexane and concentrated to 1 ml. The temperature program for GC–MS analysis was as follows: initial temperature 60°C, held for 2 min; ramped up to 240°C at 5°C/min, held for 2 min; ramped up to 280°C at 10°C/min, held for 2 min.

**Total RNA isolation, de novo assembly, and gene functional annotation**

Total RNA was extracted from C. coccinea var. mollis using the TRIzol reagent (Invitrogen) according to the manufacturer’s protocol. The RNA quality was assessed using a NanoDrop 2000 instrument and gel electrophoresis (1% agarose). mRNA purified using Sera-mag magnetic oligo(dT) beads (Ambion) was sequenced on the Illumina HiSeq 2000 platform at GENE DENOVO in Guangzhou, China. Clean reads were obtained by removing redundant and low-quality sequences, and de novo transcriptome assembly was performed with a short-read assembly program (Trinity) (Grabherr et al., 2011). Sequences obtained from the Trinity assembly were called unigenes. All unigenes were annotated with four databases (Nr, Swiss-Prot, COG, and KEGG) using the BLASTx program with e-value < 10−5. Unigene expression was calculated and normalized to RPKM (Mortazavi et al., 2008).

**Cloning of the full-length cDNA of CcTPS1**

Based on the specific fragments obtained from the transcriptome data, the full-length sequence of the candidate TPS was cloned according to the manufacturer’s instructions for the SMART RACE cDNA amplification kit (Clontech). The full-length ORF of CcTPS1 was amplified by PCR using...
A promiscuous terpenoid synthase from a Lamiaceae plant

PrimeSTAR HS DNA Polymerase (Takara) and then cloned into a pCold-TF vector (Takara) containing an N-terminal six-histidine tag at the Ndel and BamHI restriction sites, resulting in the plasmid pCold-TF/CcTPS1. The recombinant plasmid was transformed into E. coli TOP10 cells for sequence verification, then transferred into E. coli strain Rosetta (DE3) for heterologous expression. All primers used are listed in Supplemental Table 6.

Real time qPCR analysis of CcTPS1 transcript levels

For tissue-specific analysis of gene expression, total RNA was extracted from the leaves, flowers, and epidermis-free tender stems of mature C. coccinea var. mollis grown in the Botanical Garden of Kunming Institute of Botany, CAS using TRizol reagent in December 2020. For MeJA treatment, 3- to 4-month-old C. coccinea var. mollis plants grown in the greenhouse were sprayed with 100 μM MeJA (Sigma-Aldrich) in 0.1% (v/v) aqueous ethanol. After treatment for 3, 6, 12, 24, 48, and 72 h, the second and third fully elongated leaves were collected for RNA extraction and metabolite analysis. The control plants were sprayed with the same amount of 0.1% (v/v) aqueous ethanol. Real time qPCR was performed using UltraSYBR mixture reagent (ROX) (TRANSge) on an Applied Biosystems 7500 instrument (Life Technologies) according to the manufacturer’s instructions. The actin gene was used as an internal standard. The relative expression levels were calculated using the ΔΔCT method. Real time qPCR experiments were carried out with three independent biological replicates, each of which was performed in three technical replicates.

GC–MS analysis of CcTPS1 products in engineered E. coli

MmGFDPs was ligated into the Ndel and Xho restriction sites of the pET-28a vector, generating the plasmid pET-28a/MmGFDPs. The recombinant pCold-TF/CcTPS1 and pET28a/MmGFDPs plasmids were co-transformed into E. coli expression strain BL21(DE3) that harbored the pBBa5c-Mev-T-MBIS vector containing the complete set of MVA pathway and FDP synthase genes (Peralta-Yahya et al., 2011). pCold-TF empty vector and pET28a/MmGFDPs were co-transformed into the same strain as a negative control. The positive transformants were screened using LB solid plates containing 100 mg/l ampicillin (AMP), 34 mg/l chloramphenicol (CM), and 50 mg/l kanamycin. After verification by PCR, the positive clones were inoculated into 500 ml Terrific Broth with the corresponding antibiotic and cultured at 37 °C in incubator shakers. When OD600 reached 0.6, 0.3 mM IPTG was added to the culture medium, and the recombinant protein was expressed at 28 °C for 3 days. The cell cultures were extracted twice with an equal volume of PE. The organic layer was concentrated using a rotary evaporator and then analyzed using GC–MS. Samples (5 μl) were injected into an Agilent GC–MS instrument at a constant flow rate of 1.8 ml/min. The temperature program for GC–MS analysis was as follows: initial temperature 80 °C; ramped up to 220 °C at 15 °C/min; ramped up to 270 °C at 4 °C/min, and held for 2 min.

Isolation and structural elucidation of CcTPS1 products from large-scale cultivation of engineered E. coli

To identify the structures of the CcTPS1 enzymatic products, 150 l of engineered E. coli cells harboring pCold-TF/CcTPS1, pET28a/MMGFDPs, and pBBa5c-Mev-T-MBIS were cultivated. The cultures were extracted with PE (3 x 150 l) at room temperature, and the organic phase was combined and concentrated in vacuo. The crude extract was subjected to CC on silica gel eluted with PE and then with PE/ethyl acetate in a stepwise-gradient system (from 100:0 to 1:1, v/v) to obtain five fractions (Fr.1–Fr.5). Fr.1 (1.2 g) was subjected to silica gel CC using 100% cyclohexane as an eluent to yield compounds 5 (400 mg), 6 (155 mg), and 7 (302 mg). Fr.2 (100 mg) was purified on a silica gel column eluted with 100% cyclohexane to obtain two sub-fractions (Fr.2-1 and Fr.2-2). Fr.2-1 was repeatedly chromatographed on a silica gel column eluted with n-hexane to yield compound 9 (10 mg). Fr.2-2 was subjected to silica gel CC and eluted with hexane to obtain compounds 1 (2 mg) and 2 (1 mg). Fr.4 was separated by gel CC eluted with PE/EtOAc (75:1, v/v) to obtain compound 4 (85 mg).

Functional identification of CcTPS1 using in vitro enzyme assays

The pCold-TF/CcTPS1 plasmid was transformed into E. coli expression strain Rosetta (DE3), and the clones were screened in an LB solid plate containing AMP (100 mg/l) and CM (34 mg/l). The positive transformants were grown to an OD600 of approximately 0.5 and were then incubated with 0.3 mM IPTG overnight at 16 °C. The cells were collected by centrifugation and disrupted by sonication in chilled extraction buffer containing 20 mM Tris (pH 7.4), 10% glycerol, 0.5 M NaCl, and 5 mM DTT. After centrifugation at 12 000 g for 10 min to remove cellular debris, the supernatant was collected. Affinity purification was carried out using Ni-NTA agarose columns (QiAGEN), and the recombinant protein was purified with elution buffers containing 10, 20, and 250 mM imidazole. The crude lysate and purified protein were then examined by SDS–PAGE. The protein concentration was determined using the Bradford method with BSA as a standard.

For the in vitro enzyme activity assays, purified CcTPS1 recombinant protein was added to the reaction buffer with a final volume of 250 μl containing 50 mM potassium phosphate (pH 7.4), 10% glycerol, 2 mM DTT, 5 mM MgCl2, and 10 μg GDP, FDP, or GGDP, then incubated at 30 °C for 3 h. Subsequently, the reaction mixture was extracted three times with n-hexane and concentrated under nitrogen gas. Negative controls included assays with affinity-purified enzyme but without prenyl diphosphate substrate and boiled recombinant protein with GDP, FDP, or GGDP as a substrate under the same reaction conditions. Six microliters of each sample was injected in splitless mode for GC–MS analysis as described in the “GC–MS analysis of terpenoid components from C. coccinea var. mollis” section.

Expression of CcTPS1 in N. benthamiana for function and subcellular localization analysis

The cDNA of CcTPS1 and AtGFDPs2 (At3g14550) were amplified and subcloned into the Cowpea mosaic virus-based (CPMV-HT) plant expression vector pEAQ-HT (Sainsbury et al., 2009) to obtain pEAQ-HT/CcTPS1 and pEAQ-HT/AtGFDPs2 using the ClonExpress II One Step Cloning Kit (Vazyme). The recombinant expression vectors were transformed into A. tumefaciens strain LBA4404 using a freeze–thaw protocol. The positive clones were cultured at 28 °C in LB liquid medium containing 25 μg/ml rifampicin, 50 μg/ml streptomycin, and 50 μg/ml kanamycin to an OD600 of 1.2–1.8. After centrifugation at 4000 g for 10 min, A. tumefaciens cells were gently resuspended in MMA buffer containing 10 mM 2-(N-morpholino) ethanesulfonic acid (MES) (pH 5.6), 10 mM MgCl2, and 100 μM acetosyringone to make a solution with a final OD600 of 0.8 for each transformant. The A. tumefaciens solutions harboring pEAQ-HT/CcTPS1 and pEAQ-HT/AtGFDPs2 plasmids were mixed in equal volumes. After incubation at room temperature for about 2 h, the prepared solution was aspirated into a sterile needle-free syringe (1 ml) to infiltrate the leaves of 5- to 8-week-old N. benthamiana. Plants infiltrated with solutions of A. tumefaciens harboring empty pEAQ-HT and pEAQ-HT/AtGFDPs2 were used as negative controls.

After growth in the greenhouse for 5 days, 9.5 g of fresh leaves were harvested from infiltrated plants and immediately frozen in liquid nitrogen. After grinding into powder, 50 ml acetone was used to extract infiltrated leaves in an ultrasonic bath for 30 min. Subsequently, the extracts were filtered and concentrated using a rotary evaporator, followed by extraction with n-hexane. The extracts were concentrated to 1 ml and then analyzed by GC–MS with the method described above in the “GC–MS analysis of CcTPS1 products in engineered E. coli” section.

For subcellular localization analysis, the full-length cDNA of CcTPS1 was ligated into the pCAMBIA1300-GFP vector. The resultant plasmid was
Plant Communications

transformed into A. tumefaciens LBA4404 and then infiltrated into N. benthamiana leaves by the method described above. Three days after infiltration, the transgenic tobacco leaves were observed and imaged under a confocal laser scanning microscope (Olympus).

Site-directed mutagenesis of CcTPS1

Site-directed mutagenesis was performed by PCR amplification using pCold-TF/CcTPS1 as the template and the mutated complementary sequences as primers. PCR amplicons were purified using a DNA quick purification kit and digested with BamHI and Ndel, then ligated into the digested pCold-TF vector. CcTPS1 variants were transformed into an E. coli expression system harboring pBBR1MCS-5-MevT-MBIS and pET-28a/+ vector. CcTPS1 variants were transformed into A. tumefaciens LBA4404 and then infiltrated into N. benthamiana leaves by the method described above. The products of the CcTPS1 variants were analyzed as described above.

Phylogenetic analysis

The amino acid sequences of plant TPSs (Supplemental Table 7) were obtained from the NCBI protein database and aligned using the Clustal X program (Larkin et al., 2007). A maximum-likelihood phylogenetic tree was constructed using MEGA X software (Kumar et al., 2018). Bootstrap values (based on 1000 replicates) >75% are shown at the corresponding nodes.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at Plant Communications Online.

FUNDING

This work was supported financially by the National Natural Science Foundation of China (21937006, 31770390, 31770340 and 31525005), the Yunnan Key Research and Development Program (no. 2019ZF011-2), the Science Foundation of Yunnan (2018FA017), and the Youth Innovation Promotion association and “Western Light” Program of the CAS (to Y. Liu).

AUTHOR CONTRIBUTIONS

D.-S.L. conducted the gene cloning, protein expression, site-directed mutagenesis, enzymatic assays, compound purification, and gene expression experiments. J.H., S.-H. Luo, Y.-G.C., and K.G. carried out the structural elucidation. Y.-C.L. and Y. Ling prepared materials for RNA sequencing analysis. S.H. Li and Y. Liu designed the research, and S.H. Li, Y. Liu, D.-S.L., and J.H. wrote the manuscript.

ACKNOWLEDGMENTS

No conflict of interest declared.

REFERENCES

Aharoni, A., Giri, A.P., Verstappen, F.W.A., Bertea, C.M., Sevenier, R., Sun, Z., Jongsma, M.A., Schwab, W., and Bouwmeester, H.J. (2004). Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. Plant Cell 16:3110–3131.

Akihisa, T., Koike, K., Kimura, Y., Sashida, N., Matsumoto, T., Ukiya, M., and Nikaido, T. (1999). Acyclic and incompletely cyclized triterpene alcohols in the seed oils of Theaceae and Gramineae. Lipids 34:1151–1157.

Armenteros, J.J.A., Salvatore, M., Emanuelli, O., Vinther, O., von Heijne, G., Elofsson, A., and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2:e201900429.

Bohmann, J., and Keeling, C.I. (2008). Terpenoid biomaterials. Plant J. 54:656–669.

Bohmann, J., Meyer-Gaumen, G., and Croteau, R. (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. U S A 95:4126–4133.

Chen, F., Tholl, D., Bohmann, J., and Pichersky, E. (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J. 66:212–229.

Chen, Q.W., Jiang, T., Liu, Y.X., Liu, H.L., Zhao, T., Liu, Z.X., Gan, X.C., Hallab, A., Wang, X.M., He, J., et al. (2019). Recently duplicated sesquiterpene (C25) gene clusters in Arabidopsis thaliana modulate root microbiota. Sci. China Life Sci. 62:947–958.

Chen, Q.W., Li, J.X., Liu, Z.X., Mitsuhashi, T., Zhang, Y.T., Liu, H.L., Ma, Y.H., He, J., Shinada, T., Sato, T., et al. (2020). Molecular basis for sesquiterpene diversity produced by plant terpene synthases. Plant Com 1:100051.

Chen, Q.W., Li, J.X., Ma, Y.H., Yuan, W.L., Zhang, P., and Wang, G.D. (2021). Occurrence and biosynthesis of plant sesquiterpenes (C25), a new addition to terpene diversity. Plant Com 2:100184.

Christianson, D.W. (2017). Structural and chemical biology of terpenoid cyclases. Chem. Rev. 117:11570–11648.

Consortium, M.E.G. (2018). Phylogenomic mining of the mints reveals multiple mechanisms contributing to the evolution of chemical diversity in Lamiaceae. Mol. Plant 11:1084–1096.

Davidovich-Rikanati, R., Lewinsohn, E., Bar, E., Iijima, Y., Pichersky, E., and Sitrit, Y. (2008). Overexpression of the lemon basil α-zingiberene synthase gene increases both mono- and sesquiterpene contents in tomato fruit. Plant J. 56:229–238.

Degenhardt, J., Kollner, T.G., and Gershenzon, J. (2009). Monoterpenes and sesquiterpenes as terpenoids and sesquiterpene synthases. The origin of terpene skeletal diversity in plants. Phytochemistry 70:1621–1637.

Dhandapani, S., Tjiang, J.G., and Jang, I.-C. (2020). Production of multiple terpenes of different chain lengths by subcellular targeting of multi-substrate terpene synthase in plants. Metab. Eng. 61:397–405.

Delay, F., and Ohloff, G. (1979). Syntheses and absolute configuration of (E)- and (Z)-bisabolones. Helv. Chim. Acta 62:369–377.

Gercke, O., Hansen, N.L., Pedersen, G.B., Kjaerulf, L., Luo, D., Staerk, D., Moller, B.L., Pateraki, I., and Heskes, A.M. (2020). Neryllyeryl diphosphate is the precursor of serrulatane, viscidane and camphene-type diterpenoids in Eremophila species. BMC Plant Biol. 20:91.

Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. Nat. Chem. Biol. 3:408–414.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., and Zeng, Q. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 29:644–652.

Green, S., Baker, E.N., and Laing, W. (2011). A non-synonymous nucleotide substitution can account for one evolutionary route to sesquiterpene synthase activity in the TPS-b subgroup. FEBS Lett. 585:1841–1846.

Green, S., Friel, E.N., Match, A., Beuning, L.L., Cooney, J.M., Rowan, D.D., and MacRae, E. (2007). Unusual features of a recombinant apple alpha-farnesene synthase. Phytochemistry 68:176–188.

Gu, H.S., Ma, S.G., Li, Y.H., Wang, Y.D., Liu, Y.B., Li, L., Li, Y., Qu, J., Lv, H.N., Chen, X.G., et al. (2014). Cloxyllones A-I, prenylbisabolane diterpenoids with anti-Coxsackie B virus activity from the branches and leaves of Claoxylon polot. Tetrahedron 70:7476–7483.
A promiscuous terpenoid synthase from a Lamiaceae plant

Guo, K., Liu, Y., and Li, S.H. (2021). The untapped potential of plant sesterterpenoids: chemistry, biological activities and biosynthesis. Nat. Prod. Rep. https://doi.org/10.1039/D1NP00021G.

Hansen, N.L., Heskes, A.M., Hamberger, B., Olsen, C.E., Hallstrom, B.M., Andersen-Ranberg, J., and Hamberger, B. (2017). The terpene synthase gene family in Tripterygium wilfordii harbors a labdane-type diterpene synthase among the monoterpene synthase TPS-b subfamily. Plant J. 89:429–441.

Herde, M., Gaertner, K., Koellner, T.G., Fode, B., Boland, W., Gershenzon, J., Gatz, C., and Tholl, D. (2008). Identification and regulation of TPS04/GES, an Arabidopsis geranyllylalcohol synthase catalyzing the first step in the formation of the insect-induced volatile C(16)-homoterpene TMTT. Plant Cell 20:1152–1168.

Huang, A.C., Hong, Y.J., Bond, A.D., Tantillo, D.J., and Osbourn, A. (2018). Diverged plant terpene synthases reroute the carbocation cyclization path towards the formation of unprecedented 6/11/5 and 6/6/7/5 sesterterpene scaffolds. Angew. Chem. Int. Ed. 57:1291–1295.

Huang, A.C., Kautsar, S.A., Hong, Y.J., Medema, M.H., and Osbourn, A. (2017). Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. Proc. Natl. Acad. Sci. U S A 114:E6005–E6014.

Huang, M., Abel, C., Sohrabi, R., Petri, J., Haupt, I., Cosimano, J., Huang, A.C., Kautsar, S.A., Hong, Y.J., Medema, M.H., and Osbourn, A. (2019). Variation of herbivore-induced volatile terpenes among Arabidopsis ecotypes depends on allelic differences and subcellular targeting of two terpene synthases, TPS02 and TPS03. Plant Physiol. 153:1293–1310.

Jing, S.X., Fu, R., Li, C.H., Zhou, T.T., Liu, Y.C., Liu, Y., Luo, S.H., Li, X.N., Zeng, F., and Li, S.H. (2021). Immunosuppressive sesterterpenoids and nosesterterpenoids from Colquohuina coccinea var. mollis. J. Org. Chem. https://doi.org/10.1021/acs.joc.1c00374.

Johnson, S.R., Bhat, W.W., Sadre, R., Miller, G.P., Garcia, A.S., and Hamberger, B. (2019). Promiscuous terpene synthases from Prunella vulgaris highlight the importance of substrate and compartment switching in terpene synthesis evolution. New Phytol. 223:323–335.

Kodama, K., Higuchi, R., Miyamoto, T., and Van Soest, R.W. (2012). (-)-Axinyssene: a novel cytotoxic diterpene from a Japanese marine sponge Axinyssa sp. Org. Lett. 14:2947–2948.

Kumar, S., Stecher, G., Li, M., Knzy, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35:1547–1549.

Landmann, C., Fink, B., Festner, M., Dregus, M., Engel, K.H., and Schwab, W. (2007). Cloning and functional characterization of three terpene synthases from lavender (Lavandula angustifolia). Arch. Biochem. Biophys. 465:417–429.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R.M., Mcgettigan, P.A., Mcwilliam, H., Valentin, F., Wallace, I.M.W., Wilm, A., and Lopez, R. (2007). Clustal W. Clustal X version 2.0. Bioinformatics 23:2947–2948.

Li, C.H., Jing, S.X., Luo, S.H., Shi, W., Hua, J., Liu, Y., Li, X.N., Schneider, B., Gershenzon, J., and Li, S.H. (2013). Peltate glandular trichomes of Colquohuina coccinea var. mollis harbor a new class of defensive sesterterpenoids. Org. Lett. 15:1694–1697.

Li, K., and Gustafson, K.R. (2021). Sesterterpenoids: chemistry, biology, and biosynthesis. Nat. Prod. Rep. 2021. https://doi.org/10.1039/D0NP00070A.

Luo, F., Ling, Y., Li, D.S., Tang, T., Liu, Y.C., Liu, Y., and Li, S.H. (2019). Characterization of a sesquiterpene cyclase from the glandular trichomes of Leucoceputrum canum for sole production of cedrol in Escherichia coli and Nicotiana benthamiana. Phytochemistry 162:121–128.

Luo, S.H., Luo, Q., Niu, X.M., Xie, M.J., Zhao, X., Schneider, B., Gershenzon, J., and Li, S.H. (2010). Glandular trichomes of Leucoceputrum canum harbor defensive sesterterpenoids. Angew. Chem. Int. Ed. 49:4471–4475.

Minami, A., Ozaki, T., Liu, C.W., and Oikawa, H. (2018). Cyclopentane-forming di/ Lesterterpene synthases: widely distributed enzymes in bacteria, fungi, and plants. Nat. Prod. Rep. 35:1330–1346.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5:621–628.

Nagel, R., Bernholz, C., Vranová, E., Kosuth, J., Bergau, N., Ludwig, W., Wessjohann, L., Gershenzon, J., Tissier, A., and Schmidt, A. (2015). Arabidopsis thaliana isoprenyl diphosphate synthase produces the C25 intermediate geranylfarnesyl diphosphate. Plant J. 84:847–859.

Pazouki, L., and Niinemets, U. (2016). Multi-substrate terpene synthases: their occurrence and physiological significance. Front. Plant Sci. 7:1019.

Peralta-Yahya, P.P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J.D., and Lee, T.S. (2011). Identification and microbial production of a terpene-based advanced biofuel. Nat. Commun. 2:483.

Pichersky, E., and Raguso, R.A. (2018). Why do plants produce so many terpenoid compounds? New Phytol. 220:692–702.

Rucker, G., Manns, D., and Wilbert, S. (1992). Hornditerpene peroxides from Artemisia absinthium. Phytochemistry 31:340–342.

Sainsbury, F., Thuenemann, E.C., and Lomonossoff, G.P. (2009). pEAG: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol. J. 7:682–693.

Scheffrahn, R.H., Gaston, L.K., Sims, J.J., and Rust, M.K. (1983). Identification of the defensive secretion from soldiers of the North American termite, Amtermes wheeleri (Desneux) (Isoptera: Termitidae). J. Chem. Ecol. 9:1293–1305.

Schreper, P., Buettner, A., Goerner, C., Hertel, M., van Rijn, J., Wallrapp, F., Eisenreich, W., Sieber, V., Kourist, R., and Bruck, T. (2016). Identification of amino acid networks governing catalysis in the closed complex of class I terpene synthases. Proc. Natl. Acad. Sci. U S A 113:E958–E967.

Shao, J., Chen, Q.W., Lv, H.J., He, J., Liu, Z.F., Lu, Y.N., Liu, H.L., Wang, G.D., and Wang, Y. (2017). (+)-Thalianatriene catalyzed by sesterterpene synthases from Arabidopsis thaliana. Org. Lett. 19:1816–1819.

Sharkey, T.D., Gray, D.W., Peli, H.K., Breneman, S.R., and Topper, L. (2013). Isoprene synthase genes form a monophyletic clade of acyclic terpene synthases in the TPS-b terpene synthase family. Evolution 67:1026–1040.

Snider, B.B., and Beal, R.B. (1988). Total synthesis of sesquiterpenes via intramolecular ketene cycloadditions. Isocoumarin and 6-cis- and 6-trans-bergamotenes. An approach to seychellene. J. Org. Chem. 53:4508–4515.

Sullivan, B.W., Faulkner, D.J., Okamoto, K.T., Chen, M.H.M., and Clardy, J. (1986). (6R,7S)-7-Amino-7,8-dihydro-6H-bisabolone, an antimicrobial metabolite from the marine sponge Halichondria sp. J. Org. Chem. 51:5134–5136.
Plant Communications

A promiscuous terpenoid synthase from a Lamiaceae plant

Yang, Y.S., Chung, I.M., kim, S.H., and Ahmad, A. (2014). New sesterterpene diolyl butanoate from rice straw of Oryza sativa. Asian J. Chem. 26:7792–7794.

Ying, Z., Zhang, L.X., Yan, Z., and Huang, G.D. (2010). Unusual sesquiterpene lactones with a new carbon skeleton and new acetylenes from Ajania przewalskii. Food Chem. 118:228–238.

Yoshikuni, Y., Ferrin, T.E., and Keasling, J.D. (2006). Designed divergent evolution of enzyme function. Nature 440:1078–1082.

Zhou, F., and Pichersky, E. (2020). The complete functional characterisation of the terpene synthase family in tomato. New Phytol. 226:1341–1360.

Ogawa, T., Yoshimura, T., and Hemmi, H. (2010). Geranylfarnesyl diphosphate synthase from Methanosarcina mazei: Different role, different evolution. Biochemical and Biophysical Research Communications 393:16–20.

Li, X.W., and Hedge, I.C. (1994). Colquhounia. In Flora of China, C.Y. Wu and P.H. Raven, eds. (Beijing, St. Louis: Science Press & Missouri Botanical Garden Press), pp. 185–187.