Plant Physiology Preview. Published on June 15, 2015, as DOI:10.1104/pp.15.00695

Running head: Functional divergence of diterpene synthases

Author for correspondence:

Name: Xiaoquan Qi

Address: Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Phone: +86 10 62836671

Fax: +86 10 82599701

E-mail: xqi@ibcas.ac.cn

Reuben J. Peters

Address: Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA, 50011, USA

Phone: +01 515-294-8580

E-mail: rjpeters@iastate.edu

Functional divergence of diterpene syntheses in the medicinal plant *Salvia miltiorrhiza* Bunge

Guanghong Cui,[a, b] Lixin Duan,[a] Baolong Jin,[b] Jun Qian,[c] Zheyong Xue,[a] Guoan Shen,[a] John Hugh Snyder,[a] Jingyuan Song,[c] Shilin Chen,[c] Luqi Huang,[b] Reuben J. Peters,[d,1] Xiaoquan Qi [a,1]

[a] Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, People's Republic of China; [b] State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, People's Republic of China; [c] Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, People's Republic of China; [d] Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA, 50011, USA

Summary

The medicinal plant *Salvia miltiorrhiza* contains the most abundant diterpenoid metabolites that normally have pharmacological activities such as vasorelaxation, against ischemia-reperfusion injury, and antiarrhythmic effects. However, it remains unclear what is the driving force behind the observed diterpene synthase functional radiation -e.g., from the gibberellin biosynthesis required in all vascular plants to the varied stereochemical outcome. In this study we report that *Salvia miltiorrhiza* has at least four diterpenoid pathways including the separate tanshinone pathways in roots and aerial tissues, as well as a novel *ent*-13-*epi*-manoyl oxide pathway in floral sepals, and positive selection driving the fast

divergence of diterpene synthase underlies the biosynthesis of specialized diterpenes in *S. miltiorrhiza*.

Footnote: This work was supported by the Key Project of Chinese National Programs for Fundamental Research and Development (2013CB127000 to X. Q.), National Science Fund for Distinguished Young Scholars (81325023 to L. H.), the National Natural Science Foundation of China (No. 81001604 to G. C.), the Independent Studies Supported by China Academy of Chinese Medical Sciences (ZZ20090301 to G. C.), and the US National Institutes of Health (NIH grant GM076324 to R.J.P.).

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Xiaoquan Qi (xqi@ibcas.ac.cn).

ABSTRACT

The medicinal plant *Salvia miltiorrhiza* produces various tanshinone diterpenoids that have pharmacological activities such as vasorelaxation, against ischemia-reperfusion injury, and antiarrhythmic effects. Their biosynthesis is initiated from the general diterpenoid precursor (*E*,*E*,*E*)-geranylgeranyl diphosphate by sequential reactions catalyzed by copalyl diphosphate synthase (CPS) and kaurene synthase-like (KSL) cyclases. Here is reported characterization of these enzymatic families from *S. miltiorrhiza*, which has led to the identification of novel pathways, including roles for separate CPSs in tanshinone production in roots versus aerial tissues (SmCPS1 and SmCPS2, respectively), as well as the novel production of *ent*-13-*epi*-manoyl oxide by SmCPS4 and SmKSL2 in floral sepals. The conserved SmCPS5 is involved in gibberellin plant hormone biosynthesis. Down-regulation of SmCPS1 by RNAi resulted in substantial reduction of tanshinones, and metabolomics analysis revealed 21 potential intermediates, indicating a complex network for tanshinone metabolism defined by certain key biosynthetic steps. Notably, the correlation between conservation pattern and stereochemical product outcome of the CPSs observed here, suggests a degree of correlation that, especially when combined with the identity of certain key residues, may be predictive. Accordingly, this study provides molecular insights into the evolutionary diversification of functional diterpenoids in plants.

INTRODUCTION

Salvia miltiorrhiza Bunge, a Lamiaceae species known as red sage or tanshen, is a traditional Chinese medicinal herb that is described in the Shen Nong Ben Cao Jing, the oldest classical Chinese herbal book, which dates from between 25 and 220 C.E. The lipophilic pigments from the reddish root and rhizome consist of abietane quinone diterpenoids (Nakao and Fukushima, 1934), largely tanshinone IIA, cryptotanshinone and tanshinone I (Zhong et al., 2009). These are highly bioactive. For example, tanshinone IIA exerts vasorelaxative activity, has antiarrhythmic effects, provides protection against ischemia-reperfusion injury (Zhou et al., 2005; Gao et al., 2008; Sun et al., 2008), and exhibits anti-cancer activities (Efferth et al., 2008; Lee et al., 2008; Wang et al., 2008; Gong et al., 2010). In addition, tanshinones have been reported to have a broad spectrum of antimicrobial activities against various plant pathogens including rice blast fungus *Magnaporthe oryzae* (Zhao et al., 2011). Although tanshinones are mainly accumulated in the roots, trace amounts of tanshinones have been detected in aerial organs as well (Hang et al., 2008).

Diterpenoid biosynthesis is initiated by diterpene synthases (diTPS), which catalyze cyclization and/or rearrangement of the general acyclic precursor (*E*,*E*,*E*)-geranylgeranyl diphosphate (GGPP) to form various hydrocarbon backbone structures that are precursors to more specific families of diterpenoids (Zi et al., 2014). Previous work has indicated that tanshinone biosynthesis is initiated by cyclization of GGPP to copalyl diphosphate (CPP) by a CPP synthase (SmCPS1), and subsequent further cyclization to the abietane miltiradiene, by a kaurene synthase-like cyclase (SmKSL1), so named for its homology to the *ent*-kaurene synthases (KSs) required for gibberellin (GA) plant hormone biosynthesis (Gao et al., 2009). Miltiradiene is a precursor to at least cryptotanshinone (Guo

et al., 2013), and RNAi knock-down of *SmCPS1* expression reduces tanshinone production, at least in hairy root cultures (Cheng et al., 2014). The identification of SmCPS1 and SmKSL1 has been followed by that of many related diTPSs from other Lamiaceae plant species (Caniard et al., 2012; Sallaud et al., 2012; Schalk et al., 2012; Brückner et al., 2014; Pateraki et al., 2014). These largely exhibit analogous activity, particularly the CPSs, which produce CPP or the stereochemically related 8α-hydroxy-labd-13*E*-en-15-yl diphosphate (LDPP) rather than the enantiomeric (*ent*) CPP relevant to GA biosynthesis.

To further investigate diterpenoid biosynthesis in *S. miltiorrhiza*, we report here a more thorough characterization of its diTPS family. A previously reported whole genome shotgun sequencing survey (Ma et al., 2012) has indicated that there are at least five CPS, although only two KS(L) genes in *S. miltiorrhiza* (Supplemental Table S1). Intriguingly, based on a combination of biochemical and genetic (RNAi gene silencing) evidence, we find that these diTPSs nevertheless account for at least four different diterpenoid biosynthetic pathways, each dependent on a unique CPS, with the KS presumably involved in GA biosynthesis seeming to be responsible for alternative diterpenoid metabolism as well. In addition, our studies clarify the evolutionary basis for the observed functional diversity, with investigation of gene structure, positive selection, molecular docking, and mutational analysis used to explore the driving force for the functional divergence of these diTPSs. Moreover, we report metabolomic analysis, also carried out with *SmCPS1* RNAi lines, which enables prediction of the downstream steps in tanshinone biosynthesis.

113

RESULTS

Molecular analysis of the *S. miltiorrhiza* diTPSs

Using the working draft genome sequences of the *S. miltiorrhiza* (Ma et al., 2012), we identified seven cDNA sequences encoding diTPSs from an inbred line (bh2-7). The deduced amino acid sequences showed that five contain the '(D,E)XDD' motif required for the protonation-initiated cyclization reactions catalyzed by CPSs (Prisic et al., 2004), and are defined here as SmCPS1 to SmCPS5, while two have the 'DDXXD' motif involved in binding the divalent magnesium ions required for heterolytic cleavage/ionization of the allylic diphosphate ester bond catalyzed by terpene synthases such as KSs (Christianson, 2006), and are defined here as SmKSL1 and SmKSL2, respectively (Supplemental Table S1).

Phylogenetic analysis revealed distinct conservation of these diTPSs. SmCPS1-3 cluster with other previously characterized CPSs from the Lamiaceae, all of which (including SmCPS1) produce CPP or the stereochemically analogous LDPP. On the other hand, SmCPS4 and 5 cluster with CPSs that have been previously shown to produce *ent*-CPP, generally for GA biosynthesis. As previously reported (Hillwig et al., 2011), SmKSL1 has undergone loss of the N-terminal γ-domain usually found in KS(L)s (although not most other terpene synthases), which has been a characteristic of the KSLs involved in more specialized diterpenoid metabolism from the Lamiaceae. By contrast, SmKSL2 retains the γ-domain, and clusters with other dicot KS(L)s, many of which have been shown to act as KSs involved in GA biosynthesis (Fig. 1).

Biochemical characterization of the *S. miltiorrhiza* diTPSs

To determine the biochemical activity of these diTPSs, *in vitro* enzyme assays were carried out

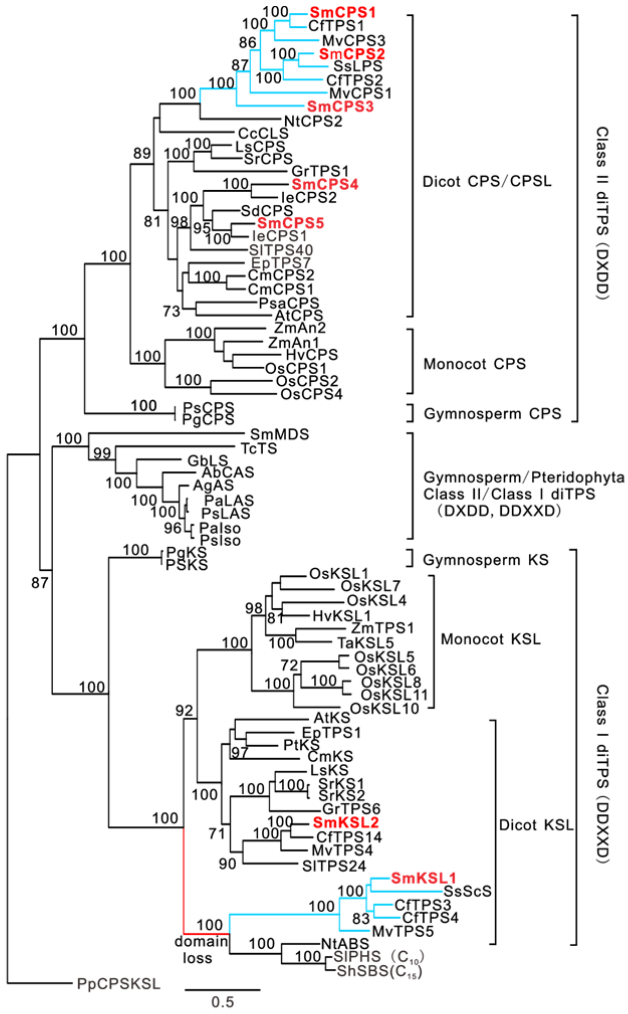


Figure 1. Phylogeny of diTPS genes in *S. miltiorrhiza*. The phylogenetic relationship was reconstructed using the JTT models by PhyML 3.0 with 68 representative characterized diTPS (Supplemental Table S2). Numbers on branches indicate the bootstrap percentage values calculated from 100 bootstrap replicates. *Physcomitrella patens* copalyl diphosphate synthase/kaurene (PpCPSKS) was used as outgroup. Blue lines show diTPS genes involved in more specialized diterpenoid metabolism from Lamiaceae. Red lines show the loss of N-terminal γ-domain found in KS(L)s. Red marked enzymes show diTPS from *S. miltiorrhiza*.

using crude extracts of recombinantly expressed proteins, separately combining each SmCPS with SmKSL1 or SmKSL2, and feeding GGPP as substrate. As expected, combining SmCPS1 and SmKSL1 led to the previously reported production of miltiradiene. Notably, combining SmCPS2 and SmKSL1 also led to production of miltiradiene (Fig. 2A and Supplemental Fig. S1A). Kinetic analysis indicates that SmCPS1 has both higher affinity (K_m = 0.54 versus 0.95 μM) and activity (>18-fold

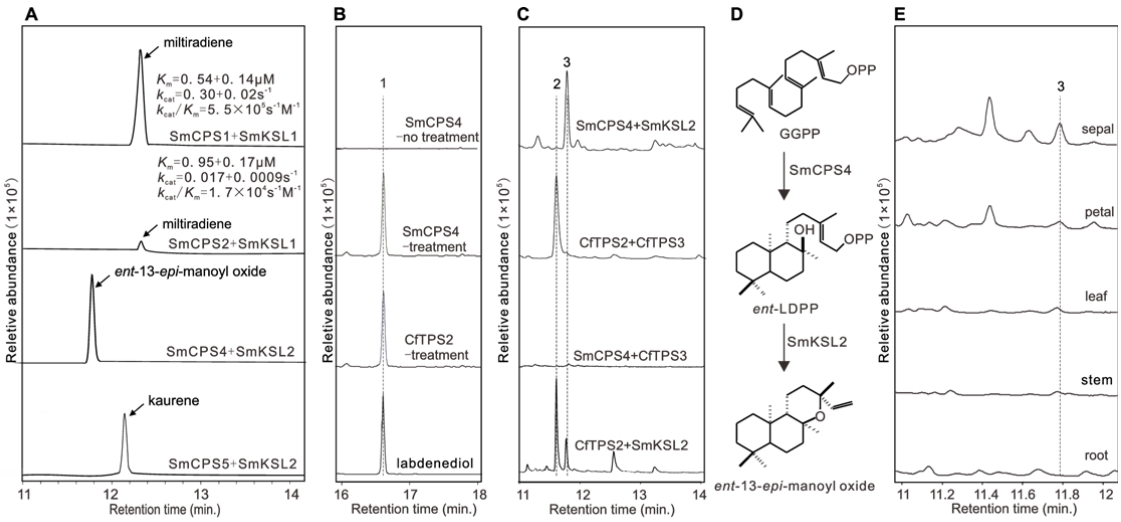


Figure 2. Functional identification of diTPS in *S. miltiorrhiza*. (A) Total ion chromatography of diterpene products from *in vitro* assays. (B) Total ion chromatography of the assay with purified SmCPS4 and CfTPS2. The assay product was extracted directly by hexane (no treatment) or after the treatment of alkaline phosphatase (treatment). (C) Total ion chromatography of the assay combines SmCPS4 with SmKSL2, together with the identified enzyme CfTPS2 (CPS) and CfTPS3 (KSL) from *Coleus forskohlii*. Assays all with 30 μM GGPP as substrate. (D) Proposed pathway to *ent*-13-*epi*-manoyl oxide in *S. miltiorrhiza*. (E) Total ion chromatography of hexane extracts from different organs in *S. miltiorrhiza*. The compounds are 1, labdenediol, 2, manoyl oxide, and 3, *ent*-13-*epi*-manoyl oxide.

higher k_{cat}) with GGPP than SmCPS2 (Fig. 2A). Nevertheless, this observation raises the potential for redundancy between SmCPS1 and SmCPS2. No product was observed when either SmCPS1 or 2 was combined with SmKSL2. No product was observed when SmCPS3 was combined with either SmKSL1 or 2. Intriguingly, combining SmCPS4 and SmKSL2 led to production of an unknown compound, while combining SmCPS5 and SmKSL2 led to the production of *ent*-kaurene, the diterpene precursor to GAs (Fig. 2A and Supplemental Fig. S1A).

To identify the compound produced by SmCPS4 and SmKSL2, the observed mass spectra was used to search a number of publically available databases, revealing a close match to that of the known diterpenoid 13-*epi*-manoyl oxide (Supplemental Fig. S1B; Demetzos et al., 2002). This presumably arises from cyclization of LDPP, suggesting that SmCPS4 produces this intermediate. However,

SmCPS4 is more closely related to *ent*-CPP producing CPSs rather than the previously characterized LDPP synthases from Lamiaceae, whose products are enantiomeric. Previous mutational analysis has shown that the CPS from *Arabidopsis thaliana* involved in GA metabolism (AtCPS), which then produces *ent*-CPP, can be easily diverted to the production of *ent*-LDPP (Potter et al., 2014). Thus, the stereochemistry of the SmCPS4 product was further investigated. Among the previously identified LDPP synthases is one from *Coleus forskohlii*, CfTPS2 (a CPS homolog), and this Lamiaceae species also encodes a subsequently acting KSL, CfTPS3, that reacts with LDPP to produce manoyl oxide (Pateraki et al., 2014). Both SmCPS4 and CfTPS2 reacted with GGPP to produce LDPP, detected as labdenediol by GC-MS following dephosphorylation (Fig. 2B and Supplemental Fig. S1B). To determine if these were enantiomeric, CfTPS2 and SmCPS4 were separately incubated with either SmKSL2 or CfTPS3, respectively, and GGPP as substrate. Notably, combining SmCPS4 with CfTPS3 did not lead to the production of manoyl oxide (Fig. 2C). In addition, combining CfTPS2 with SmKSL2 led to predominant production of manoyl oxide with relatively less 13-*epi*-manoyl oxide (Fig. 2C). These results indicate that SmCPS4 produces *ent*-LDPP (8β-hydroxy-*ent*-CPP). Therefore, the product of SmCPS4 and SmKSL2 appears to be *ent*-13-*epi*-manoyl oxide (Fig. 2D).

Physiological roles of *S. miltiorrhiza* diTPSs

The tissue specific expression pattern of the *S. miltiorrhiza* diTPSs was investigated by quantitative reverse transcription PCR (qRT-PCR) analysis of various organs, including leaves and roots of 3-day-old seedlings, and the root periderm, cortex and xylem, as well as stem, leaf, sepal, petal, stamen, pistil, and immature seeds of adult plants at the flowering stage (Fig. 3). *SmCPS1* and *SmKSL1* exhibit closely coordinated expression, and their transcripts were found at extremely high

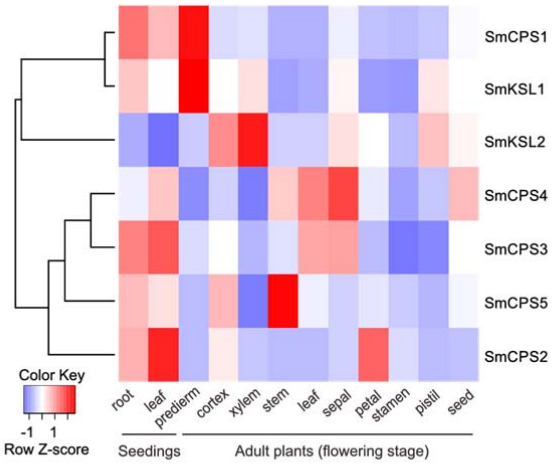


Figure 3. qRT-PCR analysis of transcript levels of seven diTPS genes in 12 organs including leaves and roots of 3-day-old seedlings, and the root periderm, cortex and xylem, as well as stem, leaf, sepal, petal, stamen, pistil, and immature seeds of adult plants at the flowering stage. The expression level was normalized to that of Actin. Data are means from three technical replicates of at least 3 biological replicates.

levels in the root periderm, consistent with their role in biosynthesis of the tanshinone pigments, which accumulate in this tissue. *SmCPS2* and *SmCPS3* were most highly expressed in seedling leaves, although *SmCPS2* also is highly expressed in the petals of adult plants. *SmCPS4* was most highly expressed in the sepal and *SmCPS5* in the stem. *SmKSL2* was most highly expressed in the root xylem, although this was expressed throughout all tissues of adult plants to some extent.

To investigate the diterpenoid natural product arsenal of *S. miltiorrhiza*, both untargeted and targeted metabolomics analyses were carried out using LC-qTOF-MS and GC-QqQ-MS, respectively. Notably, although *S. miltiorrhiza* has not been previously reported to produce (*ent*-)13-*epi*-manoyl oxide, this diterpenoid was found here, mainly in sepals, with lower amounts found in petals, leaves and stems, but not in roots (Fig. 2E; Supplemental Fig. S1B). This tissue specific accumulation pattern matches *SmCPS4* mRNA expression with a correlation coefficient of 0.97, indicating that the production of *ent*-13-*epi*-manoyl oxide in *S. miltiorrhiza* depends on expression of *SmCPS4* (as well as the more ubiquitous expression of *SmKSL2*).

As previously reported (Hang et al., 2008), tanshinones were found not only in the root and rhizome, but also aerial tissues of *S. miltiorrhiza*. Given the distinct tissue specific expression patterns of SmCPS1 and SmCPS2, both of which can produce the relevant CPP intermediate, it seemed possible that these are not redundant, but instead contribute separately to tanshinone biosynthesis in the roots and aerial tissues, respectively. Moreover, as the only identified *ent*-CPP synthase it also seemed likely that SmCPS5 is involved in GA metabolism.

Genetic evidence for distinct diterpenoid pathways in *S. miltiorrhiza*

To further investigate the distinct roles of the various SmCPSs in *S. miltiorrhiza* diterpenoid biosynthesis, RNAi gene silencing was carried out targeting either *SmCPS1* or *SmCPS5*, separately. RNAi knock-down of *SmCPS5* (Supplemental Fig. S2A) resulted in dwarf transgenic plants with significantly shorter pinnately compound leaves, shorter and narrower top leaves, and smaller flowers as compared to wild type (WT) plants and could be rescued (i.e., normal growth restored) by applying GA_3 to the T_0 generation plants (Supplemental Fig. S2 B-D). Further, T_1 generation plants showed a 3:1 segregation ratio for the dwarf phenotype with severe stunting of shoots as well as dark green leaves (Fig. S2E). These phenotypes are associated with GA deficiency (Margis-Pinheiro et al., 2005), indicate that SmCPS5, presumably together with SmKSL2, function in the GA biosynthetic pathway in *S. miltiorrhiza*.

To investigate the potential redundancy between *SmCPS1* and *SmCPS2*, RNAi targeting *SmCPS1* was carried out, generating five lines exhibiting silencing of ~90% (Fig. 4A). These *SmCPS1*-RNAi plants exhibited an obvious white color root phenotype in comparison to the WT roots, which had the characteristic reddish color associated with tanshinones (Fig. 4B). Root directed metabolite analysis

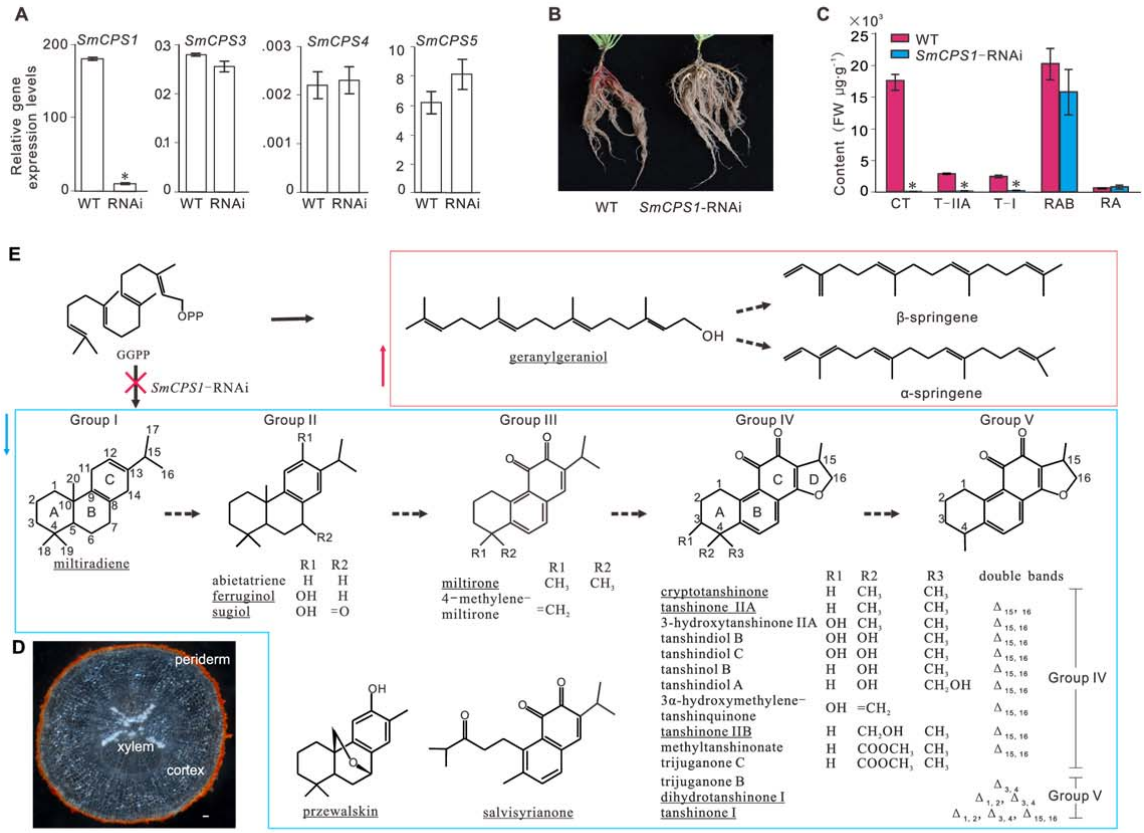


Figure 4. Phenotype and metabolic profiles caused by down regulation of *SmCPS1* in T_0 generation plants. (A) qRT-PCR analysis of transcript levels of five *CPS* genes in root of *SmCPS1*-RNAi and WT. Expression was normalized to that of *Actin*. The error bars show the SDs from mean value; $n = 3$ experiments. *SmCPS2* is not expressed in these samples. (B) The phenotype of down regulation of *SmCPS1*. (C) Quantitative analysis of five major compounds including cryptotanshinone (CT), tanshinone IIA (T-IIA), tanshinone I (T-I) and lithospermic acid B (LAB), rosmarinic acid (RA) in root of *SmCPS1*-RNAi lines and WT. The error bars show the SDs from mean value; $n = 3$ experiments. Asterisks indicate significant difference at $P < 0.01$ compared with WT by Student's t test. (D) Cross-section of the root. Bar = 100μm. (E) Summary of metabolite flux caused by down regulation of *SmCPS1*. Red arrow showed the accumulated metabolites and blue arrow showed the reduced metabolites in root of *SmCPS1*-RNAi lines compared with WT. Underlined metabolites are identified by standard reference.

208 demonstrated that cryptotanshinone, tanshinone IIA, and tanshinone I were dramatically reduced, from

209 14465 ± 995, 2311 ± 13, and 2010 ± 1170 ug/g in WT plants to 0.10 ± 0.01, 1.2 ± 0.2, and 14 ± 1 μg/g

210 in the *SmCPS1*-RNAi plants, respectively (Fig. 4C). As a control, the water-soluble polyphenolic acid

211 content of the plants, including lithospermic acid B and rosmarinic acid, also was investigated. The

212 lack of any statistically significant differences between the WT and *SmCPS1*-RNAi samples (Fig. 4C)
213 indicates that down-regulation of *SmCPS1* expression specifically affects tanshinone biosynthesis (i.e.,
214 this did not interfere with the production of at least these polyphenolic acid metabolites).

215 Notably, tanshinone levels in the aerial tissues were not affected in the *SmCPS1*-RNAi plants
216 (Table 1). Given that *SmCPS2* is predominantly expressed in these tissues (Fig. 3), these results
217 suggest that SmCPS2 mediates tanshinone biosynthesis in aerial organs of *S. miltiorrhiza*
218 independently of the SmCPS1 dependent pathway in the root periderm, such that *SmCPS1* and
219 *SmCPS2* are not redundant.

220

221 **Metabolomic analysis clarifies tanshinone biosynthesis**

222 Metabolomics analysis was carried with not only WT but also *SmCPS1*-RNAi plant roots, with
223 UPLC-ESI-qTOF-MS revealing 39, and GC-EI-QqQ-MS 19, metabolites with significantly reduced
224 accumulation (i.e., a > 2 fold change with p value < 0.05). On the other hand, 4 metabolites (by using
225 GC-EI-QqQ-MS) with elevated accumulation also were identified (Supplemental Table S3 and S4).
226 By comparison to known compounds, 21 of the metabolites with reduced accumulation in the roots of
227 *SmCPS1*-RNAi plants were identified as diterpenoids, all of which were predominantly accumulated
228 (>94.5%) in the periderm (Fig. 4D, Supplemental Table S3 and S4). Of the metabolites exhibiting
229 increased accumulation, three were identified as diterpenoids (geranylgeraniol, α-springene and
230 β-springene), and these were not detected in the periderm, cortex and xylem of wild type plant roots
231 (Supplemental Table S3 and S4). These compounds presumably appear due to the down-regulation of
232 *SmCPS1*, with accumulation of GGPP leading to hydrolysis to geranylgeraniol, with subsequent
233 dehydration leading to α-springene and β-springene (Fig. 4E).

234 Of the 21 identified SmCPS1 dependent metabolites, 19 are tanshinones or plausible biosynthetic
235 intermediates, while 2 are rearranged abietane diterpenoids – i.e., przewalskin and salvisyrianone (Fig.
236 4E). The 19 tanshinones and biosynthetically relevant metabolites could be further divided into five
237 main groups according to the progressive modification of their carbon skeletons (Fig. 4E). Group I is
238 simply composed of miltiradiene, with its planar cyclohexan-1,4-diene 'C' ring. Group II compounds
239 are dehydro-abietanes, with a characteristic aromatic 'C' ring – i.e., abietatriene, ferruginol and sugiol.
240 Group III compounds are nor-abietatetraen-11,12-diones, with conversion of the 'C' ring to an
241 ortho-quinone (i.e., keto groups at C-11 and C-12), as well as aromatization of the 'B' ring, along with
242 loss of C-20 – e.g., miltirone and 4-methylene-miltirone. The 11 metabolites comprising group IV all
243 are variously named tanshinones, with addition of the (dihydro)furan ring 'D'. Group V contains 3
244 metabolites that have been additionally modified by the loss of one of the geminal methyl groups from
245 the C-4 position, along with the presence of at least one double bond in the 'A' ring.
246
247 **Positive selection for divergent CPS activity**
248 Given the ease with which diTPSs can be diverted to alternative activity (Wilderman and Peters,
249 2007; Xu et al., 2007; Keeling et al., 2008; Morrone et al., 2008; Criswell et al., 2012; Potter et al.,
250 2014), it is not clear what underlies the expanded nature and divergent activity observed in the *S.*
251 *miltiorrhiza* diTPS family (i.e., selective pressure or genetic drift). Gene structure, specifically the
252 number and placement of introns, has been associated with evolutionary descent in TPSs (Trapp and
253 Croteau, 2001). Accordingly, the *CPSs* and *KSLs* of *S. miltiorrhiza*, Arabidopsis and rice can be
254 divided into three groups. Group I genes associated with GA metabolism, each have the typical 15
255 exons and 14 introns for *CPS* and 14 exons and 13 introns for the *KS* genes (Supplemental Fig. S3),

256 which corresponds to the ancestral plant diTPS gene structure (Trapp and Croteau, 2001). Group II
257 *CPS/KSL* genes, including *SmCPS1*, *SmCPS3, SmCPS4*, *SmKSL1, OsCPS2*, *OsCPS4*, *OsKSL5*,
258 *OsKSL6*, *OsKSL8*, *OsKSL10,* show diverged sequences and genomic architecture. In particular, intron
259 loss (Zhang et al., 2014) relative to the conserved Group I genes are often observed. For example,
260 *SmCPS1* has lost the 10th and 12th introns, *SmCPS4* the 5th intron, and *SmCPS3* the first four introns,
261 while *OsCPS2* has lost the 2nd and 3rd introns, and *OsKSL5*, *6*, *8*, and *10* all have lost the last intron
262 (Supplemental Fig. S3). These architecturally divergent group II genes are involved in the biosynthesis
263 of specialized metabolites in *S. miltiorrhiza* and rice. On the other hand, the Group III genes, *SmCPS2*,
264 *OsKSL4* and *OsKSL7* exhibit conserved genomic architecture, but divergent sequences and functions
265 relative to the Group I genes associated with GA biosynthesis (Supplemental Fig. S3).
266 In order to investigate whether the observed functional divergence of the CPS genes involved in
267 biosynthesis of the tanshinones and other more specialized labdane-related diterpenoids is a function
268 of positive selection, phylogenetic analysis of the protein-coding DNA sequences of CPS members
269 from an array of angiosperms was carried out. The resulting phylogenetic tree (Fig. 5A), provided the
270 basis for branch-site model analysis conducted with the PAML package (Zhang, 2004; Zhang et al.,
271 2005; Yang and Reis, 2011). This revealed a statistically significant signature for positive selection
272 associated with functional divergence from the production of *ent*-CPP required for GA biosynthesis to
273 the enantiomeric (i.e., normal) CPP involved in production of the tanshinones, as well as
274 stereochemically related LDPP for other more specialized diterpenoid biosynthesis (branch c, d_N/d_S (ω)
275 = 5.64, P<0.01) (Fig. 5A and B; Supplemental Table S5).
276 To more closely examine the basis for the observed functional diversification, the analysis was
277 extended to individual codons in SmCPS1 versus SmCPS5. The positively selected sites identified by

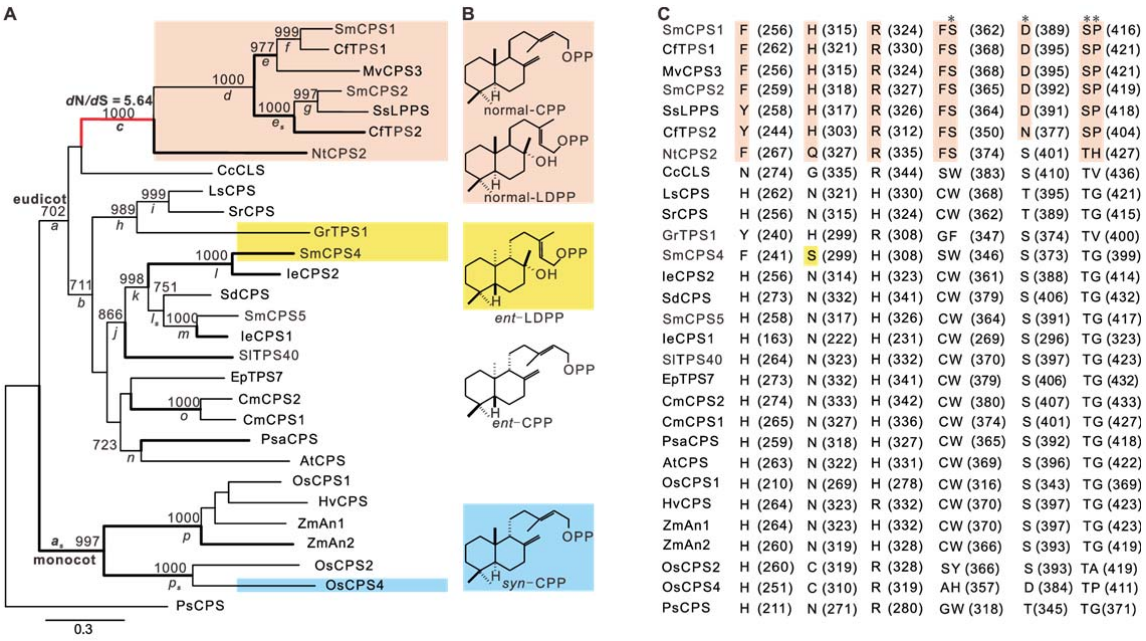


Figure 5. Molecular evolution of CPS genes. (A) Bold lines illustrate branches or genes evolved under positive selection with significant statistical support at P<0.05 both in branch-site model test 1 and test2. (B) Different stereoisomers of CPP and LDPP. (C) Alignment of conserved residues. The asterisks indicate positive selection sites with posterior probability >95% by Bayes Empirical Bayes analysis.

Bayes Empirical Bayes analysis correspond to Ser362, Asp389, Ser415 and Pro416 in SmCPS1, which are Trp364, Ser391, Thr416 and Gly417 in SmCPS5 (Fig. 5C). Models for both SmCPS1 and SmCPS5 were generated, and revealed that three of these residues, Ser362:Trp364 (S-W) and Ser415/Pro416:Thr416/Gly417 (SP-TG) are part of the active site cavity, whereas Asp389:Ser391 (D-S) is more than 6Å away from the active site cavity (Supplemental Fig. S4 A and B). Site-directed mutagenesis was performed to swap these residues between SmCPS1 and SmCPS5 except Thr416 in SmCPS5, as this corresponds to Thr421 in AtCPS, which has already been reported to be involved in catalysis (Köksal et al., 2014). Assays using the mutant enzymes reacting with GGPP alone or in combination with SmKSL1 or SmKSL2 showed that the mutant SmCPS1:S362W, in which Ser (S) 362 of SmCPS1 was replaced by the Trp (W) found at the same position in SmCPS5, was >100-fold lower efficient in the production of CPP, and miltiradiene when assayed with SmKSL1, relative to the

289 wild-type SmCPS1. Moreover, this mutant also did not yield any detectable product when incubated
290 with SmKSL2 and GGPP (Supplemental Fig. S5), indicating that no *ent*-CPP is produced. The mutant
291 SmCPS5:W364S also showed about 80-fold lower efficiency in production of *ent*-CPP, and
292 *ent*-kaurene when assayed with SmKSL2, than wild-type SmCPS5. Similarly, this mutant also did not
293 yield any product when incubated with SmKSL1. The other four mutant enzymes SmCPS1:D389S,
294 SmCPS1:P416G, SmCPS5:S391D and SmCPS5:G417P did not cause significant change in enzymatic
295 activity relative to the parental/wild-type CPSs (Supplemental Fig. S5). Thus, while no change in
296 product stereochemistry was observed, it does seem that the residue corresponding to the S-W position
297 is important for catalysis in both SmCPS1 and SmCPS5 (Supplemental Fig. S4 C and D), despite their
298 difference in product outcome, suggesting that this substitution have a role in enabling the production
299 of normal versus *ent*-CPP and, hence, biosynthesis of the derived tanshinones.
300
301

302

DISCUSSION

The combined biochemical and genetic work reported here has defined the roles of the diTPS family in *S. miltiorrhiza*. Of the five SmCPSs, while SmCPS3 appears to be inactive, each of the other SmCPSs defines separate diterpenoid pathways. Rather than reflecting redundancy, the similar biochemical activity of SmCPS1 and SmCPS2 is coupled to their distinct roles in tanshinone biosynthesis in the roots versus aerial tissues, respectively. This discovery provides the possibility of using metabolic engineering strategies to enhance the production of tanshinones in aerial organs. Cultivation of the resulting plants could then provide annually renewable source materials for extraction of tanshinones by harvesting aerial tissues without destroying the entire plant.

Both SmCPS1 and SmCPS2 react with GGPP to form normal CPP in *S. miltiorrhiza*. Our phylogenic analysis indicates that SmCPS1 and SmCPS2 are indeed in the same clade that also contains other CPSs from the Lamiaceae and closely related Solanaceae (Fig. 1) that similarly produce CPP and LDPP with analogous stereochemistry. These CPSs were likely derived from an ancestral CPS that produced *ent*-CPP for gibberellin biosynthesis via early gene duplication and neo-functionalization that occurred at least before the divergence of the Lamiaceae and Solanaceae (Fig. 1 and 5), which has been estimated to have been 64 million years ago (48 - 75Ma) (Zhang et al., 2012). During the evolution of Lamiaceae, two more gene duplication events likely happened, producing SmCPS3, as well as SmCPS1 and SmCPS2, which individually are representative of two more widespread clades (e.g., homologs to both are found in *C. forskohlii*), respectively. SmCPS2 retained the ancestral gene architecture (Fig. S3) and exhibits a similar gene expression pattern as the SmCPS5 involved in GA biosynthesis (Fig. 3), whereas SmCPS1 has undergone more divergence, including intron loss (Fig. S3) and altered transcriptional regulation (Fig. 3), along with exhibiting

significantly higher catalytic activity than SmCPS2. These data suggest that tanshinone biosynthesis may have evolved early in the Lamiaceae, and the gene duplication leading to SmCPS1 and SmCPS2 enabled further localization and upregulation of tanshinone biosynthesis in root periderm cells in some *Salvia* species, including *S. miltiorrhiza*. The presence of distinct SmCPS1- and SmCPS2- dependent tanshinone pathways in the root periderm and aerial tissues, respectively, indicates that these labdane-related diterpenoids may play important roles in plant development and in adaptation to different stress conditions, which is a topic worth future investigation.

The ability of SmCPS4 to produce the enantiomeric form of LDPP was suggested by the ability of SmKSL2 to react with both *ent*-CPP (to produce *ent*-kaurene) and this *ent*-LDPP, but not the CPP product of SmCPS1 (or SmCPS2). The observed production of *ent*-13-*epi*-manoyl oxide in *S. miltiorrhiza* then suggests dual function for SmKSL2 – i.e., in gibberellin and this more specialized diterpenoid biosynthesis. Regardless, identification of the ability of SmCPS4 and SmKLS2 provides access to this novel diterpenoid. In addition, the ability of SmCPS4 to produce *ent*-LDPP was presaged by its phylogenetic relationship to *ent*-CPP producing CPSs. Based on the similar clustering of a previously identified LDPP synthase from *Grindelia robusta* (Zerbe et al., 2013), as well as ability of a KS to selectively react with its product, it is suggested here that this GrTPS1 may produce *ent*-LDPP (Fig. 2D).

It is interesting to note that certain previously identified residues are consistent with the CPS enzymatic activity observed here. SmCPS5, shown here to play a role in gibberellin biosynthesis, contains the histidine (His326) associated with susceptibility to inhibition by Mg^{2+} (Mann et al., 2010), which has been suggested to serve a regulatory role in such CPSs (Prisic et al., 2007). In addition, SmCPS5 contains the histidine-asparagine dyad (His258/Asn317) (Fig. 5C) that has been suggested to

act as the catalytic base and is conserved in CPSs that produce *ent*-CPP (Potter et al., 2014). By contrast, the *ent*-LDPP producing SmCPS4 contains a serine (Ser299) in place of the corresponding asparagine (Fig. 5C), consistent with the ability of such substitution of smaller residues to enable production of hydroxylated CPP (Potter et al., 2014). In addition, the investigation of residues showing signs of positive selection here indicates other positions important for altering stereochemical product outcome, although further experiments are required to determine how many and which residues are required. Nevertheless, altogether the results reported here suggest that combining phylogenetic relationship with the identity of residues at positions of known catalytic relevance may be predictive for CPS catalytic activity.

MATERIALS AND METHODS

Plant Materials

The *S. miltiorrhiza* species has two different flower colors: purple and white. Varieties with purple flowers are distributed throughout China. The variety with white flowers (*S. miltiorrhiza* f. *alba*.) is only found in Shandong Province. *S. miltiorrhiza* f. *alba* is a rare and thus more valuable variety of Danshen. The white flower line bh2-7, inbred for five generations, was used in our study.

Plant Growth and Culture Conditions

bh2-7 seeds were surface-sterilized with 5% sodium hypochlorite and cultured on solid hormone-free Murashige and Skoog (MS) basal medium containing 30 $g \cdot L^{-1}$ sucrose and 8 $g \cdot L^{-1}$ agar. Cultures were maintained at 25°C under a 16 h light/8 h dark photoperiod. Seedlings were transferred to pots filled with 3:1 soil:vermiculite mix, and grown under the same temperature and light regime in a plant growth room. Dwarf plants of *SmCPS5*-RNAi were sprayed with 150 μM GA_3 solution to

complement the phenotype in T_0 generation.

Genomic Sequence of DiTPS Gene Family Members

In order to show the structural divergence occurred between *SmCPS1*, *SmCPS3,* and *SmCPS4*, we cloned the full genomic sequence of each gene from bh2-7. CTAB method was used to extract the genomic DNA and amplified with specific primers (Supplemental Table S6). Intron/exon structures were predicted using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/chinese.php) (Guo et al., 2007).

Gene Expression Analysis

Plant samples were harvested and immediately frozen in liquid nitrogen. Total RNA was extracted using a modified CTAB protocol and treated with RNase-free DNase I (Takara) to remove residual genomic DNA. RNA integrity and quality were checked by denaturing gel electrophoresis, and the absence of genomic DNA was confirmed by PCR using primers for *Smactin,* prior to reverse transcription. 1~5 μg of total RNA was reverse transcribed into cDNA using the SuperScript Ⅲ reverse transcriptase and oligo $(dT)_{12\text{-}18}$ primer (Invitrogen), according to the manufacturer's instructions. The synthesized cDNA was then diluted 10-fold. 1 μL of this diluted template was used for subsequent quantitative reverse transcription PCR (qRT-PCR) analysis with a total PCR reaction volume of 20 μL. qRT-PCR was performed with a SuperReal PreMix for SYBR Green kit (TIANGEN) on a Corvett Rotor-Gene 3000 real-time PCR detection system. All reactions were performed using the following PCR conditions: initial denaturation step of 95°C for 10 min, followed by 40 cycles each of 95°C for 5 s, 60°C for 15 s and 72°C for 20 s, with a final melting stage from 55°C to 95°C. A final dissociation step was performed to assess the quality of the amplified product. cDNA from a series of 5-fold dilutions were used for calibration, and the efficiency of the PCR amplifications was found to

391 be in the range of 90 % - 110 %, which is considered desirable for qPCR (Taylor et al., 2010). Relative
392 expression levels were calculated as the ratio of the target gene transcript level to the transcript level
393 of the housekeeping gene *Actin* (*Smactin*) (Yang et al., 2010). Primer specificity was confirmed by
394 direct cloning and sequencing of individual PCR amplification products. qRT-PCR was performed
395 with three technical replicates of at least 3 biological replicates for each tissue or transformed plant
396 line. The hot map of gene expression data was generated in *R* software package.

397 **Phylogenetic Analysis**

398 Sixty eight diterpene synthases with characterized functions were included in the phylogenetic
399 analysis (Supplemental Table 2). To carry out the phylogenic reconstruction, multiple protein
400 sequence alignments were performed with MAFFT v7.012 employing the E-INS-I method (Katoh et
401 al., 2005). Maximum likelihood (ML) trees were built using PhyML v3.0 (Guindon et al., 2010).
402 Specifically, PhyML analyses were conducted with the JTT substitution model, four rate substitution
403 categories, and 100 or 1000 bootstrap replicate analyses. The phylogeny was displayed using FigTree
404 software (http://tree.bio.ed.ac.uk/software/figtree/).

405 **Hairy root and Plant Transformation for Knockdown of *SmCPS1* and *SmCPS5* by RNA**
406 **Interference**

407 A 289 bp gene-specific sequence including the 3'-UTR of *SmCPS1* and a 399 bp gene-specific
408 sequence in the 3'- end of *SmCPS5* were amplified by PCR using cDNA as a template and then cloned
409 using Gateway technology into the pK7GWIWG (II) binary vector (Limpenset al., 2005). Positive
410 plasmids of pK7GWIWG-*CPS1* and pK7GWIWG-*CPS5* were identified with sequencing and
411 restriction enzyme analysis, and then introduced into *Agrobacterium tumefaciens* strain EHA105 and
412 *Agrobacterium rhizogens* strain ACCC 10060 by electroporation. Transformation of leaf explants from

413 *S. miltiorrhiza* bh2-7 plants was carried out following previously described methods (Yan et al., 2007)
414 with minor modifications. Single colonies of *Agrobacterium tumefaciens* strain EHA105 cells
415 harboring the various RNAi vectors were inoculated into 10 mL of liquid LB medium with 50 $mg \cdot L^{-1}$
416 spectinomycin and 100 $mg \cdot L^{-1}$ rifampicin, and then grown on a shaker (180 rpm) at 28 °C for 16-18 h.
417 Cells were collected by centrifugation when the OD_{600} reached 0.6, and were re-suspended in 20 mL
418 of liquid MS medium. Leaves or petioles were cut into 0.5 × 0.5 cm pieces discs and pre-cultured for 2
419 days on MS basal medium supplemented with 2.0 $mg \cdot L^{-1}$ 6-BA. The discs were then submerged with
420 shaking in a bacterial suspension for 15 min and co-cultured on the MS basal medium for 2 days. The
421 leaf discs were then transferred to selection MS basal medium supplemented with 2.0 $mg \cdot L^{-1}$ 6-BA, 50
422 $mg \cdot L^{-1}$ kanamycin, and 225 $mg \cdot L^{-1}$ timentin. After 2~3 rounds of selection (10 days each), the
423 regenerated buds with expression of GFP were transferred to MS basal medium supplemented with 25
424 $mg \cdot L^{-1}$ kanamycin for root formation and elongation. Rooted plantlets were further cultured on MS
425 basal medium for about one month. The plantlets (7-8 cm tall, with roots 5-6 cm long) were
426 transplanted to soil and vermiculite (3:1) and covered by beakers to maintain humidity for 1 week,
427 then gradually hardened off in pots in a greenhouse for further growth.

428 Hairy root cultures can be successfully obtained with *S. miltiorrhiza*, and we used both hairy root
429 cultures and full plants to analyze the silencing effect of S*mCPS5*. Though hairy root cultures with
430 silencing of *SmCPS5* showed no significant phenotype, we found that it was an easy and fast system to
431 analysis the effect of different silencing vectors. The *Agrobacterium rhizogens*-based transformation
432 had similar procedure as the *Agrobacterium tumefaciens*-based transformation described above. When
433 the hairy roots were 2~3 cm in length, expression of green fluorescent protein (GFP) was observed
434 under a fluorescence microscope to identify positive lines. The positive hairy roots were excised and

435 cultured on solid, hormone-free MS basal medium containing 50 mg·L^{1} kanamycin and reduced
436 timentin from 225 mg·L^{-1} to zero in two or three selection cycles (15 days each). The rapidly growing
437 kanamycin-resistant and GFP visible lines with no bacterial contamination were then maintained at
438 25 °C in the dark with MS medium without ammonium nitrate and routinely sub-cultured every 25-30
439 days.

440 **Metabolomics Profiling using LC-qTOF-MS and GC-QqQ-MS**

441 The metabolomics profiling data was acquired using a combination of two independent analytical
442 platforms. LC-qTOF-MS analysis of methanol extracts was used for global unbiased metabolite
443 detection. GC-QqQ-MS analysis of hexane extracts, optimized for detection of targeted intermediates,
444 was used for the analysis of metabolites that were not readily detectable with the LC-qTOF-MS
445 platform.

446 For LC-qTOF-MS, fresh plant root samples were frozen in liquid nitrogen and ground to a fine
447 powder under continuous cooling. 100 mg fresh weight of the powder was extracted in 2 mL of
448 methanol, containing an internal standard (umbelliferone 20 µg·mL^{-1}). The extracts were sonicated
449 twice for 15 min, centrifuged (1500 x g) for 10 min, and then filtered through a 0.2 µm PTFE syringe
450 filter (Agilent). An aliquot of each filtrate (5 µL) was separated using an Agilent 1290 Infinity UHPLC
451 system consisting of a binary pump, an autosampler, a column temperature controller, and a VWD
452 detector at 285 nm. The chromatography was performed using a ZORBAX RRHD SB-C18 column
453 from Agilent Technologies (2.1 × 100 mm, 1.8 µm). The mobile phase consisted of (A) 0.01 % formic
454 acid in acetonitrile and (B) water containing 0.01 % (v/v) formic acid. A gradient program was used as
455 follows: linear gradient from 10 to 20 % A (0-5 min), linear gradient from 20 % to 40 % A (5-7 min),
456 linear gradient from 40 to 100 % A (7-10 min), isocratic at 100 % A (10-14 min), and linear gradient

457 from 100 to 10 % A (14-15 min). The mobile phase flow rate was 0.25 $mL \cdot min^{-1}$, and the column
458 temperature was set at 30 °C. Five reference standard compounds, namely salvianolic B, rosmarinic
459 acid, crypotanshinone, tanshinone IIA, and tanshinone I were dissolved in methanol to create standard
460 curves for use in absolute quantification calculations. In order to ensure that analytes were in the linear
461 range, and in order to exclude some artifactual peaks, samples were also diluted by 10-fold and
462 50-fold and re-analyzed.

463 Mass spectrometry was performed using an Agilent 6540 q-TOF equipped with an electrospray
464 ionization (ESI) source operating in positive ion mode. The nebulization gas was set to 40 psi. The
465 drying gas was set to 10 $L \cdot min^{-1}$ at a temperature of 350 °C; the sheath gas was set to 11 $L \cdot min^{-1}$ at a
466 temperature of 350 °C. The capillary voltage was set to 4000 V. The q-TOF acquisition rate was set to
467 0.5 s. For full-scan MS analysis, the spectra were recorded in the range of m/z 100-1000.
468 Chromatographic separation, followed by full-scan mass spectra, was performed to record retention
469 time and m/z values of all detectable ions present in the samples.

470 For GC-QqQ-MS, plant material was lyophilized for 48 h. 100 mg of the lyophilized powder was
471 extracted in 2 mL of hexane that contained two internal standards (tricosane 6.5 $ug \cdot mg^{-1}$ and
472 tetracosane 0.8 $ug \cdot mL^{-1}$). The extracts were sonicated twice for 15 min and centrifuged (3000 × g) for
473 10 min. The supernatant was evaporated under nitrogen, re-suspended in 50 µL of hexanes or
474 derivatized with 80 µL N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) and 8 µL pyridine at
475 80 °C for 40 minutes, and then analyzed by GC-MS.

476 GC-MS analyses were performed on an Agilent 7890A GC system connected to an Agilent
477 7000B triple quadruple MS with electron impact ionization. A 1 µL portion of the extract was injected
478 in splitless mode onto the column. The column used was a DB-5ms (30 m x 0.25 mm i.d., 0.25 - µm

479 film thickness, Agilent J&W ScientiWc, USA) fused silica capillary column. Helium was used as the
480 carrier gas for GC at a flow rate of 1.0 $mL \cdot min^{-1}$. The injector temperature was 280 °C. The oven
481 program was as follows: 50 °C for 2 min, linear ramp at a rate of 20 $°C \cdot min^{-1}$ to 200 °C, then followed
482 with a linear ramp at a rate of 5 $°C \cdot min^{-1}$ to 300 °C, held at 300 °C for 10 min. The transfer line
483 temperature was 280 °C.

484 Raw data were processed with Mass hunter Qualitative Analysis software (Agilent). Mass
485 Profiler Professional (MPP, Agilent) software was used to identify significantly different ion features.
486 A series of filtration steps were performed to further filter the initial results. First, only features with
487 abundances above 1000 ion counts were selected. Second, features were passed through a QC
488 tolerance window of 0.1 % + 0.15 min and 5 ppm + 2.0 mDa chosen for alignment of RT and *m/z*
489 values, respectively. Third, features that were not present in all biological replicates of any single
490 sample group were removed. The data were normalized to the detected values for the internal standard
491 peak; GC-QqQ-MS data was normalized to tetracosane. After alignment and normalization of the
492 peaks of each sample, a single data set stored as a matrix was prepared. Fold change values were
493 calculated as the ratio of mean *SmCPS1*-RNAi line feature values compared with the mean values for
494 these features in the wild type lines. Student's t-tests were then used to determine whether each feature
495 was increased or decreased significantly. The aligned data was exported to SIMCA-P+ 12.0 for
496 multivariate analyses (PLS-DA).

497 Differentially accumulated features detected by LC-qTOF-MS were identified by automatic
498 comparison to a personalized metabolite database established with METLIN software (Sana et al.,
499 2008). 762 diterpenoids were collected from different *Saliva* species, among them, 86 tanshinones and
500 biosynthetically related metabolites came from *S. miltiorrhiza*. The putative metabolite peaks were

501 tentatively identified by comparison with MS/MS spectra (Yang et al., 2006; Zhou et al., 2009) and
502 reference standard compounds including sugiol, cryptotanshinone, tanshinone IIA, tanshinone IIB,
503 tanshinone I, dihydrotanshinone I, przewalskin, salvisyrianone (BioBiopha Co., Ltd.), and miltirone
504 (Faces Biochemical Co., Ltd). Differentially accumulated features detected by GC-QqQ-MS were
505 compared with the NIST 05 standard mass spectral databases and four reference standard compounds
506 including geranylgeraniol (Sigma), miltiradiene (Luqi Huang's lab), and ferruginol and sugiol
507 (BioBiopha Co., Ltd.).

508 ***In Vitro* Assays**

509 For *in vitro* functional assays, the full coding sequence of *S. miltiorrhiza* diTPS genes with specific
510 restriction enzyme sites (Supplemental Table S6) were cloned into the pGEM-T vector (Promega),
511 digested with corresponding restriction enzymes, and sub-cloned into the expression plasmid pET32a
512 (Merck) to create pET32-CPSs and pET32-KSLs. AtCPS (AAA53632) and AtKS (AAC39443) were
513 cloned from *Arabidopsis thaliana*. CfTPS2 (KF444507) and CfTPS3 (KF444508) were full
514 synthesized. The expression, purification and kinetic analysis of the recombinant proteins were
515 performed as described previously (Hillwig et al., 2011). The constructs were transformed into Tuner
516 (DE3) or Origami B (DE3) competent cells (Merck). 3-5 positive colonies were cultured in LB
517 medium with 50 $mg \cdot L^{-1}$ carbenicillin, 0.1~0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) was
518 added to induce the expression of the protein. Subsequently, cell pellets were collected and
519 re-suspended in assay buffer (50 mM phosphate pH 7.4, 10 % glycerol, 2 mM DTT, 10 mM $MgCl_2$)
520 and sonicated for 10 sec six times on ice. Lysate from the samples was centrifuged at 12000 x g, and
521 the resulting supernatant was used for the assays. The conversion of GGPP to CPP or LDPP was
522 carried out by incubating 500 μg of pET32-CPS sample protein extract with 20~50 μM GGPP (Sigma)

in a final volume of 250μL assay buffer for 2~4 h at 30 °C. Assay mixtures were hydrolysed (dephosphorylated) with 75 U bacterial alkaline phosphatase at pH 8 for 16 h at 37 °C to produce hexane-soluble products. GGPP was converted to kaurene, miltiradiene, or manoyl oxide by mixing 250 μg of pET32-CPS protein extract and 250 μg of pET32-KSL protein extraction with 20~50 μM of GGPP, and incubated for 2~4 h at 30°C. Assay mixtures were extracted three times with an equal volume of hexane. The hexane fractions were pooled, evaporated under nitrogen, re-suspended in 50 μL of hexanes or derivatized with 80 μL N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) and 8 μL pyridine at 80°C for 40 minutes, and then analyzed by GC-MS. In order to identify possible products, we obtained and analyzed a series of standard reference compounds including geranylgeraniol (sigma), geranyllinalool (sigma), (13*E*)-labda-8α,15-diol (BioBiopha Co., Ltd.), 13-*epi*-manool (BioBiopha Co., Ltd.) and sclareol (sigma).

Purification of Recombinant CPS and KSL

Cell pellets were re-suspended in 16 mL of protein lysis buffer (50 mM phosphate pH 7.4, 300 mM NaCl, 10 % glycerol, 10 mM $MgCl_2$, 20 mM imidazole) and sonicated for 10 s, six times, on ice. Lysate from the samples was centrifuged at 12000 x g for 20 min at 4 °C. The cleared lysate was transferred to pre-washed nichel-NTA beads and incubated for 30 min at 4 °C. Thereafter, the nickel-NTA beads were rinsed three times with 15 mL of washing buffer (20 mM phosphate pH 7.4, 300 mM NaCl, 100 mM imidazole). The tagged protein was then eluted by the addition of 6 mL of elution buffer (20 mM phosphate pH 7.4, 300 mM NaCl, 500 mM imidazole) to the bead bed. The buffer of the eluted proteins was then exchanged using a PD-10 column equilibrated with assay buffer (50 mM phosphate pH 7.4, 10% glycerol, 2 mM DTT, 10 mM $MgCl_2$). The purified proteins were then identified by SDS-PAGE gel and quantified by Bradford assays (Genstar).

Enzyme Kinetic Analysis of SmCPS1 and SmCPS2

20 nM of purified SmCPS1 and SmCPS2 was used for kinetic assays, with a 2-min reaction time at 25 °C. Single-vial assays were used as described above. Assays were completed in triplicate with 0.5-20 μM GGPP for SmCPS1 and 0.5-8 μM GGPP for SmCPS2. Enzymes were deactivated at the end of the 2-min reactions by incubating the reaction vial at 80 °C for 3 min, followed by quenching on ice. 300 nM purified SmKSL1 enzyme was then added to the vial in the single-vial assay for another reaction at 30 °C for 2 h. Assays were analyzed via GC-QqQ using selected ion monitoring of m/z 57 (for internal standard tetracosane) and m/z 134 for the miltiradiene product. Miltiradiene concentrations were determined relative to the internal standard by using excess SmCPS1 and SmKSL1 in single-vial assays and allowing the reaction to proceed to completion (2 h). Kinetic parameters were determined by nonlinear regression using a Michaelis-Menten model implemented in GraphPad Prism 6.03.

Positive Selection Analysis

For molecular evolution analysis, 29 protein-coding DNA sequences for CPSs were used to construct phylogenetic trees, using PhyML3.0 (Guindon et al., 2010) under the GTR nucleotide substitution model with four rate substitution categories. The branches with bootstrap values higher than 700 were used for branch-site model tests 1 and 2 in the PAML package to detect whether positive selection had acted on particular amino acid sites within specific lineages (Zhang, 2004; Zhang et al., 2005; Yang and Reis, 2011). Branch-site model uses a maximum-likelihood approach to calculate nonsynonymous to synonymous rate ratios (ω = Ka/Ks or d_N/d_S). Likelihood ratio tests were performed and $2\Delta l$ values were posteriorly transformed into exact *p*-values using PAML 4.6 (Yang, 2007). The chi-squared distribution with d.f. = 2 and d.f. = 1, which have been shown to be

conservative under conditions of positive selection (Zhang, 2004), were used to perform Tests 1 and 2, respectively. Probabilities of sites under positive selection were obtained using Bayesian approaches (Yang, 2005) implemented in PAML.

Homology Modelling, Molecular Docking, and Mutagenesis

Homology models were constructed within the SWISS-MODEL Workspace using the automatic alignment algorithm (Arnold, 2006). The crystal structures of CPS from Arabidopsis (PDB no. 3pya and 3pyb) were used as the templates (Köksal et al., 2012, 2014). The ligand docking modeling was performed with AutoDock Vina (Trott and Olson, 2010). Model visualization and binding site analysis were performed using PyMOL (www.pymol.org). Mutants were generated by whole-plasmid PCR amplification with overlapping mutagenic primers of the pGEM-T vector (Promega) clones, and verified by complete gene sequencing prior to sub-cloning into the expression vector pET32a (Merck). The resulting constructs were heterologous expressed and analyzed as described above.

Supplemental Data

Supplemental Figure S1. Electron impact (EI) mass spectrum of diterpenes in this study. (A) Miltiradiene and kaurene were comparison with the known enzyme SmCPS1, SmKSL1 and AtCPS, AtKS. (B) Labdenediol were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes. The only difference between peak 2 and 3 is the ratio of intensities of peaks m/z 275:257.

Supplemental Figure S2. The dwarf phenotype caused by downregulation of *SmCPS5*. (A) qRT-PCR analysis of transcript levels of five *CPS* genes in *SmCPS5*-RNAi line and WT in hairy root. Expression was normalized to that of *Actin*. Data are means ± SD (n = 5). Asterisk indicates significant difference at $P < 0.01$ compared with WT by Student's *t* test. *SmCPS2* is not expressed in these samples. (B) The length of pinnately compound leaf (Pcl), the length (Le) and width (Wi) of the top leaf are all significantly shorter in of T_0 generation plants than WT. Data are means ± SD (n = 30). (C) The dwarf phenotype of *SmCPS5*-RNAi and the recovery phenotype at 10d

after exogenous GA3 treatment in of T_0 generation plants. (D) *SmCPS5*-RNAi T_0 generation plants with exogenous GA3 treatment at flowering stage. Bar = 1cm. (E) T_1 generation of *SmCPS5*-RNAi shows characteristic dwarf phenotype.

Supplemental Figure S3. The intron/exon structure divergence of diTPS genes involved in labdane-related diterpenoid biosynthesis in *Arabidopsis thaliana*, rice and *S. miltiorrhiza.* The phylogenetic relationship was reconstructed using the JTT model by PhyML 3.0. The product of each enzyme was given in parenthesis. *At-Arabidopsis thaliana, Os*-rice and *Sm-S. miltiorrhiza.*

Supplemental Figure S4. Homology modelling and molecular docking result show the positive selection sites. (A-B) Homology modelling of SmCPS1 and SmCPS5 based on 3pya (A) and the partial enlarged view of the interface of the βγ domains in the purple square frame (B). The α，β and γ domain of SmCPS5 is blue, green and yellow. In SmCPS1, all atoms are cyan. (C-D) The reaction cavity of *ent*-CPP into SmCPS5 (C) and normal-CPP into SmCPS1 by molecular docking (D). Residues with blue and purple lines in SmCPS5 and SmCPS1 show positive selection sites, respectively. The ligand is the substrate GGPP.

Supplemental Figure S5. GC-MS analysis of the six mutant enzymes activity of SmCPS1 and SmCPS5. (A) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 combine SmKSL1 or SmKSL2 with GGPP as substrate. (B) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 with GGPP as substrate only. Samples were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes.

Supplemental Table S1. Information of diTPS gene family in *S. miltiorrhiza*

Supplemental Table S2. List of plant diTPS used in this paper

Supplemental Table S3. Different changed metabolites obtained from metabolomics analyze of WT and *CPS1*-RNAi lines by LC-qTOF-MS

Supplemental Table S4. Different changed metabolites obtained from metabolomics analyze of WT and CPS1-RNAi lines by GC-QqQ-MS

Supplemental Table S5. Summary of statistics for detection of positive selection for CPS group

Supplemental Table S6. Information about primers used in this study

622

623 **Table 1.** Main compounds detected in aerial part and seedling of WT and *SmCPS1*-RNAi

Organ	concentration (ng/g) FW[a]				Relative concentration （ng/g） FW[b]					
	cryptotanshinone		tanshinone IIA		miltirone		trijuganone B		tanshinone I	
	WT	RNAi	WT	RNAi	WT	RNAi	WT	RNAi	WT	RNAi
petal	0.2	0.2	2.2	2.2	12.1	12.0	19.9	20.2	10.6	11.0
sepal	1.2	1.2	22.1	23.0	72.2	73.0	143	144.4	14.4	15.0
young leaf	0.5	0.5	7.3	7.6	25.9	26.1	67.1	68.0	13.5	13.0
young root	0.8	0.8	8.9	9.0	61.5	62.3	59.3	60.2	20.3	21.0

624 a. the concentration obtained by calibration of standards.

625 b. the relative concentration obtained by compared with internal standard umbelliferone.

626

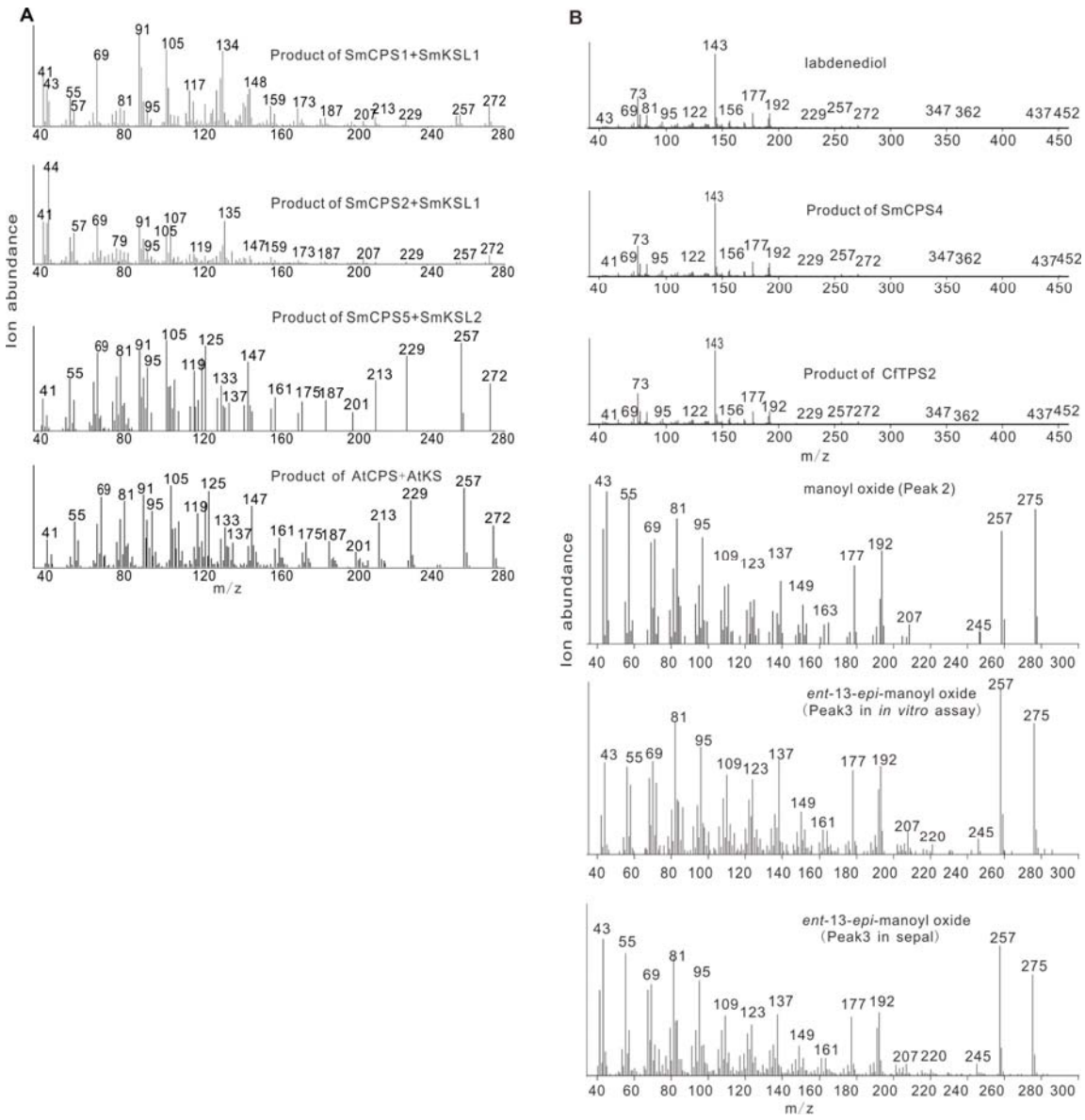


Figure S1. Electron impact (EI) mass spectrum of diterpenes in this study.

(A) Miltiradiene and kaurene were comparison with the known enzyme SmCPS1, SmKSL1 and AtCPS, AtKS. (B) Labdenediol were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes. The only difference between peak 2 and 3 is the ratio of intensities of peaks m/z 275:257.

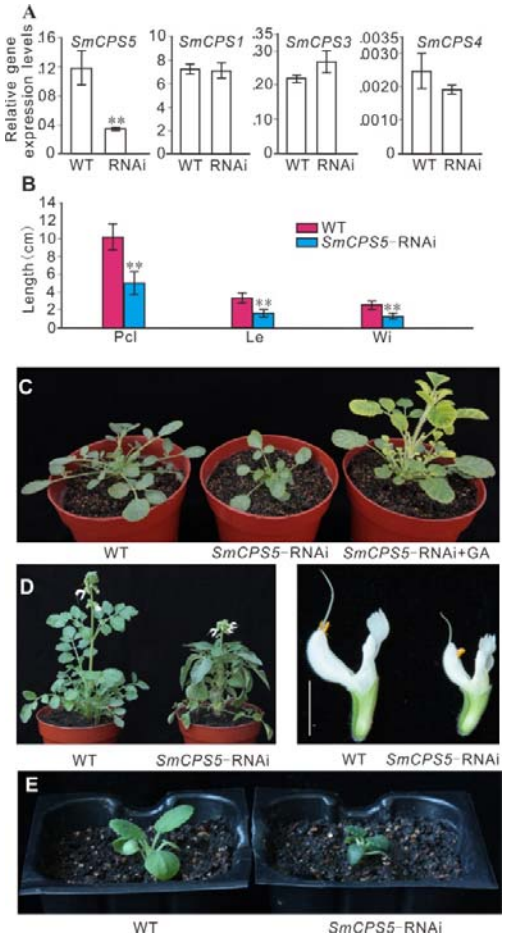
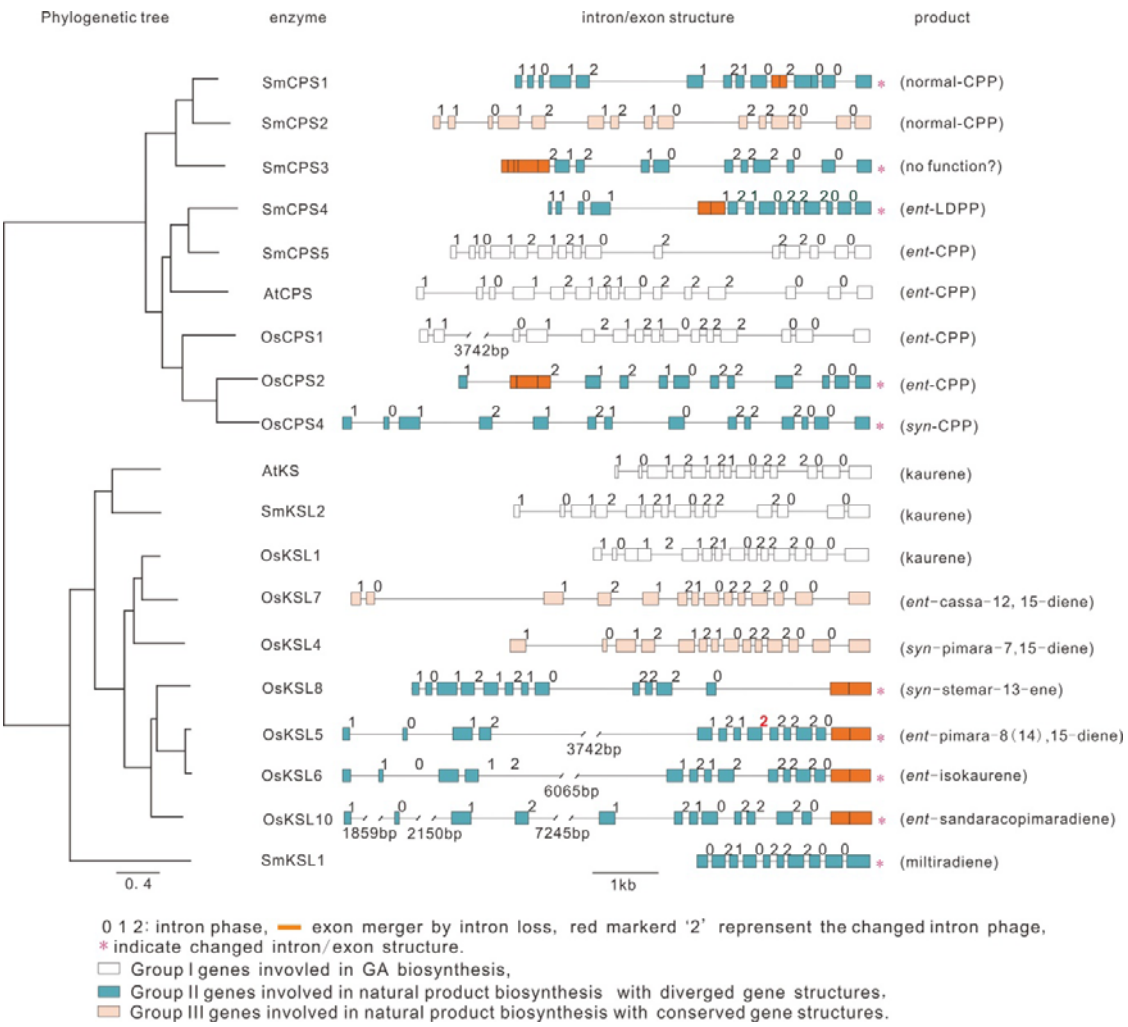


Figure S2. The dwarf phenotype caused by downregulation of *SmCPS5*. (A) qRT-PCR analysis of transcript levels of five *CPS* genes in *SmCPS5*-RNAi line and WT in hairy root. Expression was normalized to that of *Actin*. Data are means ± SD (n = 5). Asterisk indicates significant difference at $P < 0.01$ compared with WT by Student's *t* test. *SmCPS2* is not expressed in these samples. (B) The length of pinnately compound leaf (Pcl), the length (Le) and width (Wi) of the top leaf are all significantly shorter in of T_0 generation plants than WT. Data are means ± SD (n = 30). (C) The dwarf phenotype of *SmCPS5*-RNAi and the recovery phenotype at 10d after exogenous GA3 treatment in of T_0 generation plants. (D) *SmCPS5*-RNAi T_0 generation plants with exogenous GA3 treatment at flowering stage. Bar = 1cm. (E) T1 generation of *SmCPS5*-RNAi shows characteristic dwarf phenotype.



Supplemental Figure S3. The intron/exon structure divergence of diTPS genes involved in labdane-related diterpenoid biosynthesis in *Arabidopsis thaliana*, rice and *S. miltiorrhiza.* The phylogenetic relationship was reconstructed using the JTT model by PhyML 3.0. The product of each enzyme was given in parenthesis. *At-Arabidopsis thaliana, Os*-rice and *Sm-S. miltiorrhiza.*

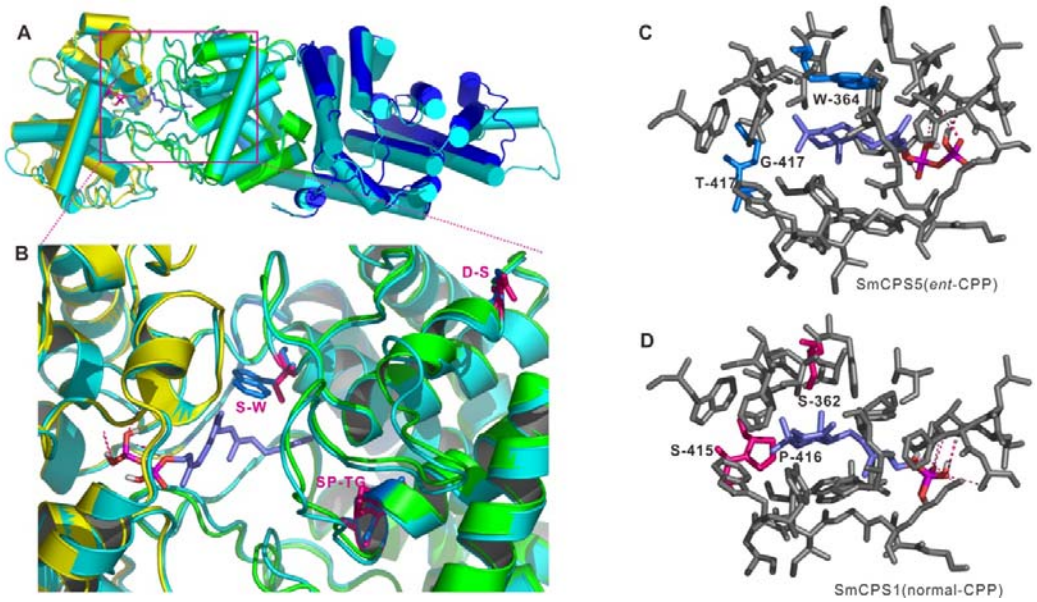


Figure S4. Homology modelling and molecular docking result show the positive selection sites. (A-B) Homology modelling of SmCPS1 and SmCPS5 based on 3pya (A) and the partial enlarged view of the interface of the βγ domains in the purple square frame (B). The α，β and γ domain of SmCPS5 is blue, green and yellow. In SmCPS1, all atoms are cyan. (C-D) The reaction cavity of *ent*-CPP into SmCPS5 (C) and normal-CPP into SmCPS1 by molecular docking (D). Residues with blue and purple lines in SmCPS5 and SmCPS1 show positive selection sites, respectively. The ligand is the substrate GGPP.

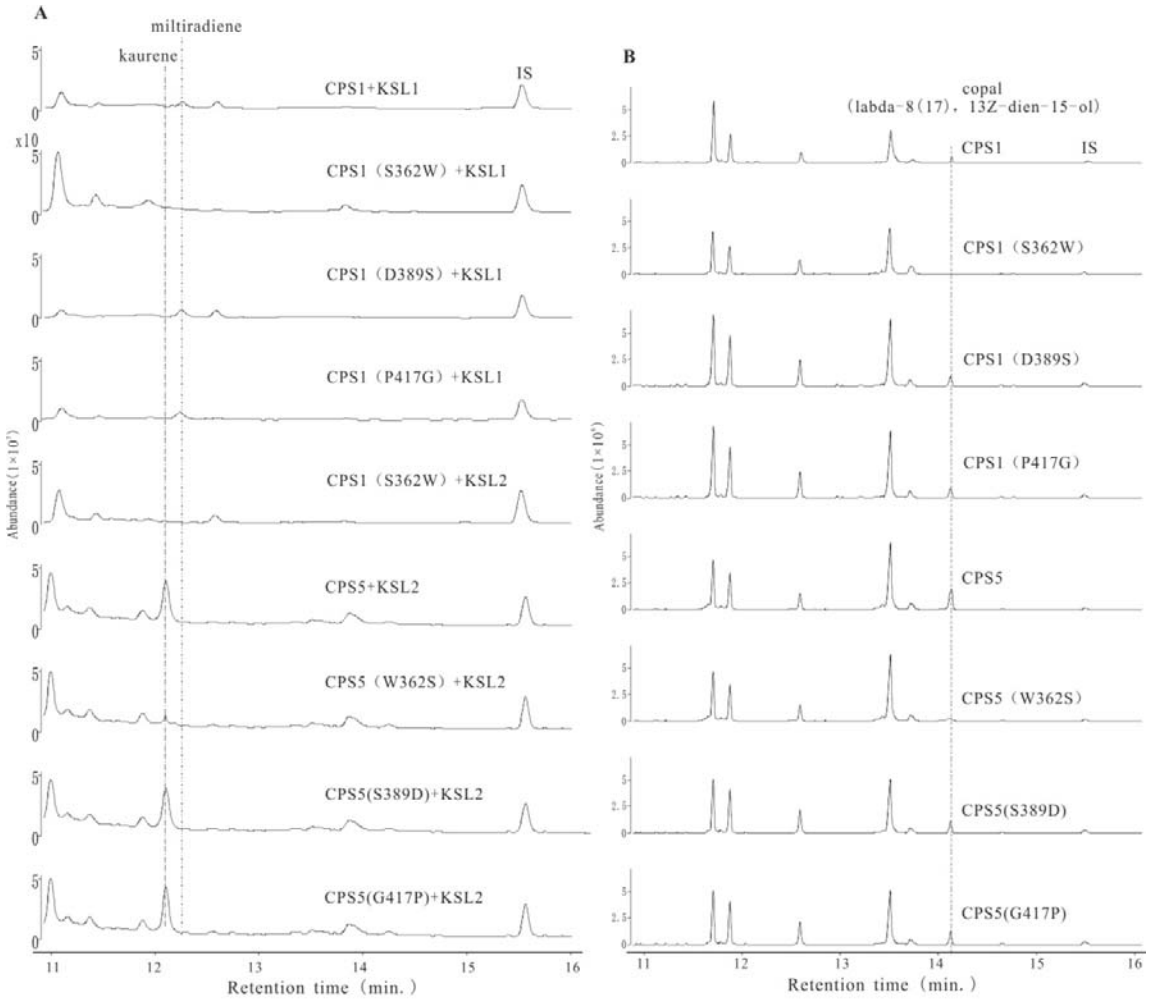


Figure S5. GC-MS analysis of the six mutant enzymes activity of SmCPS1 and SmCPS5.

(A) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 combine SmKSL1 or SmKSL2 with GGPP as substrate. (B) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 with GGPP as substrate only. Samples were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes.

Table S1. Information of diTPS gene family in *S. miltiorrhiza*

Gene	Accession No.	Length[a]			Function annotation	reference
		cDNA (bp)	Protein	gDNA (bp)		
CPS1	KC814639	2382	793aa	5422	copalyl diphosphate synthase	Gao et al.(2009)
CPS2	KC814640	2397	798aa	6633	copalyl diphosphate synthase	This study
CPS3	KC814641	2385	794aa	5619	no function	This study
CPS4	KP063138	2331	776aa	4908	labd-13-en-8-ol diphosphate synthase	This study
CPS5	KC814642	2382	793aa	6256	*ent*-copalyl diphosphate synthase	This study
KSL1	EF635966	1788	595aa	2667	miltiradiene synthase	Gao et al.(2009)
KSL2	KC814643	2424	807aa	5405	kaurene synthase	This study

[a] The cDNA sequences of CPS1 to CPS5 and KSL2 are from bh2-7, genomic sequence of CPS1, CPS3 and CPS4 are from bh2-7, and the other sequences from genomic sequencing material published by Ma et al., 2012.

6 **Supplemental Table 2. List of plant DiTPS involved in this paper**

Gene	Species	GenBank Accession	Funtional Annotation*	Ref.
AgAS	*Abies grandis*	AAB05407	abietadiene synthase	Vogel et al.,1996
AtCPS	*Arabidopsis thaliana*	AAA53632	*ent*-copalyldiphosphatesynthas	Sun et al., 1994
AtKSL	*A. thaliana*	AAC39443	*ent*-kaurene synthase	Yamaguchi et al., 1998
CcCLS	*Cistus creticus*	ADJ93862	copal-8-ol diphosphate synthase	Falara et al., 2010
CmCPS1	*Cucurbita maxima*	AAD04292	*ent*-copalyldiphosphate synthase	Smith et al.,1998
CmCPS2	*C. maxima*	AAD04293	*ent*-copalyldiphosphate synthase	Smith et al.,1998
CmKS	*C. maxima*	AAB39482	*ent*-kaurene synthase	Yamaguchi et al., 1996
GbLS	*Ginkgo biloba*	AAL09965	levopimaradiene synthase	Schepmannet al., 2001
IeCPS1	*Isodon eriocalyx*	AEP03177	*ent*-copalyldiphosphate synthase	Li et al., 2012
IeCPS2	*I. eriocalyx*	AEP03175	*ent*-copalyldiphosphate synthase	Li et al., 2012
OsCPS1	*Oryza sativa*	BAD42449	*ent*-copalyldiphosphate synthase	Otomo et al., 2004
OsCPS2	*O. sativa*	AAT11021	*ent*-copalyldiphosphate synthase	Prisic et al., 2004
OsCPS4	*O. sativa*	AAS98158	*syn*-copalyldiphosphate synthase	Xu et al., 2004
OsKS1	*O. sativa*	AAQ72559	ent-kaurene synthase	Margis-Pinheiro et al., 2005
OsKSL4	*O. sativa*	AAU05906	pimara-7,15-diene synthase	Wilderman et al., 2004
OsKSL5	*O. sativa*	Q6Z5J6	*ent*-pimara-8(14),15-diene synthase	Margis-Pinheiro et al., 2005
OsKSL6	*O. sativa*	A4KAG8	*ent*-isokaur-15-ene synthase	Xu et al., 2007
OsKSL7	*O. sativa*	BAC56714	*ent*-cassa-12,15-diene synthase	Cho et al., 2004
OsKLS8	*O. sativa*	Q6BDZ9	stemar-13-ene synthase	Nemoto et al., 2004
OsKSL10	*O. sativa*	Q2QQJ5	*ent*-sandaracopimara-8(14),15-diene synthase	Peters. 2006
OsKSL11	*O. sativa*	Q1AHB2	stemod-13(17)-ene synthase	Morrone et al., 2006
PaIso	*Picea abies*	AAS47690	isopimaradiene synthase	Martin et al., 2004
PaLAS	*P. abies*	AAS47691	levopimaradiene/abietadiene synthase	Martin et al., 2004
PsIso	*P. sitchensis*	ADZ45512	Isopimaradiene synthase	Keeling et al., 2011
PsLAS	*P. sitchensis*	ADZ45517	Levopimaradiene/abietadiene synthase	Keeling et al., 2011
PgCPS	*P. glauca*	ADB55707	*ent*-copalyldiphosphate synthase	Keeling et al., 2010
PgKS	*P. glauca*	ADB55708	*ent*-kaurene synthase	Keeling et al., 2010
PsCPS	*P. sitchensis*	ADB55709	*ent*-copalyldiphosphate synthase	Keeling et al., 2010
PsKS	*P. sitchensis*	ADB55710	*ent*-kaurene synthase	Keeling et al., 2010
PpCPS/KS	*Physcomitrella patens*	BAF61135	*ent*-kaurene synthase	Hayashi et al., 2006
PsaCPS	*Pisum sativum*	AAB58822	Bifunctional *ent*-copalyldiphosphate synthase	Ait-Ali et al.,1997
SdCPS	*Scoparia dulcis*	BAD91286	*ent*-copalyldiphosphate synthase	Nakagiri et al., 2005
SmCPS1	*S. miltiorrhiza*	KC814639	copalyl diphosphate synthase	Gao et al., 2009
SmCPS2	*S. miltiorrhiza*	KC814640	copalyl diphosphate synthase	this study
SmCPS3	*S. miltiorrhiza*	KC814641	No function	this study

SmCPS4	*S. miltiorrhiza*		copal-8-ol diphosphate synthase	this study
SmCPS5	*S. miltiorrhiza*	KC814642	*ent*-copalyldiphosphate synthase	this study
SmKSL1	*S. miltiorrhiza*	KC814643	miltiradiene synthase	Gaoet al.,2009
SmKSL2	*S. miltiorrhiza*	ABV08817	*ent*-kaurene synthase, *ent*-13-*epi*-manoyl oxide synthase	this study
SrCPS	*Stevia rebaudiana*	AAB87091	*ent*-copalyldiphosphate synthase	Richman et al., 1999
SrKS1	*S. rebaudiana*	AAD34294	*ent*-kaurene synthase	Richman et al., 1999
SrKS2	*S. rebaudiana*	AAD34295	*ent*-kaurene synthase	Richman et al., 1999
TcTS	*Taxus brevifolia*	AAC49310	taxadiene synthase	Wildung and Croteau.1996
ZmAn1	*Zea mays*	AAA73960	*ent*-copalyldiphosphate synthase	Bensen et al., 1995
ZmAn2	*Z. mays*	AAT70083	*ent*-copalyldiphosphate synthase	Bensen et al., 1995
HvCPS	*Hordeum vulgare*	AAT49065	*ent*-copalyldiphosphate synthase	Spielmeyer et al., 2004
HvKSL1	*H. vulgare*	AAT49066	*ent*-kaurene synthase	Spielmeyer et al., 2004
SsLPS	*Salvia sclarea*	*AET21247*	copal-8-ol diphosphate synthase	Schalk et al., 2012
SsScS	*S. sclarea*	AET21246	sclareol synthase	Schalk et al., 2012
NtCPS2	*Nicotiana tabacum*	CCD33018	copal-8-ol diphosphate synthase	Christopheet al., 2012
NtABS	*N. tabacum*	CCD33019	cis-abienol synthase	Christophe et al., 2012
TaKSL5	*Triticum aestivum*	BAL41692	nerolidol synthase	Hillwig et al., 2011
ZmTPS1	*Zea mays*	NP_001105097	sesquiterpene synthase	Schnee et al., 2002
AbCAS	*Abies balsamea*	AEL99953	cis-abienol synthase	Zerbe et al., 2012
SlPHS	*Solanum lycopersicum*	ACO56896	phellandrene synthase	Schilmiller et al., 2009
ShSBS	*S. habrochaites*	B8XA41	santalene and bergamotene synthase	Sallaud et al., 2009
PtKS	*Populus trichocarpa*	XP_002311286	*ent*-kaurene synthase	Tuskan et al., 2006
LsKS	*Lactuca sativa*	BAB12441	*ent*-kaurene synthase	Sawada et al., 2008
LsCPS	*L. sativa*	BAB12440	*ent*-kaurene synthase	Sawada et al., 2008
GrTPS1	*Grindelia robusta*	AGN70886	13-labden-8,15-diol pyrophosphate synthase	Zerbe et al., 2013
GrTPS6	*G. robusta*	AGN70887	manoyl oxide	Zerbe et al., 2013
EpTPS1	*Euphorbia peplus*	KC702395	*ent*-kaurene synthase	Zerbe et al., 2013
CfTPS14	*C. forskohlii*	AGN70881	*ent*-kaurene synthase	Zerbe et al., 2013
EpTPS7	*Euphorbia peplus*	AGN70883	*ent*-copalyldiphosphate synthase	Zerbe et al., 2013
SlTPS40	*Solanum lycopersicum*	JN412074	*ent*-copalyldiphosphate synthase	Falara et al., 2011
SlTPS14	*S. lycopersicum*	JN412091	*ent*-kaurene synthase	Falara et al., 2011
CfTPS1	*Coleus forskohlii*	KF444506	copalyldiphosphate synthase	Pateraki et al., 2014
CfTPS2	*C. forskohlii*	KF444507	copal-8-ol diphosphate synthase	Pateraki et al., 2014
CfTPS3	*C. forskohlii*	KF444508	(13R) manoyl oxide synthase	Pateraki et al., 2014
CfTPS4	*C. forskohlii*	KF444509	miltiradiene synthase	Pateraki et al., 2014
MvCPS1	*Marrubium vulgare*	KJ584450	peregrinol diphosphate synthase	Zerbe et al., 2014
MvCPS3	*M. vulgare*	KJ584452	copalyl diphosphate synthase	Zerbe et al., 2014
MvCPS4	*M. vulgare*	KJ584453	*ent*-kaurene synthase	Zerbe et al., 2014

MvCPS5	*M. vulgare*	KJ584454	9,13-epoxy-labd-14-en synthase	Zerbe et al., 2014
SmMDS	*Selaginella moellendorffii*	AB668998	miltiradiene synthase	Sugai et al., 2011

*Functional annotation is based on the main terpenoid product(s) of recombinant enzymes expressed in *E. coli*. Many TPSs produced multiple products.

Supplemental References

Ait-Ali T, Swain SM, Reid JB, Sun TP, Kamiya Y. (2003). The LS locus of pea encodes the gibberellin biosynthesis enzyme *ent*-kaurene synthase A. Plant J. 11(3): 443-454.

Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, Briggs SP. (1995). Cloning and characterization of the maize An1 gene. Plant Cell 7(1): 75-84.

Cho EM, Okada A, Kenmoku H, Otomo K, Toyomasu T, Mitsuhashi W, Sassa T, Yajima A, Yabuta G, Mori K, Oikawa H, Toshima H, Shibuya N, Nojiri H, Omori T, Nishiyama M, Yamane H. (2004). Molecular cloning and characterization of a cDNA encoding ent-cassa-12, 15-diene synthase, a putative diterpenoidphytoalexin biosynthetic enzyme, from suspension-cultured rice cells treated with a chitin elicitor. Plant J. 37(1): 1-8.

Falara V, Pichersky E, Kanellis AK. (2010). A copal-8-ol diphosphate synthase from the angiosperm *Cistus creticus* subsp. creticus is a putative key enzyme for the formation of pharmacologically active, oxygen-containing labdane-type diterpenes. Plant physiol 154(1): 301-310.

Falara V, Akhtar TA, Nguyen TT, Spyropoulou EA, Bleeker PM, Schauvinhold I, Pichersky E. (2011). The tomato terpene synthase gene family. *Plant physiol 157*(2): 770-789.

Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong JQ, Yang B, Peters RJ. (2009). A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org Lett 11(22): 5170-5173.

Hayashi KI, Kawaide H, Notomi M, Sakigi Y, Matsuo A, Nozaki H. (2006). Identification and functional analysis of bifunctional *ent*-kaurene synthase from the moss Physcomitrella patens. FEBS Lett 580(26): 6175-6181.

Hillwig ML, Xu M, Toyomasu T, Tiernan MS, Wei G, Cui G, Huang LQ, Peters RJ. (2011). Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J. 68(6): 1051-1060.

Keeling CI, Dullat HK, Yuen M, Ralph SG, Jancsik S, Bohlmann J. (2010). Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. Plant Physiol 152(3): 1197-1208.

Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J. (2011). Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). BMC Plant Boil 11(1): 43.

Li JL, Chen QQ, Jin QP, Gao J, Zhao PJ, Lu S, Zeng Y. (2012). IeCPS2 is potentially involved in the biosynthesis of pharmacologically active Isodon diterpenoids rather than gibberellin. Phytochemistry 76: 32-39.

Margis-Pinheiro M, Zhou XR, Zhu QH, Dennis ES, Upadhyaya NM. (2005). Isolation and characterization of a Ds-tagged rice (*Oryza sativa* L.) GA-responsive dwarf mutant defective in an early step of the gibberellin biosynthesis pathway. Plant Cell Rep 23(12): 819-833.

Martin DM, Fäldt J, Bohlmann J. (2004).Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiol 135(4): 1908-1927.

Morrone D, Jin Y, Xu M, Choi SY, Coates RM, Peters RJ.(2006). An unexpected diterpene cyclase from rice: functional identification of a stemodene synthase. Arch. Biochem. Biophys 448 (1-2):133-140.

Nakagiri T, Lee JB, Hayashi T. (2005). cDNA cloning, functional expression and characterization of *ent*-copalyldiphosphate synthase from *Scoparia dulcis* L. Plant Sci 169(4): 760-767.

NemotoT, Cho EM, Okada A, Okada K, Otomo K, Kanno Y, Toyomasu T, Mitsuhashi W, Sassa T, Minami E, Shibuya N, Nishiyama M, Nojiri H, Yamane H. (2004) Stemar-13-ene synthase, a diterpene cyclase involved in the biosynthesis of the phytoalexin oryzalexin S in rice. FEBS Lett 571 (1-3): 182-186.

Otomo K, Kenmoku H, Oikawa H, König WA, Toshima H, Mitsuhashi W, Yamane H, Sassa T, Toyomasu T. (2004). Biological functions of *ent*-and *syn*-copalyldiphosphate synthases in rice: key enzymes for the branch point of gibberellin and phytoalexin biosynthesis. Plant J 39(6): 886-893.

Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Møller BL, Bohlmann J, Hamberger B. (2014). Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. Plant Physiol 164: 1222-1236.

Prisic S, Xu M, Wilderman PR, Peters RJ. (2004). Rice contains two disparate *ent*-copalyldiphosphate synthases with distinct metabolic functions. Plant physiol 136(4): 4228-4236.

Peters, R. J.(2006) Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. Phytochemistry 67 (21):2307-2317.

Richman AS, Gijzen M, Starratt AN, Yang Z, Brandle JE. (1999). Diterpene synthesis in Stevia rebaudiana: recruitment and up‐regulation of key enzymes from the gibberellin biosynthetic pathway. Plant J 19(4): 411-421.

Sallaud C, Rontein D, Onillon S, Jabes F, Duffe P, Giacalone C, Thoraval S, Escoffier C, Herbette G, Leonhardt N, Causse M, Tissier A. (2009). A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell 21 (1):301-317.

Sallaud C, Giacalone C, Töpfer R, Goepfert S, Bakaher N, Rösti S, Tissier A. (2012). Characterization of two genes for the biosynthesis of the labdanediterpene Z‐abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. Plant J 72(1): 1-17.

SawadaY, Katsumata T, Kitamura J, Kawaide H, Nakajima M, Asami T, Nakaminami K, Kurahashi T, Mitsuhashi W, Inoue Y, Toyomasu T. (2008).Germination of photoblastic lettuce seeds is regulated via the control of endogenous physiologically active gibberellin content, rather than of gibberellin responsiveness. J Exp Bot 59 (12), 3383-3393.

Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L. (2012). Toward a biosynthetic route to sclareol and amber odorants. J Am Chem Soc 134: 18900-18903.

Schepmann HG, Pang J, Matsuda S. (2001). Cloning and Characterization of *Ginkgo biloba* Levopimaradiene Synthase, Which Catalyzes the First Committed Step in Ginkgolide Biosynthesis. Arch Biochem Biophys 392(2): 263-269.

Schilmiller AL, Schauvinhold I, Larson M, Xu R, Charbonneau AL, Schmidt A, Wilkerson C, Last RL, Pichersky E. (2009). Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc Natl Acad Sci USA 106(26):10865-10870.

Schnee C, Köllner TG, Gershenzon J, Degenhardt J. (2002). The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)-β-farnesene, (E)-nerolidol, and (E, E)-farnesol after herbivore damage. Plant physiol 130(4): 2049-2060.

Smith MW, Yamaguchi S, Ait-Ali T, Kamiya Y. (1998). The first step of gibberellin biosynthesis in pumpkin is catalyzed by at least two copalyl diphosphate synthases encoded by differentially regulated genes. Plant physiol 118(4): 1411-1419.

89 Spielmeyer W, Ellis M, Robertson M, Ali S, Lenton JR, Chandler PM. (2004). Isolation of gibberellin metabolic
90 pathway genes from barley and comparative mapping in barley, wheat and rice. Theor Appl Genet 109(4):
91 847-855.
92 Sugai Y, Ueno Y, Hayashi K, Oogami S, Toyomasu T, Matsumoto S, Natsume M, Nozaki H, Kawaide H.
93 (2011). Enzymatic ^{13}C labeling and multidimensional NMR analysis of miltiradiene synthesized by
94 bifunctional diterpene cyclase in *Selaginella moellendorffii*. J Biol Chem 86(50): 42840-42847.
95 Sun TP, Kamiya Y. (1994). The Arabidopsis GA1 locus encodes the cyclaseent –kaurene synthetase A of
96 gibberellin biosynthesis. Plant Cell 6(10): 1509-1518.
97 Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S,
98 Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, BoerjanW, Brun A, Brunner A,
99 Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J,
100 Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B,
101 Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A,
102 Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N,
103 Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R,
104 Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F,
105 Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, PeredaV, Peter G, Philippe R, Pilate G,
106 Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J,
107 Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K,
108 Wessler S, Yang G Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D. (2006).The genome
109 of black cottonwood, Populus trichocarpa (Torr. & Gray) Science 313 (5793):1596-1604.
110 Vogel BS, Wildung MR, Vogel G, Croteau R. (1996). Abietadiene synthase from grand fir (*Abies grandis*)
111 cDNA isolation, characterization, and bacterial expression of a bifunctionalditerpenecyclase involved in resin
112 acid biosynthesis. J Biol Chem 271(38): 23262-23268.
113 Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ. (2004). Identification of *syn*-pimara-7, 15-diene synthase
114 reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis.
115 Plant physiol 135(4): 2098-2105.
116 Wildung MR, Croteau R. (1996).A cDNA clone for taxadiene synthase, the diterpenecyclase that catalyzes the
117 committed step of taxol biosynthesis. J Biol Chem 271(16): 9201-9204.
118 Xu MM, Hillwig ML, Prisic S, Coates RM, Peters RJ. (2004). Functional identification of rice *syn*-copalyl
119 diphosphate synthase and its role in initiating biosynthesis of diterpenoidphytoalexin/allelopathic natural
120 products. Plant J 39(3): 309-318.
121 Xu MM, Wilderman PR, Morrone D, Xu J, Roy A, Margis-Pinheiro M, Upadhyaya NM, Coates RM，Peters RJ.
122 (2007). Functional characterization of the rice kaurene synthase-like gene family. Phytochemistry 68 (3):
123 312-326.
124 Yamaguchi S, Saito T, Abe H, Yamane H, Murofushi N, Kamiya Y. (1996). Molecular cloning and
125 characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from
126 pumpkin (*Cucurbita maxima* L.).Plant J 10(2): 203-213.
127 Yamaguchi S, Sun TP, Kawaide H, Kamiya Y. (1998). The GA2 Locus of *Arabidopsis thaliana* encodes
128 *ent*-kaurene synthase of gibberellin biosynthesis. Plant Physiol 116(4): 1271-1278.
129 Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Bohlmann J. (2012).
130 Bifunctionalcis-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its
131 implications for diterpenoid fragrance production. J Biol Chem 287(15): 12121-12131.

132 **Table S3.** Different changed metabolites obtained from metabolomics analyze of WT and *CPS1*-RNAi lines by LC-qTOF-MS

No.	t_R (min)	positive ion mode			redicted molecular formula	Predicted assigned identity	MS/MS product ion (m/z)	relative content (ug/g) FW		Percent (%)[b]
		$[M+H]^+$	$[M+Na]^+$	major fragment				WT+SE	RNAi+SE	
1	5.72	313.1070	335.0896	647.1890	C18H16O5	tanshindiol B[a]	295.0969, 277.0860, 267.1024, 249.0914	37.4 ± 7.0	n.d.	99.6
2	5.87	313.1071	335.0896	647.1892	C18H16O5	tanshindiol C[a]	295.0965, 277.0864, 267.1021, 249.0913	270.3 ± 46.9	n.d.	98.9
3	6.28	313.1073	335.0896	647.1889	C18H16O5	tanshindiol A[a]	295.0969, 277.0865, 267.1024, 249.0902	47.3 ± 8.8	n.d.	99.2
4	6.89	343.1547	365.1378	-	C20H22O5	prineoparaquinone*	-	41.6 ± 5.2	n.d.	98.9
5	6.95	295.0968	317.0796	611.1683	C18H14O4	3α-hydroxymethylenetanshinquinone[a]	277.0864, 267.1022, 251.0709, 249.0912, 221.0965	204.1 ± 30.1	n.d.	99.2
6a	7.03	297.1122	319.0952	615.1986, 342.1698	C18H16O4	tanshinol B[a]	279.1016, 261.0911, 233.0963, 218.0717, 205.1017, 190.0774	768.0 ± 105.6	7.0 + 3.1	99.4
7	7.16	281.1179	303.1005	263.1078	C18H16O3	danshenxinkun B	263.1078, 248.0837, 235.1121, 220.0887, 192.0929	191.9 ± 24.5	2.2 + 0.7	98.9
8	7.32	311.1280	333.1103	643.2302, 356.1859	C19H18O4	tanshinone IIB[a]	293.1164, 275.1070, 265.0852, 251.1075, 247.1127	2088.3 ± 219.8	40. + 17.8	94.5
9	7.37	283.1338	305.1154	587.2417	C18H18O3	no hits	-	488.9 ± 1004	3.1 + 0.8	98.2
10	7.64	341.1389	363.1217	703.2527, 386.1907	C20H20O5	trijuganone C[a]	281.1175, 263.1068，235.1124, 220.0882	142.4 ± 23.7	3.1 + 0.8	98.1
11	7.85	299.1645	321.1463	344.2226, 619.3023	C19H22O3	1R-hydroxymiltirone*	281.1539, 266.1296, 263.1435, 221.0961, 193.1020	1357.3 ± 116.7	1.9 + 0.9	99.4
12	7.92	313.1439	335.1264	-	C19H20O4	3-hydroxycryptotanshinone	-	10.0 ± 1.6	n.d.	98.1

13	8.02	311.1279	333.1106	643.2289, 356.1860	C19H18O4	hydroxytanshinone IIA[a]	293.1189, 278.0941, 247.1122, 219.1179, 204.0925	200.4 ± 23.8	0.8 + 0.4	99.4
14	8.09	309.1124	331.0953	639.1989, 354.1703	C19H16O4	no hits	-	252.0 ± 9.6	3.5 + 1.2	98.6
15	8.17	299.2011	-	-	C20H26O2	5,6-dehydrosugiol	257.1541, 243.1341, 229.1228, 187.0751	8.0 ± 0.7	2.4 + 0.9	98.7
16	8.18	279.1018	301.0843	579.1778	C18H14O3	dihydrotanshinone I[a]	261.0910, 233.0965, 218.0730, 205.1015, 190.0777	1233.6 ± 156.9	84.4 + 12.0	98.3
17	8.20	273.1846	-	-	C18H24O2	przewalskin[a]	255.1745, 243.1743	11.3 ± 1.3	0.2 + 0.1	96.9
18	8.21	287.2017	-	-	C19H26O2	przewalskin Y-1	-	157.8 ± 12.0	8.5 + 1.6	98.5
19	8.26	315.1605	337.1421	-	C19H22O4	neocryptotanshinone	-	7.9 ± 1.5	n.d.	96.7
20	8.27	327.1595	349.1417	-	C20H22O4	7-hydroxy-12-methoxy-20-nor-abieta-1,5(10),7,9,12-pentaen-6,14-dione*	-	13.5 ± 1.5	0.5 + 0.2	98.3
21	8.41	339.1234	361.1051	669.2209, 384.1806	C20H18O5	methyltanshinonate[a]	279.0968, 261.0909, 233.0959, 218.0724, 190.0769	2137.8 ± 209.9	120.0 + 40.6	98.9
22	8.52	281.1175	303.0995		C18H16O3	trijuganone B[a]	263.1069, 248.0826, 235.1129, 220.0880, 192.0933	6957.5 ± 590.7	306.4 + 95.1	97.3
23	8.77	327.1589	349.1413	372.2171	C20H22O4	prionoid E*	-	105.8 ± 10.9	0.0	96.8
24	8.84	317.2118	339.1932		C20H28O3	cryptanol*	-	11.17 ± 1.4	n.d.	96.9
25	8.90	301.2169	-	-	C20H28O2	sugiol[a]	-	59.3 ± 5.0	1.1 + 0.5	95.9
26	9.10	313.1805	335.1626	647.3344, 358.2383	C20H24O3	salvisyrianone[a]*	-	24.5 ± 2.0	n.d.	99.3
27	9.13	277.0862	299.0688	575.1466, 322.1436	C18H12O3	tanshinone I[a]	249.0918, 234.0675, 221.0969, 193.1013, 178.0780	2252.3 ± 258.3	321.1 + 48.9	97.9

28	9.29	297.1490	319.1314	615.2726	C19H20O3	cryptotanshinone[a]	282.1254, 279.1384, 267.1018, 254.0945, 221.0961, 193.1019	15959.6 ± 1088.1	270.8 + 88.8	96.3
29	9.43	295.1333	317.1153	611.2409, 340.1914	C19H18O3	isotanshinone IIA	277.1224, 262.0980, 249.1274	107.1 ± 9.1	6.2 + 2.8	98.3
30	9.43	317.2118	339.1937	362.2698	C20H28O3	7β-hydroxy-8,13-abietadiene-11,12-dione*	-	92.1 ± 30.8	n.d.	97.6
31	10.06	267.1382	289.1208	555.2509, 312.1964	C18H18O2	4-methylenemiltirone[a]	249.1272, 225.0910, 207.0809	246.8 ± 47.8	2.1 + 0.9	98.9
32	10.44	329.2458	-		C22H32O2	pomiferin G*	-	12.0 ± 1.2	n.d.	96.8
33	10.62	295.1329	317.1148	611.2403, 340.1913	C19H18O3	tanshinone IIA[a]	277.1225, 262.0989, 249.1274, 234.1039, 221.1325, 206.1089	19520.4 ± 964.5	1820.6 + 706.7	95.4
34	10.91	301.2165	-	-	C20H28O2	16-hydroxy-6,7-didehydroferruginol*	-	7.9 ± 1.5	n.d.	97.9
35	11.03	283.1693	305.1513	578.2132, 328.2278	C19H22O2	miltirone[a]	265.1591, 223.1119, 195.1168	6271.2 ± 988.8	29.1 + 10.7	98.4
36	11.42	309.1492	331.1309	-	C20H20O3	3-hydroxysalvilenone*	291.1385, 267.1022	7.9 ± 1.5	n.d.	96.9
37	11.66	331.2271	353.2093	-	C21H30O3	cryptojaponol*	-	7.8 ± 1.5	n.d.	99.2
38	11.69	297.1854	319.1672	342.2430	C20H24O2	saprorthoquinone*	279.1750, 269.1905, 255.1399	101.2 ± 12.4	0.7 + 0.3	97.8
39	12.53	299.2011	-	-	C20H26O2	12-deoxy-6,7-dehydroroyleanone*	-	10.4 ± 1.2	n.d.	96.8

[a]Identified via retention time and MS/MS spectrum compared to reference standard.

[b]the proposition of metabolite content in periderm to the whole WT root containing periderm, cortex and xylem.

* compounds didn't isolate in *S. miltiorrhiza* before.

n.d. not detected.

Table S4. Different changed metabolites obtained from metabolomics analyze of WT and CPS1-RNAi lines by GC-QqQ-MS

No.	metabolites	Molecular formula	similatity	t_R (min.)	quantified ion	relative content (ug/g) DW		Percent (%) [b]
						WT ± SE	RNAi ± SE	
1	β-springene[a]	C20H32	94.9	12.35	93.00	n.d.	5.6 ± 0.02	n.d.
2	α-springene[a]	C20H32	94.0	12.84	69.10	n.d.	1.3 ± 0.03	n.d.
3	geranylgeraniol[a]	C20H34O	87.6	15.41	69.10	n.d.	0.5 ± 0.01	n.d.
5	unknown			13.72	135.00	n.d.	0.8 ± 0.02	100
4	Neophytadiene	C20H38	89.2	11.62	57.10	0.03	0.01	100
6	abietatriene[a]	C20H30	92.0	14.00	173.10	0.6 ± 0.01	n.d.	99.9
7	miltiradiene[a]	C20H32	94.7	14.06	80.20	0.2 ± 0.01	n.d.	100
8	ferruginol[a]	C20H30O	94.7	17.25	286.20	29.5 ± 0.6	0.15	99.9
9	sugiol[a]	C20H28O2	87.2	21.20	285.00	1.1 ± 0.02	n.d.	100
10	cryptotanshinone[a]	C19H20O3	90.0	21.59	296.10	6.8 ± 0.14	n.d.	100
11	Retinol	C20H30O	82.8	15.74	174.10	0.08	n.d.	100
12	1,4-dimethyl-8-isopropylidenetricyclo[5.3.0.094,100]decane	C15H24	88.2	14.32	204.10	1.6 ± 0.04	0.02	100
13	unknown			15.87	123.00	0.6 ± 0.02	n.d.	100
14	unknown			16.73	274.00	0.4 ± 0.01	n.d.	100
15	unknown			18.40	282.00	2.2 ± 0.04	0.04	100
16	unknown			18.61	281.00	6.4 ± 0.08	0.32	100
17	unknown			20.64	249.00	2.0 ± 0.06	n.d.	100
18	unknown			21.04	265.10	0.4 ± 0.01	n.d.	100
19	unknown			21.32	279.10	0.7 ± 0.02	n.d.	100
20	unknown			21.84	282.10	2.0 ± 0.05	n.d.	100
21	unkwown			21.90	296.00	0.9 ± 0.02	n.d.	100
22	unknown			22.34	265.10	3.2 ± 0.03	n.d.	100
23	unknown			17.16	284.10	3.7 ± 0.06	n.d.	100

n.d. not detected.

[a] Identified via retention time and EI-MS spectrum compared to reference standard.

[b] the proposition of metabolite content in periderm to the whole WT root containing periderm, cortex and xylem.

n.d. not detected.

151 **Table S5. Summary of statistics for detection of positive selection for CPS group**

Lineages	Branches	lnL (MA)	lnL (MA1)	2Δ/ (LRT)[a] Test1 (MA/M1a)	2Δ/ (LRT)[a] Test2 (MA/MA1)	ω_2[b]
CPS	*a*	-38972.35	-38974.20	3.95	3.69	350.86
	a_s	-38974.33	-38974.33	0.00	0. 00	1.00
	b	-38972.61	-38972.61	3.43	0.00	1.00
	c	-38961.74	-38966.66	25.18**	9.84**	5.64
	d	-38966.61	-38966.61	15.44**	0.00	1.00
	e	-38964.06	-38966.16	20.54**	4.21*	16.38
	e_s	-38962.98	-38965.51	22.69**	5.07*	7.09
	f	-38973.76	-38973.76	1.14	0.00	1.00
	g	-38971.57	-38971.57	5.52	0.00	1.00
	h	-38974.33	-38974.33	0.00	0.00	1.00
	i	-38974.33	-38974.33	0.00	0.00	1.00
	j	-38974.33	-38974.33	0.00	0.00	1.00
	k	-38974.33	-38974.33	0.00	0.00	1.00
	l	-38963.72	-38969.17	21.22**	10.90**	16.17
	l_s	-38974.33	-38974.33	0.00	0.00	1.00
	m	-38971.56	-38974.33	5.53	5.53*	∞
	n	-38971.38	-38973.99	5.89	5.23*	∞
	o	-38954.79	-38974.33	39.07**	39.07**	∞
	p	-38965.94	-38972.10	16.78**	12.33**	145.19
	p_s	-38946.54	-38962.91	55.58**	32.74**	15.46
	CfTPS1	-38971.89	-38974.33	4.88	4.88*	26.71
	SmCPS1	-38974.33	-38974.33	0.00	0.00	1.00
	MvCPS3	-38971.28	-38971.51	6.10*	0.47	3.04
	SmCPS2	-38972.57	-38972.57	3.53	0.00	1.00
	SsLPS	-38962.74	-38962.85	23.19**	0.22	1.78
	CfTPS2	-38967.34	-38970.40	13.97**	6.11*	10.37
	NtCPS	-38960.37	-38965.68	27.92**	10.63**	5.75
	LsCPS	-38974.33	-38974.33	0.00	0.00	1.00
	SrCPS	-38974.15	-38974.15	0.00	0.00	1.00
	GrTPS1	-38958.80	-38959.16	31.05**	0.72	1.15
	SmCPS4	-38956.95	-38958.97	34.75**	4.03*	3.43
	IeCPS2	-38963.46	-38968.50	21.74**	10.07**	∞
	SdCPS	-38974.33	-38974.33	0.00	0.00	1.00
	SmCPS5	-38974.33	-38974.33	0.00	0.00	1.00
	IeCPS1	-38971.02	-38973.36	6.61*	4.67*	8.06

SlTPS40	-38969.79	-38974.33	9.07*	9.07**	174.92
CmCPS1	-38970.85	-38972.23	6.95*	2.76	6.92
CmCPS2	-38970.85	-38972.23	6.95*	2.76	6.92
PsaCPS	-38964.99	-38972.42	18.67**	14.86**	34.26
AtCPS	-38972.75	-38974.33	3.15	3.15	26.09
ZmAN2	-38968.66	-38972.04	11.33**	6.75**	147.98
OsCPS2	-38962.97	-38963.00	22.72**	0.05	1.37
OsCPS4	-38957.11	-38957.11	34.43**	0.00	1.44

[a]2ΔL is twice the log-likelihood difference between MA and M1a, MA and MA1. Where under M1a (nearly neutral model), lnL was estimated to be -38974.33 for the CPS tree and -33070.77 for the KSL tree.

[b]Nonsynonymous: synonymous substitution (dN／dS) ratio of site classes 2a and 2b.

[∞]The dN/dS value was estimated to be 999.00.

[*] In Test 1, the difference at the significant level of $P < 0.01$ ([**]) and $P < 0.05$ ([*]) was compared to the chi-squared distribution with d.f. = 2 and in test 2 with d. f. = 1.

158 **Table S6.** Information about primers used in this study

Name	Sequence(5'to 3')	Description
RACE primer		
CPS2-RACE5	CACGAGCCATCCGCCATTTGGTGCTGCGC	3` primer for CPS2 5`-RACE
CPS2-RACE52	CACGTCCTTGATTAGGGCCACTATC	3` primer for CPS2 RACE
CPS3-RACE5	TTCGTCTCCCCATGAGCCATCAGGAAGC	3` primer for CPS3 5`-RACE
CPS3-RACE52	GGCCACACGTCATGTGCGCTTCAGAAG	3` primer for CPS3 5`-RACE
CPS4-RACE5N	CCAAGCTGAAGAGCAGTGATGTGGGCTC	3` primer for CPS4 5`-RACE
CPS4-RACE52	GAGCGTAGGCGGTGGACGACGGCGACGA	3` primer for CPS4 5`-RACE
CPS5-RACE53	TGTACCTTCCTCCGGTGGCGTGTTGGAA	3` primer for CPS5 5`-RACE
CPS5-RACE54	TCTTGTTGGTGTAACGCCTGGCAATTG	3` primer for CPS5 5`-RACE
CPS5-RACE56	TTGTACATTCCGGTGACCGCCTGCGTCGA	3` primer for CPS5 5`-RACE
CPS5-RACE57	GCCTGTCTACTGCCCATAGATGCTCGA	3` primer for CPS5 5`-RACE
KSL2-RACE52	CCCAGTACGTTTCCGTCTCCGACGAGTA	3` primer for KSL2 5`-RACE
KSL2-RACE53N	TTCCTGCCATGCTTGATTATGCCAGAGG	3` primer for KSL2 5`-RACE
Full length cloning primers for E. coli expression		
CPS1-EcoRV	GCGATATCATGGCCTCCTTATCCTCTACAATC	5`primer for full length CPS1
CPS1-NotI	GGAAAGCGGCCGCTCACGCGACTGGCTCGAAAAGC	3`primer for full length CPS1
CPS2+BamHI	CGCGGATCCATGACCTCTTTGTTCAGTCTA	5`primer for full length CPS2
CPS2+EagI	GCCGGCCGTCATACGACCGGTTCAAAAAGTA	3`primer for full length CPS2
CPS3+EcoRV	GCGATATCATGATCTCTCTTTGCTTTTTAGG	5`primer for full length CPS1
CPS3+NotI	TGCGGCCGCTTATACAATCGGCTCAAATAGTA	3`primer for full length CPS1
CPS4+EcoRV	GCGATATCATGTCATTTGCGTCCAACGC	5`primer for full length CPS4
CPS4+NotI	TGCGGCCGCCTAGACTATTTTTTCAAACAATAC	3`primer for full length CPS4
CPS5+EcoRV	GCGATATCATGCCTCTCGCTTCCAATCCCG	5`primer for full length CPS5
CPS5+XhoI	CCGCTCGAGCGGTTACGTACATGTACATGGTCATTG	3`primer for full length CPS5
KSL1+EcoRV	GCGATATCATGTCGCTCGCCTTCAACCCGGC	5`primer for full length KSL1
KSL1+NotI	AGCGGCCGCTCATTTCCCTCTCACATTATTAGC	3`primer for full length KSL1
KSL2+EcoRV	GCGATATCATGGCGCTTCCTCTCTCCACTTG	5`primer for full length KSL2
KSL2+NotI	GGAAAGCGGCCGCTCAACCATGAAGCTTTGATAATCC	3`primer for full length KSL2
At-CPS+EcoRI	CGGAATTCATGTCTCTTCAGTATCATGTTCTAA	5`primer for full length AtCPS
At-CPS+NotI	TTGCGGCCGCCTAGACTTTTTGAAACAAGACTTTG	3`primer for full length AtCPS
At-KSL+XhoI	CCCTCGAGTCAAGTTAAAGATTCTTCCTGTA	5`primer for full length AtKS
At-KSl+NcoI	CCCATGGCTATCAACCTTCGCTCCTC	3`primer for full length AtKS
qRT-PCR primers for gene expression		
CPS1-F	CCACATCGCCTTCAGGGAAGAAAT	5`primer for qRT-PCR *CPS1*
CPS1-R	TTTATGCTCGATTTCGCTGCGATCT	3`primer for qRT-PCR *CPS1*
CPS2-F	GGTCTCATCGCCTTCAACGAAGAT	5`primer for qRT-PCR *CPS2*
CPS2-R	TCCTTATCCTTTATGCTCCCATCCA	3`primer for qRT-PCR *CPS2*
CPS3-F	GGAGATGCCAATTCGAACATCAGA	5`primer for qRT-PCR *CPS3*
CPS3-R	TCAAATATAGTTGCGGCGGCCAAA	3`primer for qRT-PCR *CPS3*
CPS4-F	CGGCTGCCTTGGGCTACAACAATA	5`primer for qRT-PCR *CPS4*
CPS4-R	TCCCTGGTGACCTCCTCCTTCCCA	3`primer for qRT-PCR *CPS4*
CPS5-F	TAGAAGATGCAGCTACTTTCTCTGCT	5`primer for qRT-PCR *CPS5*

CPS5-R	CATCATCTTCACCGCCGTACTGTT	3`primer for qRT-PCR *CPS5*
KSL1-F	CTTCCCAAGACAATGCAAAGAT	5`primer for qRT-PCR *KSL1*
KSL1-R	ATTTCCCTCTCACATTATTAGC	3`primer for qRT-PCR *KSL1*
KSL2-F	TTAGTTTTGGAGGGCAAGAAGAGTGT	5`primer for qRT-PCR *KSL2*
KSL2-R	CTCCTGTTTGGTCGTTGAGAAGAATA	3`primer for qRT-PCR *KSL2*
actin-488	AGGAACCACCGATCCAGACA	5`primer for qRT-PCR *actin*
actin-222	GGTGCCCTGAGGTCCTGTT	3`primer for qRT-PCR *actin*
Pimers for GATEWAY RNA interference construction		
CPS1-RNAi-F	AGGAGATGGGAGTAGAGGGC	5`primer for RNAi *CPS1*
CPS1-RNAi-R	TACATTTCTTTTCCCGTGGC	3`primer for RNAi *CPS1*
CPS5-RNAi-F	GCTGCCACATTCATCAACCGTTGA	5`primer for RNAi *CPS5*
CPS5-RNAi-R	GATAGTTCCTGGATTACAATAAGCTGC	3`primer for RNAi *CPS5*
Primers for genomic sequence cloning		
CPS1g-3220F	TTGGAAGGCTCTGGGCGATCGA	5`primer for *CPS1*
CPS1g-4267R	TCGTACCATCTGCTCAAGCGACA	3`primer for *CPS1*
CPS1g-4116F	TTCGAGTGGCTCTACATGCAAGA	5`primer for *CPS1*
CPS1g-3620R	TTCCTCAGCACATCTGTGTGGA	3`primer for *CPS1*
CPS1g-2560F	TTACTATCGTTCAGCCGGACTA	5`primer for *CPS1*
CPS1g-2663R	TCTCCAGTGGAATCCTGTCGTGGA	3`primer for *CPS1*
CPS1g-956F	TATAGGAGTGACGTACATCAAGGA	5`primer for *CPS1*
CPS1g-1060R	TGCACCGCCGGAGAATCGTAAGGAA	3`primer for *CPS1*
CPS1g-2560R	TAGTCCGGCTGAACGATAGTAA	3`primer for *CPS1*
CPS3-1F	ATGATCTCTCTTTGCTTTTTAGG	5`primer for *CPS3*
CPS3-615R	GGCCACACGTCATGTGCGCTTCAGAAG	3`primer for *CPS3*
CPS3-431F	TCCTGATGGCTCATGGGGAGA	5`primer for *CPS3*
CPS3-855R	TAGTGTTGTTGTCACCTGGTGCA	3`primer for *CPS3*
CPS3-650F	TACCGCAAGGGATCATAAACTCAA	5`primer for *CPS3*
CPS3-3205R	TGCCATAGATGTGTCATCAACTTC	3`primer for *CPS3*
CPS3-3183F	GAAGTTGATGACACATCTATGGCA	5`primer for *CPS3*
CPS3-3508R	TTGGCGATCCACACTGTGGTGGA	3`primer for *CPS3*
CPS3-3428F	TGGTATGCCACGTTACCACGAGTA	5`primer for *CPS3*
CPS3-4764R	GCCATGCAGTTAGCTGAAGTGGCA	3`primer for *CPS3*
CPS3-4738F	TATGCCACTTCAGCTAACTGCA	5`primer for *CPS3*
CPS3-5633R	TTATACAATCGGCTCAAATAGTA	3`primer for *CPS3*
CPS4-1F	ATGTCATTTGCGTCCAACGCC	5`primer for CPS4
CPS4-782R	GACCCATCGTCCAACTGGTTATCCGA	3`primer for CPS4
CPS4-645F	GAGGTTGAATCGGAGAAGATGAAGGA	5`primer for CPS4
CPS4-2653R	TTGACAGGCTTGAGCAGATAGTCGA	3`primer for CPS4
CPS4-2630F	TCGACTATCTGCTCAAGCCTGTCAA	5`primer for CPS4
CPS4-3452R	TATGAGTACCTCGGCAGGAAGGTC	3`primer for CPS4
CPS4-3249F	TTCTGGTGTATGGCAGGGCAGA	5`primer for CPS4
CPS4O-R	CTAGACTATTTTTTCAAACAATAC	3`primer for CPS4

159

160

Parsed Citations

Arnold K, Bordoli L, Kopp J, Schwede T. (2006). The SWISS-MODEL workspace: a Web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Brückner K, Bozic D, Manzano D, Papaefthimiou D, Pateraki I, Scheler U, Ferrer A, de Vos RC, Kanellis AK, Tissier A. (2014). Characterization of two genes for the biosynthesis of abietane-type diterpenes in rosemary (Rosmarinus officinalis) glandular trichomes. Phytochemistry 101: 52-64.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Caniard A, Zerbe P, Legrand S, Cohade A, Valot N, Magnard JL, Bohlmann J, Legendre L. (2012). Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in Salvia sclarea (L.) and their relevance for perfume manufacture. BMC Plant Biol 12: 119.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cheng Q, Su P, Hu Y, He Y, Gao W, Huang LQ. (2014). RNA interference- mediated repression of SmCPS (copalyl diphosphate synthase) expression in hairy roots of Salvia miltiorrhiza causes a decrease of tanshinones and sheds light on the functional role of SmCPS. Biotechnol Lett 36: 363-369.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Christianson DW. (2006). Structural biology and chemistry of the terpenoid cyclases. Chem Rev 106: 3412-3442.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Criswell J, Potter K, Shephard F, Beale MB, Peters RJ. (2012). A single residue change leads to a hydroxylated product from the class II diterpene cyclization catalyzed by abietadiene synthase. Org Lett 14: 5828-5831.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Demetzos C, Kolocouris A, Anastasaki T. (2002). A simple and rapid method for the differentiation of C-13 manoyl oxide epimers in biologically important samples using GC-MS analysis supported with NMR spectroscopy and computational chemistry results. Bioorg. Med Chem Lett 12: 3605-3609.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Efferth T, Kahl S, Paulus K, Adams M, Rauh R, Boechzelt H, Hao XJ, Kaina B, Bauer R. (2008). Phytochemistry and pharmacogenomics of natural products derived from traditional Chinese medicine and chinese materia medica with activity against tumor cells. Mol Cancer Ther 7: 152-161.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gao J, Yang G, Pi R, Li R, Wang P, Zhang H, Le K, Chen S, Liu P. (2008). Tanshinone IIA protects neonatal rat cardiomyocytes from adriamycin-induced apoptosis. Transl Res 151: 79-87.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gao W, Hillwig ML, Huang LQ, Cui GH, Wang XY, Kong JQ, Yang B, Peters RJ. (2009). A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org Lett 11: 5170-5173.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gong Y, Li YL, Lu Y, Li LL, Abdolmaleky H, Blackburn GL, Zhou JR. (2010). Bioactive tanshinones in Salvia miltiorrhiza inhibit the growth of prostate cancer cells in vitro and in mice. Int J Cancer 129: 1042-1052.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307-321.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Guo AY, Zhu QH, Chen X, Luo JC. (2007). GSDS: a gene structure display server. Yi Chuan 29: 1023-1026.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Guo J, Zhou YJ, Hillwig ML, Shen Y, Yang L, Wang YJ, Zhang XN, Liu WJ, Peters RJ, Chen XY, Zhao ZB, Huang LQ. (2013). CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. Proc Natl Acad Sci USA 110:12108-12113.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hang L,Wang JR, Yang DF, Shu ZM, Liang ZS. (2008). Distribution traits of bioactives in different parts of Salvia miltiorrhiza Bunge. and Salvia miltiorrhiza Bunge. f. alba. J Northwest A&F Univ (Nat Sci Ed) 36(12): 217-222.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hillwig ML, Xu MM, Toyomasu T, Tiernan MS, Wei G, Cui GH, Huang LQ, Peters RJ. (2011). Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J 68: 1051-1060.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Katoh K, Kuma K, Toh H, Miyata T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33: 511-518.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Keeling CI, Weisshaar S, Lin RP, Bohlmann, J. (2008). Functional plasticity of paralogous diterpene synthases involved in conifer defense. Proc Natl Acad Sci USA 105: 1085-1090.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Köksal M, Hu HY, Coates RM, Peters RJ, Christianson DW. (2012). Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. Nat Chem Bio 7: 431-433.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Köksal M, Potter K, Peters RJ, Christianson, DW. (2014).1.55Å-resolution structure of ent-copalyl diphosphate synthase and exploration of general acid function by site-directed mutagenesis. Biochim Biophys Acta 1840: 184-190.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lee WY, Chi LC, Yeung JH. (2008). Cytotoxicity of major tanshinones isolated from Danshen (Salvia miltiorrhiza) on HepG2 cells in relation to glutathione perturbation. Food Chem Toxicol 46: 328-338.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Limpens E, Mirabella R, Fedorova E, Franken C, Franssen H, Bisseling T, Geurts R (2005). Formation of organelle-like N2-fixing symbiosomes in legume root nodules is controlled by DMI2. Proc Natl Acad Sci USA 102: 10375-10380.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ma YM, Yuan LC, Wu B, Li XE, Chen SL, Lu SF. (2012). Genome-wide identification and characterization of novel genes involved in terpenoid biosynthesis in Salvia miltiorrhiza. J Exp Bot 63: 2809-2823.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mann FM, Prisic S, Davenport EK, Determan MK, Coates RM, Peters RJ. (2010). A single residue switch for Mg(2+)-dependent inhibition characterizes plant class II diterpene cyclases from primary and secondary metabolism. J Biol Chem 285: 20558-20563.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Margis-Pinheiro M, Zhou XR, Zhu QH, Dennis ES, Upadhyaya NM. (2005). isolation and characterization of a Ds-tagged rice (Oryza sativa L.) GA-responsive dwarf mutant defective in an early step of the gibberellin biosynthesis pathway. Plant Cell Rep 23: 819-833.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Morrone D, Xu M, Fulton DB, Determan MK, Peters RJ. (2008). Increasing complexity of a diterpene synthase reaction with a single residue switch. J Am Chem Soc 130: 5400-5401.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nakao M, Fukushima T. (1934). On the chemical composition of Salvia miltiorrhiza (Chinese drug Tan-Shen). J Pharm Soc Jpn 54: 154-162.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Møller BL, Bohlmann J, Hamberger B. (2014). Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in Coleus forskohlii. Plant Physiol 164: 1222-1236.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Potter K, Criswell J, Zi J, Stubbs A, Peters RJ. (2014). Novel product chemistry from mechanistic analysis of ent-copalyl diphosphate synthases from plant hormone biosynthesis. Angew Chem Int Ed 53: 7198-7202.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Prisic S, Peters RJ. (2007). Synergistic substrate inhibition of ent-copalyl Diphosphate synthase: a potential feed-forward inhibition mechanism limiting gibberellin metabolism. Plant Physiol 144: 445 - 454.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Prisic S, Xu MM, Wilderman PR, Peters RJ. (2004). Rice contains two disparate ent-copalyl diphosphate synthases with distinct metabolic functions. Plant Physiol 136: 4228-4236.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sallaud C, Giacalone C, Töpfer R, Goepfert S, Bakaher N, Rösti S, Tissier A. (2012). Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (Nicotiana tabacum) glandular trichomes. Plant J 72: 1-17.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sana TR, Roark JC, Li XD, Waddell K, Fischer SM. (2008). Molecular formula and METLIN personal metabolite database matching applied to the identification of compounds generated by LC/TOF-MS. J Biom Tech 19: 258-266.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L. (2012). Toward a biosynthetic route to sclareol and amber odorants. J Am Chem Soc 134: 18900-18903.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sun DD, Wang HC, Wang XB, Luo Y, Jin ZX, Li ZC, Li GR, Dong MQ. (2008). Tanshinone IIA: a new activator of human cardiac KCNQ1/KCNE1 (IKs) potassium channels. Eur J pharmacol 590: 317-321.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M. (2010). A practical approach to RT-qPCR publishing data that conform to the MIQE guidelines. Methods 50: S1-S5.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Trapp SC, Croteau RB. (2001). Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics 158: 811-832.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Trott O, Olson AJ. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31: 455-461.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang L, Zhou GB, Liu P, Song JH, Liang Y, Yan XJ, Xu F, Wang BS, Mao JH., Shen ZX, Chen SJ, Chen Z. (2008). Dissection of mechanisms of Chinese medicinal formula Realgar- Indigo naturalis as an effective treatment for promyelocytic leukemia. Proc Natl Acad Sci USA 105: 4826-4831.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wilderman PR, Peters RJ. (2007). A single residue switch converts abietadiene synthase into a pimaradiene specific cyclase. J Am Chem Soc 129: 15736-15737.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Xu MM, Wilderman PR, Peters RJ. (2007). Following evolution's lead to a single residue switch for diterpene synthase product outcome. Proc Natl Acad Sci USA 104: 7397-7401.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yan YP, Wang ZZ. (2007). Genetic transformation of the medicinal plant Salvia miltiorrhiza by agrobacterium tumefaciens-mediated method. Plant Cell Tiss Orga Cult 88: 175-184.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang M, Liu AH, Guan SH, Sun JH, Xu M, Guo DA. (2006). Characterization of tanshinones in the roots of Salvia miltiorrhiza (Dan-shen) by high-performance liquid chromatography with electrospray ionization tandem mass spectrometry. Rapid Commun Mass Spectrom 20: 1266-1280.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang YF, Hou S, Cui GH, Chen SL, Wei JH, Huang LQ. (2010). Characterization of reference genes for quantitative real-time PCR analysis in various tissues of Salvia miltiorrhiza. Mol Biol Rep 37: 507-513.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang Z, Reis DM. (2011). Statistical properties of the branch-site test of positive selection. Mol Biol Evol 28: 1217-1228.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang Z, Wong WS, Nielsen R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22: 1107-1118.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yang Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24: 1586-1591.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zerbe P, Hamberger B, Yuen M, Chiang A, Sandhu H, Madilao L, Nguyen A, Hamberger B, Bach SS, Bohlmann J. (2013). Gene discovery of modular diterpene metabolism in nonmodel systems. Plant Physiol 162:1073-1091.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang C, Gschwend AR, Ouyang Y, Long MY. (2014). Evolution of gene structural complexity: an alternative-splicing-based model accounts for intron-containing retrogenes. Plant Physiol 165: 412-423.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang J, Nielsen R, Yang Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22: 2472-2479.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang J. (2004). Frequent false detection of positive selection by the likelihood method with branch-site models. Mol Biol Evol 21: 1332-1339.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang N, Zeng LP, Shan HY, Ma H. (2012). Highly conserved low-copy nuclear genes as effective 1157 markers for phylogenetic analyses in angiosperms. New Phytol 195: 923-937.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhao JL, Mou Y, Li PQ, Wu JY, Zhou LG. (2011). Diterpenoid tanshinones and phenolic acids from cultured gairy roots of Salvia miltiorrhiza Bunge and their antimicrobial activities. Molecules 16: 2259-2267.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhong GX, Li P, Zeng LJ, Guan J, Li DQ, Li SP. (2009). Chemical characteristics of Salvia miltiorrhiza (Danshen) collected from different locations in China. J Agric Food Chem 57: 6879-6887.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhou Y, Xu G, Choi FFK, Ding LS, Han QB, Song JZ, Qiao CF, Zhao QS, Xu HX. (2009). Qualitative and quantitative analysis of diterpenoids in Salvia species by liquid chromatography coupled with electrospray ionization quadrupole time-of-flight tandem mass spectrometry. J Chromatogr A 1216: 4847-4858.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhou, L, Zuo, Z, Chow, MS. (2005). Danshen: an overview of its chemistry, pharmacology, pharmacokinetics, and clinical use. J Clin Pharmacol 45: 1345-1359.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zi J, Mau S, Peters RJ (2014). To gibberellins and beyond! Surveying the evolution of (Di)terpenoid metabolism. Annu Rev Plant Biol 65: 259-286.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

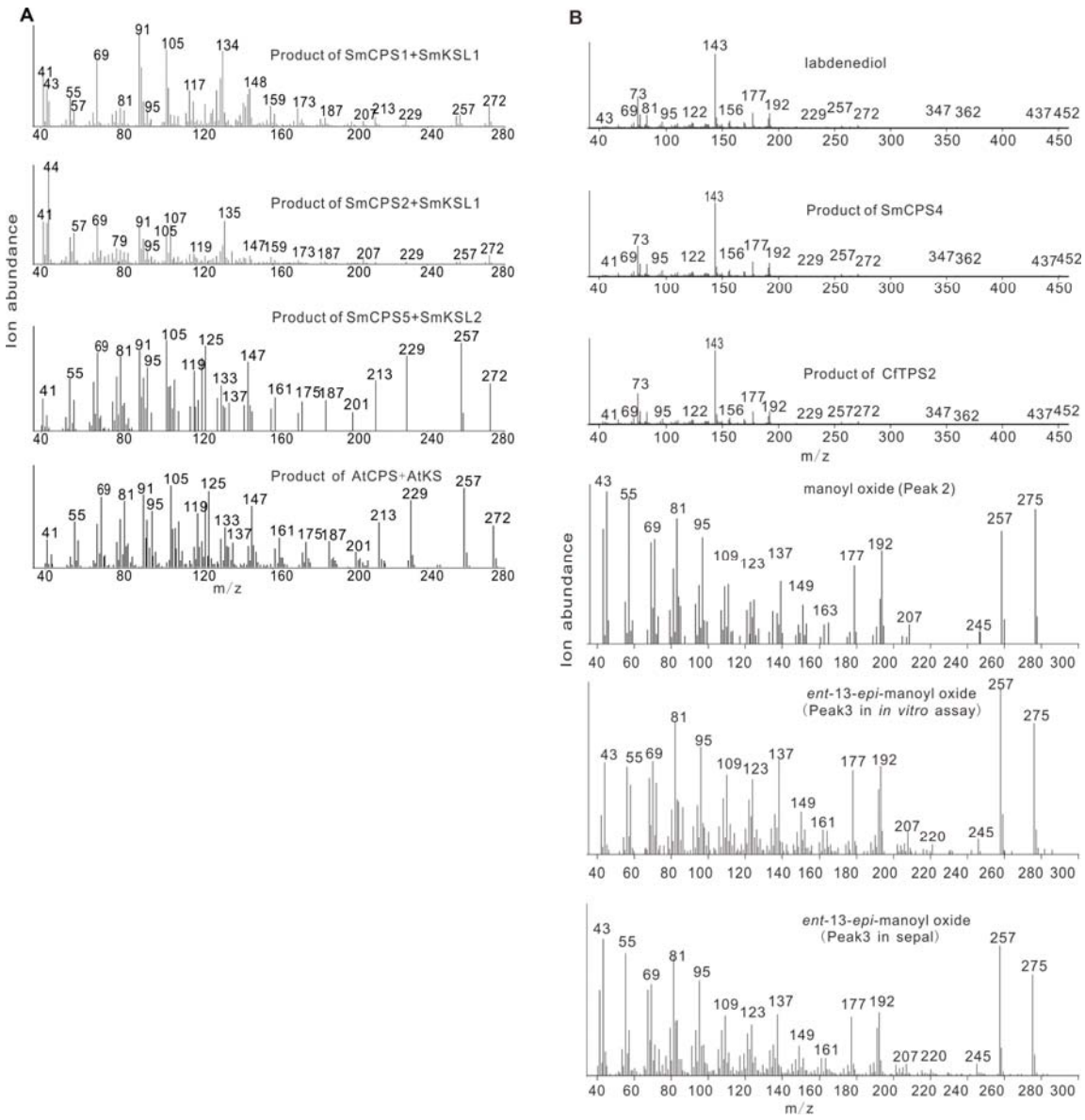


Figure S1. Electron impact (EI) mass spectrum of diterpenes in this study.

(A) Miltiradiene and kaurene were comparison with the known enzyme SmCPS1, SmKSL1 and AtCPS, AtKS. (B) Labdenediol were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes. The only difference between peak 2 and 3 is the ratio of intensities of peaks m/z 275:257.

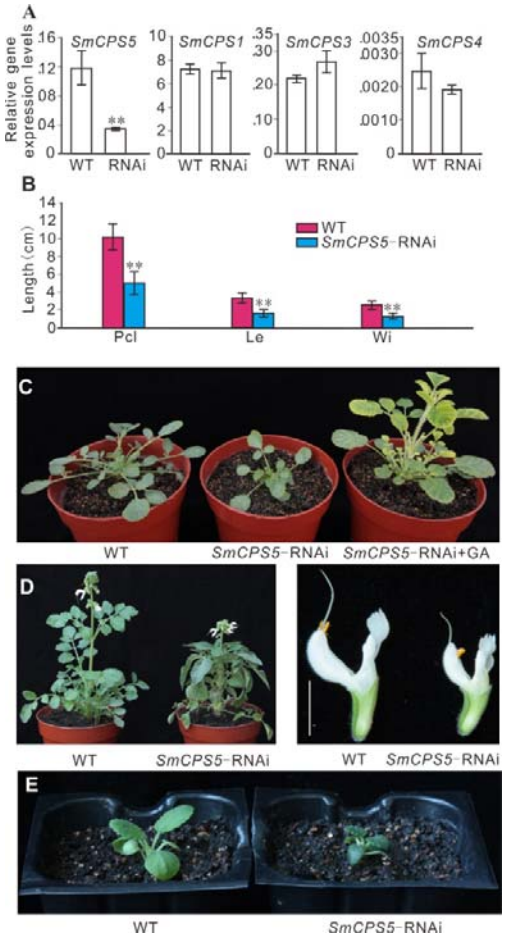
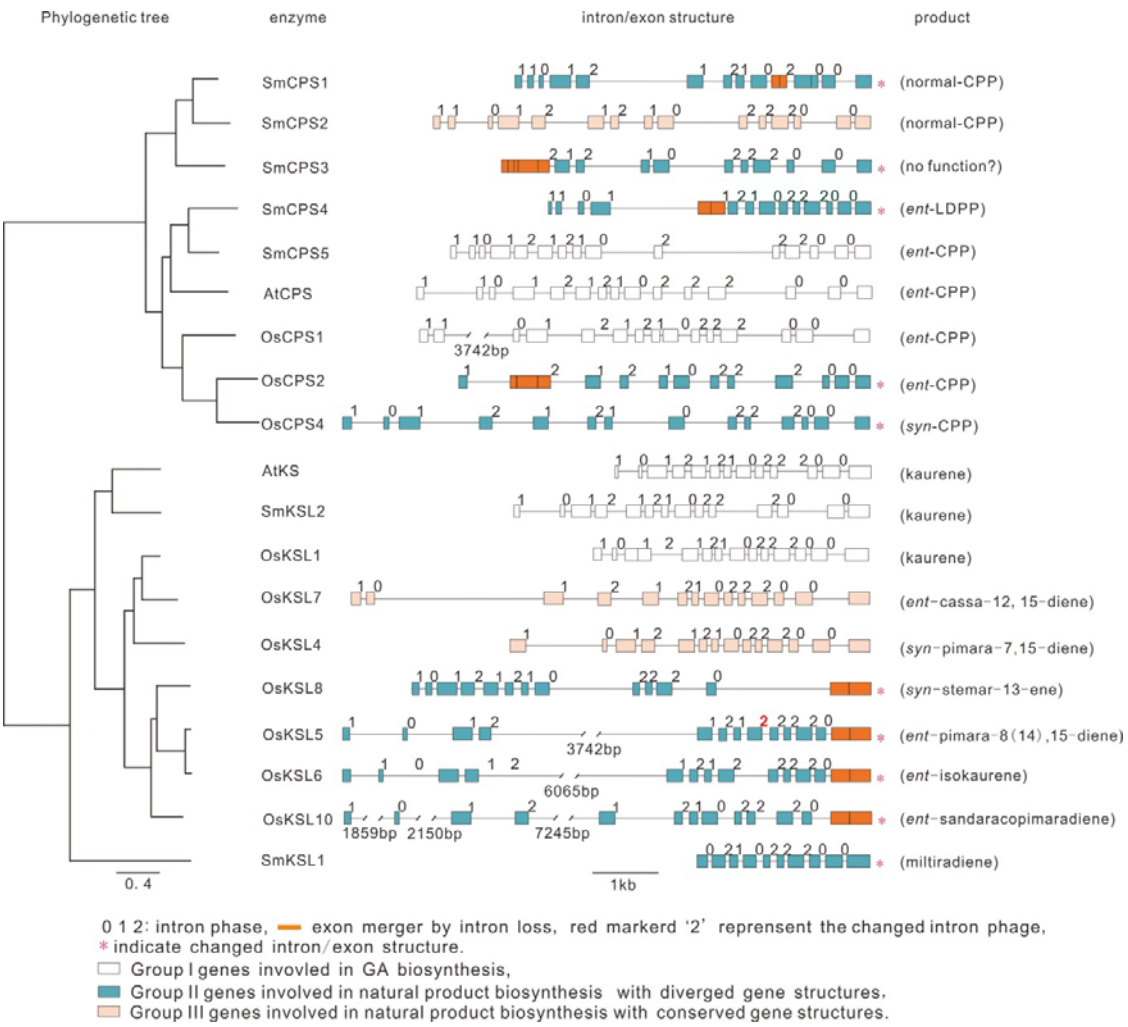


Figure S2. The dwarf phenotype caused by downregulation of *SmCPS5*. (A) qRT-PCR analysis of transcript levels of five *CPS* genes in *SmCPS5*-RNAi line and WT in hairy root. Expression was normalized to that of *Actin*. Data are means ± SD (n = 5). Asterisk indicates significant difference at $P < 0.01$ compared with WT by Student's *t* test. *SmCPS2* is not expressed in these samples. (B) The length of pinnately compound leaf (Pcl), the length (Le) and width (Wi) of the top leaf are all significantly shorter in of T_0 generation plants than WT. Data are means ± SD (n = 30). (C) The dwarf phenotype of *SmCPS5*-RNAi and the recovery phenotype at 10d after exogenous GA3 treatment in of T_0 generation plants. (D) *SmCPS5*-RNAi T_0 generation plants with exogenous GA3 treatment at flowering stage. Bar = 1cm. (E) T1 generation of *SmCPS5*-RNAi shows characteristic dwarf phenotype.



Supplemental Figure S3. The intron/exon structure divergence of diTPS genes involved in labdane-related diterpenoid biosynthesis in *Arabidopsis thaliana*, rice and *S. miltiorrhiza.* The phylogenetic relationship was reconstructed using the JTT model by PhyML 3.0. The product of each enzyme was given in parenthesis. *At-Arabidopsis thaliana, Os*-rice and *Sm-S. miltiorrhiza.*

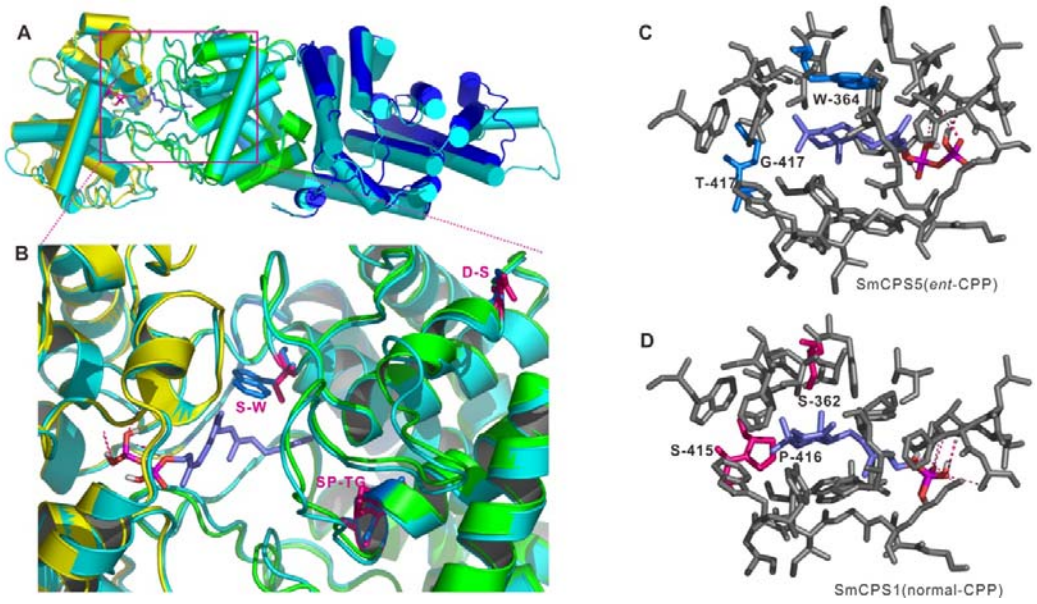


Figure S4. Homology modelling and molecular docking result show the positive selection sites. (A-B) Homology modelling of SmCPS1 and SmCPS5 based on 3pya (A) and the partial enlarged view of the interface of the βγ domains in the purple square frame (B). The α，β and γ domain of SmCPS5 is blue, green and yellow. In SmCPS1, all atoms are cyan. (C-D) The reaction cavity of *ent*-CPP into SmCPS5 (C) and normal-CPP into SmCPS1 by molecular docking (D). Residues with blue and purple lines in SmCPS5 and SmCPS1 show positive selection sites, respectively. The ligand is the substrate GGPP.

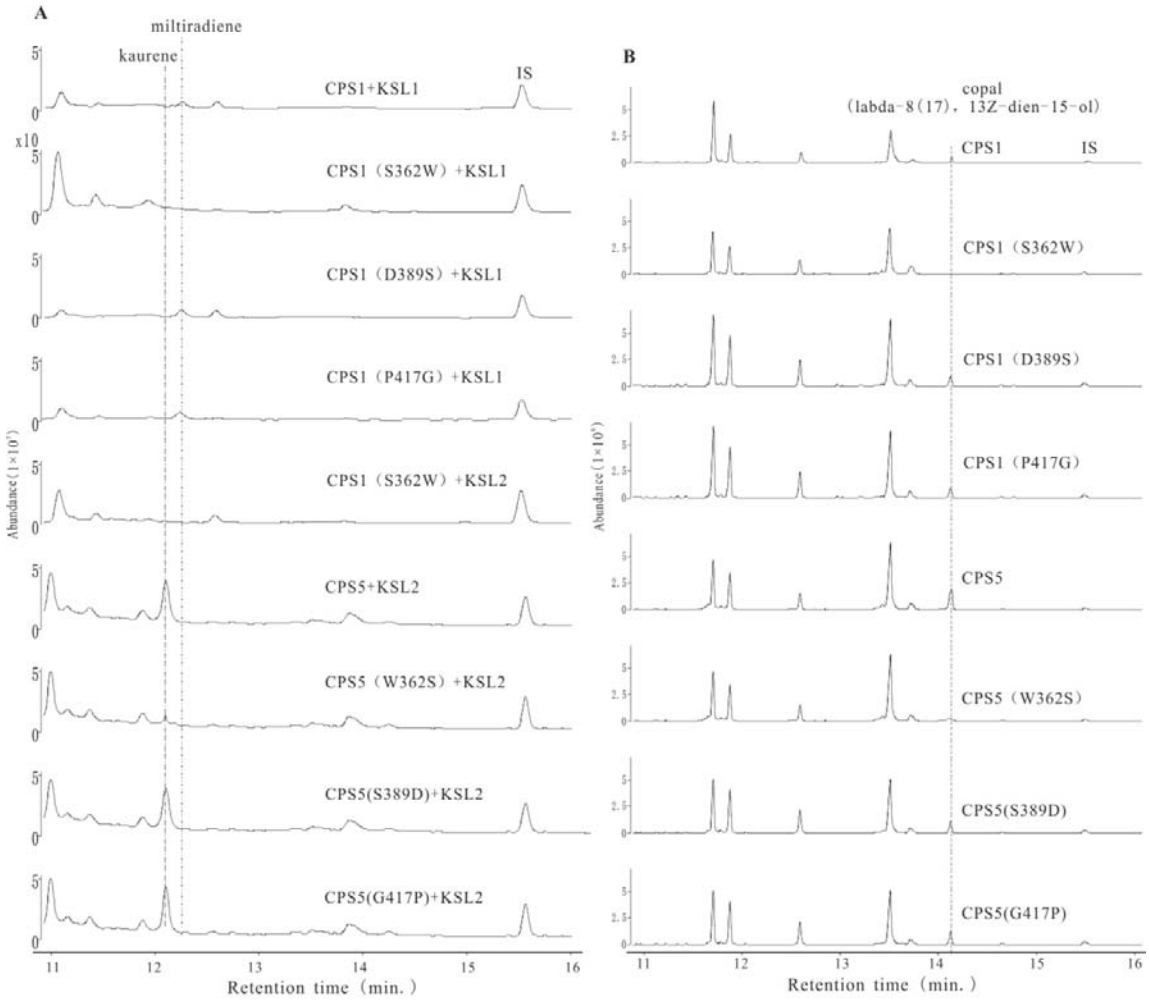


Figure S5. GC-MS analysis of the six mutant enzymes activity of SmCPS1 and SmCPS5.

(A) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 combine SmKSL1 or SmKSL2 with GGPP as substrate. (B) Total ion chromatography of the recombinant mutant enzyme of SmCPS1 and SmCPS5 with GGPP as substrate only. Samples were derivatized with 80 μL N-methyl-N-(*trimethylsilyl*) *trifluoroacetamide* (*MSTFA*) and 8 μL pyridine at 80°C for 40 minutes.

Table S1. Information of diTPS gene family in *S. miltiorrhiza*

Gene	Accession No.	Length[a]			Function annotation	reference
		cDNA (bp)	Protein	gDNA (bp)		
CPS1	KC814639	2382	793aa	5422	copalyl diphosphate synthase	Gao et al.(2009)
CPS2	KC814640	2397	798aa	6633	copalyl diphosphate synthase	This study
CPS3	KC814641	2385	794aa	5619	no function	This study
CPS4	KP063138	2331	776aa	4908	labd-13-en-8-ol diphosphate synthase	This study
CPS5	KC814642	2382	793aa	6256	*ent*-copalyl diphosphate synthase	This study
KSL1	EF635966	1788	595aa	2667	miltiradiene synthase	Gao et al.(2009)
KSL2	KC814643	2424	807aa	5405	kaurene synthase	This study

[a] The cDNA sequences of CPS1 to CPS5 and KSL2 are from bh2-7, genomic sequence of CPS1, CPS3 and CPS4 are from bh2-7, and the other sequences from genomic sequencing material published by Ma et al., 2012.

Supplemental Table 2. List of plant DiTPS involved in this paper

Gene	Species	GenBank Accession	Funtional Annotation*	Ref.
AgAS	*Abies grandis*	AAB05407	abietadiene synthase	Vogel et al.,1996
AtCPS	*Arabidopsis thaliana*	AAA53632	*ent*-copalyldiphosphatesynthas	Sun et al., 1994
AtKSL	*A. thaliana*	AAC39443	*ent*-kaurene synthase	Yamaguchi et al., 1998
CcCLS	*Cistus creticus*	ADJ93862	copal-8-ol diphosphate synthase	Falara et al., 2010
CmCPS1	*Cucurbita maxima*	AAD04292	*ent*-copalyldiphosphate synthase	Smith et al.,1998
CmCPS2	*C. maxima*	AAD04293	*ent*-copalyldiphosphate synthase	Smith et al.,1998
CmKS	*C. maxima*	AAB39482	*ent*-kaurene synthase	Yamaguchi et al., 1996
GbLS	*Ginkgo biloba*	AAL09965	levopimaradiene synthase	Schepmannet al., 2001
IeCPS1	*Isodon eriocalyx*	AEP03177	*ent*-copalyldiphosphate synthase	Li et al., 2012
IeCPS2	*I. eriocalyx*	AEP03175	*ent*-copalyldiphosphate synthase	Li et al., 2012
OsCPS1	*Oryza sativa*	BAD42449	*ent*-copalyldiphosphate synthase	Otomo et al., 2004
OsCPS2	*O. sativa*	AAT11021	*ent*-copalyldiphosphate synthase	Prisic et al., 2004
OsCPS4	*O. sativa*	AAS98158	*syn*-copalyldiphosphate synthase	Xu et al., 2004
OsKS1	*O. sativa*	AAQ72559	ent-kaurene synthase	Margis-Pinheiro et al., 2005
OsKSL4	*O. sativa*	AAU05906	pimara-7,15-diene synthase	Wilderman et al., 2004
OsKSL5	*O. sativa*	Q6Z5J6	*ent*-pimara-8(14),15-diene synthase	Margis-Pinheiro et al., 2005
OsKSL6	*O. sativa*	A4KAG8	*ent*-isokaur-15-ene synthase	Xu et al., 2007
OsKSL7	*O. sativa*	BAC56714	*ent*-cassa-12,15-diene synthase	Cho et al., 2004
OsKLS8	*O. sativa*	Q6BDZ9	stemar-13-ene synthase	Nemoto et al., 2004
OsKSL10	*O. sativa*	Q2QQJ5	*ent*-sandaracopimara-8(14),15-diene synthase	Peters. 2006
OsKSL11	*O. sativa*	Q1AHB2	stemod-13(17)-ene synthase	Morrone et al., 2006
PaIso	*Picea abies*	AAS47690	isopimaradiene synthase	Martin et al., 2004
PaLAS	*P. abies*	AAS47691	levopimaradiene/abietadiene synthase	Martin et al., 2004
PsIso	*P. sitchensis*	ADZ45512	Isopimaradiene synthase	Keeling et al., 2011
PsLAS	*P. sitchensis*	ADZ45517	Levopimaradiene/abietadiene synthase	Keeling et al., 2011
PgCPS	*P. glauca*	ADB55707	*ent*-copalyldiphosphate synthase	Keeling et al., 2010
PgKS	*P. glauca*	ADB55708	*ent*-kaurene synthase	Keeling et al., 2010
PsCPS	*P. sitchensis*	ADB55709	*ent*-copalyldiphosphate synthase	Keeling et al., 2010
PsKS	*P. sitchensis*	ADB55710	*ent*-kaurene synthase	Keeling et al., 2010
PpCPS/KS	*Physcomitrella patens*	BAF61135	*ent*-kaurene synthase	Hayashi et al., 2006
PsaCPS	*Pisum sativum*	AAB58822	Bifunctional *ent*-copalyldiphosphate synthase	Ait-Ali et al.,1997
SdCPS	*Scoparia dulcis*	BAD91286	*ent*-copalyldiphosphate synthase	Nakagiri et al., 2005
SmCPS1	*S. miltiorrhiza*	KC814639	copalyl diphosphate synthase	Gao et al., 2009
SmCPS2	*S. miltiorrhiza*	KC814640	copalyl diphosphate synthase	this study
SmCPS3	*S. miltiorrhiza*	KC814641	No function	this study

SmCPS4	*S. miltiorrhiza*		copal-8-ol diphosphate synthase	this study
SmCPS5	*S. miltiorrhiza*	KC814642	*ent*-copalyldiphosphate synthase	this study
SmKSL1	*S. miltiorrhiza*	KC814643	miltiradiene synthase	Gaoet al.,2009
SmKSL2	*S. miltiorrhiza*	ABV08817	*ent*-kaurene synthase, *ent*-13-*epi*-manoyl oxide synthase	this study
SrCPS	*Stevia rebaudiana*	AAB87091	*ent*-copalyldiphosphate synthase	Richman et al., 1999
SrKS1	*S. rebaudiana*	AAD34294	*ent*-kaurene synthase	Richman et al., 1999
SrKS2	*S. rebaudiana*	AAD34295	*ent*-kaurene synthase	Richman et al., 1999
TcTS	*Taxus brevifolia*	AAC49310	taxadiene synthase	Wildung and Croteau.1996
ZmAn1	*Zea mays*	AAA73960	*ent*-copalyldiphosphate synthase	Bensen et al., 1995
ZmAn2	*Z. mays*	AAT70083	*ent*-copalyldiphosphate synthase	Bensen et al., 1995
HvCPS	*Hordeum vulgare*	AAT49065	*ent*-copalyldiphosphate synthase	Spielmeyer et al., 2004
HvKSL1	*H. vulgare*	AAT49066	*ent*-kaurene synthase	Spielmeyer et al., 2004
SsLPS	*Salvia sclarea*	*AET21247*	copal-8-ol diphosphate synthase	Schalk et al., 2012
SsScS	*S. sclarea*	AET21246	sclareol synthase	Schalk et al., 2012
NtCPS2	*Nicotiana tabacum*	CCD33018	copal-8-ol diphosphate synthase	Christopheet al., 2012
NtABS	*N. tabacum*	CCD33019	cis-abienol synthase	Christophe et al., 2012
TaKSL5	*Triticum aestivum*	BAL41692	nerolidol synthase	Hillwig et al., 2011
ZmTPS1	*Zea mays*	NP_001105097	sesquiterpene synthase	Schnee et al., 2002
AbCAS	*Abies balsamea*	AEL99953	cis-abienol synthase	Zerbe et al., 2012
SlPHS	*Solanum lycopersicum*	ACO56896	phellandrene synthase	Schilmiller et al., 2009
ShSBS	*S. habrochaites*	B8XA41	santalene and bergamotene synthase	Sallaud et al., 2009
PtKS	*Populus trichocarpa*	XP_002311286	*ent*-kaurene synthase	Tuskan et al., 2006
LsKS	*Lactuca sativa*	BAB12441	*ent*-kaurene synthase	Sawada et al., 2008
LsCPS	*L. sativa*	BAB12440	*ent*-kaurene synthase	Sawada et al., 2008
GrTPS1	*Grindelia robusta*	AGN70886	13-labden-8,15-diol pyrophosphate synthase	Zerbe et al., 2013
GrTPS6	*G. robusta*	AGN70887	manoyl oxide	Zerbe et al., 2013
EpTPS1	*Euphorbia peplus*	KC702395	*ent*-kaurene synthase	Zerbe et al., 2013
CfTPS14	*C. forskohlii*	AGN70881	*ent*-kaurene synthase	Zerbe et al., 2013
EpTPS7	*Euphorbia peplus*	AGN70883	*ent*-copalyldiphosphate synthase	Zerbe et al., 2013
SlTPS40	*Solanum lycopersicum*	JN412074	*ent*-copalyldiphosphate synthase	Falara et al., 2011
SlTPS14	*S. lycopersicum*	JN412091	*ent*-kaurene synthase	Falara et al., 2011
CfTPS1	*Coleus forskohlii*	KF444506	copalyldiphosphate synthase	Pateraki et al., 2014
CfTPS2	*C. forskohlii*	KF444507	copal-8-ol diphosphate synthase	Pateraki et al., 2014
CfTPS3	*C. forskohlii*	KF444508	(13R) manoyl oxide synthase	Pateraki et al., 2014
CfTPS4	*C. forskohlii*	KF444509	miltiradiene synthase	Pateraki et al., 2014
MvCPS1	*Marrubium vulgare*	KJ584450	peregrinol diphosphate synthase	Zerbe et al., 2014
MvCPS3	*M. vulgare*	KJ584452	copalyl diphosphate synthase	Zerbe et al., 2014
MvCPS4	*M. vulgare*	KJ584453	*ent*-kaurene synthase	Zerbe et al., 2014

MvCPS5	*M. vulgare*	KJ584454	9,13-epoxy-labd-14-en synthase	Zerbe et al., 2014
SmMDS	*Selaginella moellendorffii*	AB668998	miltiradiene synthase	Sugai et al., 2011

*Functional annotation is based on the main terpenoid product(s) of recombinant enzymes expressed in *E. coli*. Many TPSs produced multiple products.

Supplemental References

Ait-Ali T, Swain SM, Reid JB, Sun TP, Kamiya Y. (2003). The LS locus of pea encodes the gibberellin biosynthesis enzyme *ent*-kaurene synthase A. Plant J. 11(3): 443-454.

Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, Briggs SP. (1995). Cloning and characterization of the maize An1 gene. Plant Cell 7(1): 75-84.

Cho EM, Okada A, Kenmoku H, Otomo K, Toyomasu T, Mitsuhashi W, Sassa T, Yajima A, Yabuta G, Mori K, Oikawa H, Toshima H, Shibuya N, Nojiri H, Omori T, Nishiyama M, Yamane H. (2004). Molecular cloning and characterization of a cDNA encoding ent-cassa-12, 15-diene synthase, a putative diterpenoidphytoalexin biosynthetic enzyme, from suspension-cultured rice cells treated with a chitin elicitor. Plant J. 37(1): 1-8.

Falara V, Pichersky E, Kanellis AK. (2010). A copal-8-ol diphosphate synthase from the angiosperm *Cistus creticus* subsp. creticus is a putative key enzyme for the formation of pharmacologically active, oxygen-containing labdane-type diterpenes. Plant physiol 154(1): 301-310.

Falara V, Akhtar TA, Nguyen TT, Spyropoulou EA, Bleeker PM, Schauvinhold I, Pichersky E. (2011). The tomato terpene synthase gene family. *Plant physiol 157*(2): 770-789.

Gao W, Hillwig ML, Huang L, Cui G, Wang X, Kong JQ, Yang B, Peters RJ. (2009). A functional genomics approach to tanshinone biosynthesis provides stereochemical insights. Org Lett 11(22): 5170-5173.

Hayashi KI, Kawaide H, Notomi M, Sakigi Y, Matsuo A, Nozaki H. (2006). Identification and functional analysis of bifunctional *ent*-kaurene synthase from the moss Physcomitrella patens. FEBS Lett 580(26): 6175-6181.

Hillwig ML, Xu M, Toyomasu T, Tiernan MS, Wei G, Cui G, Huang LQ, Peters RJ. (2011). Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J. 68(6): 1051-1060.

Keeling CI, Dullat HK, Yuen M, Ralph SG, Jancsik S, Bohlmann J. (2010). Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. Plant Physiol 152(3): 1197-1208.

Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J. (2011). Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). BMC Plant Boil 11(1): 43.

Li JL, Chen QQ, Jin QP, Gao J, Zhao PJ, Lu S, Zeng Y. (2012). IeCPS2 is potentially involved in the biosynthesis of pharmacologically active Isodon diterpenoids rather than gibberellin. Phytochemistry 76: 32-39.

Margis-Pinheiro M, Zhou XR, Zhu QH, Dennis ES, Upadhyaya NM. (2005). Isolation and characterization of a Ds-tagged rice (*Oryza sativa* L.) GA-responsive dwarf mutant defective in an early step of the gibberellin biosynthesis pathway. Plant Cell Rep 23(12): 819-833.

Martin DM, Fäldt J, Bohlmann J. (2004).Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiol 135(4): 1908-1927.

Morrone D, Jin Y, Xu M, Choi SY, Coates RM, Peters RJ.(2006). An unexpected diterpene cyclase from rice: functional identification of a stemodene synthase. Arch. Biochem. Biophys 448 (1-2):133-140.

Nakagiri T, Lee JB, Hayashi T. (2005). cDNA cloning, functional expression and characterization of *ent*-copalyldiphosphate synthase from *Scoparia dulcis* L. Plant Sci 169(4): 760-767.

NemotoT, Cho EM, Okada A, Okada K, Otomo K, Kanno Y, Toyomasu T, Mitsuhashi W, Sassa T, Minami E, Shibuya N, Nishiyama M, Nojiri H, Yamane H. (2004) Stemar-13-ene synthase, a diterpene cyclase involved in the biosynthesis of the phytoalexin oryzalexin S in rice. FEBS Lett 571 (1-3): 182-186.

Otomo K, Kenmoku H, Oikawa H, König WA, Toshima H, Mitsuhashi W, Yamane H, Sassa T, Toyomasu T. (2004). Biological functions of *ent*-and *syn*-copalyldiphosphate synthases in rice: key enzymes for the branch point of gibberellin and phytoalexin biosynthesis. Plant J 39(6): 886-893.

Pateraki I, Andersen-Ranberg J, Hamberger B, Heskes AM, Martens HJ, Zerbe P, Bach SS, Møller BL, Bohlmann J, Hamberger B. (2014). Manoyl oxide (13R), the biosynthetic precursor of forskolin, is synthesized in specialized root cork cells in *Coleus forskohlii*. Plant Physiol 164: 1222-1236.

Prisic S, Xu M, Wilderman PR, Peters RJ. (2004). Rice contains two disparate *ent*-copalyldiphosphate synthases with distinct metabolic functions. Plant physiol 136(4): 4228-4236.

Peters, R. J.(2006) Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. Phytochemistry 67 (21):2307-2317.

Richman AS, Gijzen M, Starratt AN, Yang Z, Brandle JE. (1999). Diterpene synthesis in Stevia rebaudiana: recruitment and up‐regulation of key enzymes from the gibberellin biosynthetic pathway. Plant J 19(4): 411-421.

Sallaud C, Rontein D, Onillon S, Jabes F, Duffe P, Giacalone C, Thoraval S, Escoffier C, Herbette G, Leonhardt N, Causse M, Tissier A. (2009). A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell 21 (1):301-317.

Sallaud C, Giacalone C, Töpfer R, Goepfert S, Bakaher N, Rösti S, Tissier A. (2012). Characterization of two genes for the biosynthesis of the labdanediterpene Z‐abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. Plant J 72(1): 1-17.

SawadaY, Katsumata T, Kitamura J, Kawaide H, Nakajima M, Asami T, Nakaminami K, Kurahashi T, Mitsuhashi W, Inoue Y, Toyomasu T. (2008).Germination of photoblastic lettuce seeds is regulated via the control of endogenous physiologically active gibberellin content, rather than of gibberellin responsiveness. J Exp Bot 59 (12), 3383-3393.

Schalk M, Pastore L, Mirata MA, Khim S, Schouwey M, Deguerry F, Pineda V, Rocci L, Daviet L. (2012). Toward a biosynthetic route to sclareol and amber odorants. J Am Chem Soc 134: 18900-18903.

Schepmann HG, Pang J, Matsuda S. (2001). Cloning and Characterization of *Ginkgo biloba* Levopimaradiene Synthase, Which Catalyzes the First Committed Step in Ginkgolide Biosynthesis. Arch Biochem Biophys 392(2): 263-269.

Schilmiller AL, Schauvinhold I, Larson M, Xu R, Charbonneau AL, Schmidt A, Wilkerson C, Last RL, Pichersky E. (2009). Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc Natl Acad Sci USA 106(26):10865-10870.

Schnee C, Köllner TG, Gershenzon J, Degenhardt J. (2002). The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)-β-farnesene, (E)-nerolidol, and (E, E)-farnesol after herbivore damage. Plant physiol 130(4): 2049-2060.

Smith MW, Yamaguchi S, Ait-Ali T, Kamiya Y. (1998). The first step of gibberellin biosynthesis in pumpkin is catalyzed by at least two copalyl diphosphate synthases encoded by differentially regulated genes. Plant physiol 118(4): 1411-1419.

89 Spielmeyer W, Ellis M, Robertson M, Ali S, Lenton JR, Chandler PM. (2004). Isolation of gibberellin metabolic
90 pathway genes from barley and comparative mapping in barley, wheat and rice. Theor Appl Genet 109(4):
91 847-855.
92 Sugai Y, Ueno Y, Hayashi K, Oogami S, Toyomasu T, Matsumoto S, Natsume M, Nozaki H, Kawaide H.
93 (2011). Enzymatic ^{13}C labeling and multidimensional NMR analysis of miltiradiene synthesized by
94 bifunctional diterpene cyclase in *Selaginella moellendorffii*. J Biol Chem 86(50): 42840-42847.
95 Sun TP, Kamiya Y. (1994). The Arabidopsis GA1 locus encodes the cyclaseent –kaurene synthetase A of
96 gibberellin biosynthesis. Plant Cell 6(10): 1509-1518.
97 Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S,
98 Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, BoerjanW, Brun A, Brunner A,
99 Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J,
100 Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B,
101 Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A,
102 Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N,
103 Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R,
104 Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F,
105 Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, PeredaV, Peter G, Philippe R, Pilate G,
106 Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J,
107 Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K,
108 Wessler S, Yang G Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D. (2006).The genome
109 of black cottonwood, Populus trichocarpa (Torr. & Gray) Science 313 (5793):1596-1604.
110 Vogel BS, Wildung MR, Vogel G, Croteau R. (1996). Abietadiene synthase from grand fir (*Abies grandis*)
111 cDNA isolation, characterization, and bacterial expression of a bifunctionalditerpenecyclase involved in resin
112 acid biosynthesis. J Biol Chem 271(38): 23262-23268.
113 Wilderman PR, Xu M, Jin Y, Coates RM, Peters RJ. (2004). Identification of *syn*-pimara-7, 15-diene synthase
114 reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis.
115 Plant physiol 135(4): 2098-2105.
116 Wildung MR, Croteau R. (1996).A cDNA clone for taxadiene synthase, the diterpenecyclase that catalyzes the
117 committed step of taxol biosynthesis. J Biol Chem 271(16): 9201-9204.
118 Xu MM, Hillwig ML, Prisic S, Coates RM, Peters RJ. (2004). Functional identification of rice *syn*-copalyl
119 diphosphate synthase and its role in initiating biosynthesis of diterpenoidphytoalexin/allelopathic natural
120 products. Plant J 39(3): 309-318.
121 Xu MM, Wilderman PR, Morrone D, Xu J, Roy A, Margis-Pinheiro M, Upadhyaya NM, Coates RM，Peters RJ.
122 (2007). Functional characterization of the rice kaurene synthase-like gene family. Phytochemistry 68 (3):
123 312-326.
124 Yamaguchi S, Saito T, Abe H, Yamane H, Murofushi N, Kamiya Y. (1996). Molecular cloning and
125 characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from
126 pumpkin (*Cucurbita maxima* L.).Plant J 10(2): 203-213.
127 Yamaguchi S, Sun TP, Kawaide H, Kamiya Y. (1998). The GA2 Locus of *Arabidopsis thaliana* encodes
128 *ent*-kaurene synthase of gibberellin biosynthesis. Plant Physiol 116(4): 1271-1278.
129 Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Bohlmann J. (2012).
130 Bifunctionalcis-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its
131 implications for diterpenoid fragrance production. J Biol Chem 287(15): 12121-12131.

Table S3. Different changed metabolites obtained from metabolomics analyze of WT and *CPS1*-RNAi lines by LC-qTOF-MS

No.	t_R (min)	positive ion mode			redicted molecular formula	Predicted assigned identity	MS/MS product ion (m/z)	relative content (ug/g) FW		Percent (%)[b]
		$[M+H]^+$	$[M+Na]^+$	major fragment				WT+SE	RNAi+SE	
1	5.72	313.1070	335.0896	647.1890	C18H16O5	tanshindiol B[a]	295.0969, 277.0860, 267.1024, 249.0914	37.4 ± 7.0	n.d.	99.6
2	5.87	313.1071	335.0896	647.1892	C18H16O5	tanshindiol C[a]	295.0965, 277.0864, 267.1021, 249.0913	270.3 ± 46.9	n.d.	98.9
3	6.28	313.1073	335.0896	647.1889	C18H16O5	tanshindiol A[a]	295.0969, 277.0865, 267.1024, 249.0902	47.3 ± 8.8	n.d.	99.2
4	6.89	343.1547	365.1378	-	C20H22O5	prineoparaquinone*	-	41.6 ± 5.2	n.d.	98.9
5	6.95	295.0968	317.0796	611.1683	C18H14O4	3α-hydroxymethylenetanshinquinone[a]	277.0864, 267.1022, 251.0709, 249.0912, 221.0965	204.1 ± 30.1	n.d.	99.2
6a	7.03	297.1122	319.0952	615.1986, 342.1698	C18H16O4	tanshinol B[a]	279.1016, 261.0911, 233.0963, 218.0717, 205.1017, 190.0774	768.0 ± 105.6	7.0 + 3.1	99.4
7	7.16	281.1179	303.1005	263.1078	C18H16O3	danshenxinkun B	263.1078, 248.0837, 235.1121, 220.0887, 192.0929	191.9 ± 24.5	2.2 + 0.7	98.9
8	7.32	311.1280	333.1103	643.2302, 356.1859	C19H18O4	tanshinone IIB[a]	293.1164, 275.1070, 265.0852, 251.1075, 247.1127	2088.3 ± 219.8	40. + 17.8	94.5
9	7.37	283.1338	305.1154	587.2417	C18H18O3	no hits	-	488.9 ± 1004	3.1 + 0.8	98.2
10	7.64	341.1389	363.1217	703.2527, 386.1907	C20H20O5	trijuganone C[a]	281.1175, 263.1068，235.1124, 220.0882	142.4 ± 23.7	3.1 + 0.8	98.1
11	7.85	299.1645	321.1463	344.2226, 619.3023	C19H22O3	1R-hydroxymiltirone*	281.1539, 266.1296, 263.1435, 221.0961, 193.1020	1357.3 ± 116.7	1.9 + 0.9	99.4
12	7.92	313.1439	335.1264	-	C19H20O4	3-hydroxycryptotanshinone	-	10.0 ± 1.6	n.d.	98.1

13	8.02	311.1279	333.1106	643.2289, 356.1860	C19H18O4	hydroxytanshinone IIA[a]	293.1189, 278.0941, 247.1122, 219.1179, 204.0925	200.4 ± 23.8	0.8 + 0.4	99.4
14	8.09	309.1124	331.0953	639.1989, 354.1703	C19H16O4	no hits	-	252.0 ± 9.6	3.5 + 1.2	98.6
15	8.17	299.2011	-	-	C20H26O2	5,6-dehydrosugiol	257.1541, 243.1341, 229.1228, 187.0751	8.0 ± 0.7	2.4 + 0.9	98.7
16	8.18	279.1018	301.0843	579.1778	C18H14O3	dihydrotanshinone I[a]	261.0910, 233.0965, 218.0730, 205.1015, 190.0777	1233.6 ± 156.9	84.4 + 12.0	98.3
17	8.20	273.1846	-	-	C18H24O2	przewalskin[a]	255.1745, 243.1743	11.3 ± 1.3	0.2 + 0.1	96.9
18	8.21	287.2017	-	-	C19H26O2	przewalskin Y-1	-	157.8 ± 12.0	8.5 + 1.6	98.5
19	8.26	315.1605	337.1421	-	C19H22O4	neocryptotanshinone	-	7.9 ± 1.5	n.d.	96.7
20	8.27	327.1595	349.1417	-	C20H22O4	7-hydroxy-12-methoxy-20-nor-abieta-1,5(10),7,9,12-pentaen-6,14-dione*	-	13.5 ± 1.5	0.5 + 0.2	98.3
21	8.41	339.1234	361.1051	669.2209, 384.1806	C20H18O5	methyltanshinonate[a]	279.0968, 261.0909, 233.0959, 218.0724, 190.0769	2137.8 ± 209.9	120.0 + 40.6	98.9
22	8.52	281.1175	303.0995		C18H16O3	trijuganone B[a]	263.1069, 248.0826, 235.1129, 220.0880, 192.0933	6957.5 ± 590.7	306.4 + 95.1	97.3
23	8.77	327.1589	349.1413	372.2171	C20H22O4	prionoid E*	-	105.8 ± 10.9	0.0	96.8
24	8.84	317.2118	339.1932		C20H28O3	cryptanol*	-	11.17 ± 1.4	n.d.	96.9
25	8.90	301.2169	-	-	C20H28O2	sugiol[a]	-	59.3 ± 5.0	1.1 + 0.5	95.9
26	9.10	313.1805	335.1626	647.3344, 358.2383	C20H24O3	salvisyrianone[a]*	-	24.5 ± 2.0	n.d.	99.3
27	9.13	277.0862	299.0688	575.1466, 322.1436	C18H12O3	tanshinone I[a]	249.0918, 234.0675, 221.0969, 193.1013, 178.0780	2252.3 ± 258.3	321.1 + 48.9	97.9

28	9.29	297.1490	319.1314	615.2726	C19H20O3	cryptotanshinone[a]	282.1254, 279.1384, 267.1018, 254.0945, 221.0961, 193.1019	15959.6 ± 1088.1	270.8 + 88.8	96.3
29	9.43	295.1333	317.1153	611.2409, 340.1914	C19H18O3	isotanshinone IIA	277.1224, 262.0980, 249.1274	107.1 ± 9.1	6.2 + 2.8	98.3
30	9.43	317.2118	339.1937	362.2698	C20H28O3	7β-hydroxy-8,13-abietadiene-11,12-dione*	-	92.1 ± 30.8	n.d.	97.6
31	10.06	267.1382	289.1208	555.2509, 312.1964	C18H18O2	4-methylenemiltirone[a]	249.1272, 225.0910, 207.0809	246.8 ± 47.8	2.1 + 0.9	98.9
32	10.44	329.2458	-		C22H32O2	pomiferin G*	-	12.0 ± 1.2	n.d.	96.8
33	10.62	295.1329	317.1148	611.2403, 340.1913	C19H18O3	tanshinone IIA[a]	277.1225, 262.0989, 249.1274, 234.1039, 221.1325, 206.1089	19520.4 ± 964.5	1820.6 + 706.7	95.4
34	10.91	301.2165	-	-	C20H28O2	16-hydroxy-6,7-didehydroferruginol*	-	7.9 ± 1.5	n.d.	97.9
35	11.03	283.1693	305.1513	578.2132, 328.2278	C19H22O2	miltirone[a]	265.1591, 223.1119, 195.1168	6271.2 ± 988.8	29.1 + 10.7	98.4
36	11.42	309.1492	331.1309	-	C20H20O3	3-hydroxysalvilenone*	291.1385, 267.1022	7.9 ± 1.5	n.d.	96.9
37	11.66	331.2271	353.2093	-	C21H30O3	cryptojaponol*	-	7.8 ± 1.5	n.d.	99.2
38	11.69	297.1854	319.1672	342.2430	C20H24O2	saprorthoquinone*	279.1750, 269.1905, 255.1399	101.2 ± 12.4	0.7 + 0.3	97.8
39	12.53	299.2011	-	-	C20H26O2	12-deoxy-6,7-dehydroroyleanone*	-	10.4 ± 1.2	n.d.	96.8

[a]Identified via retention time and MS/MS spectrum compared to reference standard.

[b]the proposition of metabolite content in periderm to the whole WT root containing periderm, cortex and xylem.

* compounds didn't isolate in *S. miltiorrhiza* before.

n.d. not detected.

Table S4. Different changed metabolites obtained from metabolomics analyze of WT and CPS1-RNAi lines by GC-QqQ-MS

No.	metabolites	Molecular formula	similatity	t_R (min.)	quantified ion	relative content (ug/g) DW		Percent (%) [b]
						WT ± SE	RNAi ± SE	
1	β-springene[a]	C20H32	94.9	12.35	93.00	n.d.	5.6 ± 0.02	n.d.
2	α-springene[a]	C20H32	94.0	12.84	69.10	n.d.	1.3 ± 0.03	n.d.
3	geranylgeraniol[a]	C20H34O	87.6	15.41	69.10	n.d.	0.5 ± 0.01	n.d.
5	unknown			13.72	135.00	n.d.	0.8 ± 0.02	100
4	Neophytadiene	C20H38	89.2	11.62	57.10	0.03	0.01	100
6	abietatriene[a]	C20H30	92.0	14.00	173.10	0.6 ± 0.01	n.d.	99.9
7	miltiradiene[a]	C20H32	94.7	14.06	80.20	0.2 ± 0.01	n.d.	100
8	ferruginol[a]	C20H30O	94.7	17.25	286.20	29.5 ± 0.6	0.15	99.9
9	sugiol[a]	C20H28O2	87.2	21.20	285.00	1.1 ± 0.02	n.d.	100
10	cryptotanshinone[a]	C19H20O3	90.0	21.59	296.10	6.8 ± 0.14	n.d.	100
11	Retinol	C20H30O	82.8	15.74	174.10	0.08	n.d.	100
12	1,4-dimethyl-8-isopropylidenetricyclo[5.3.0.094,100]decane	C15H24	88.2	14.32	204.10	1.6 ± 0.04	0.02	100
13	unknown			15.87	123.00	0.6 ± 0.02	n.d.	100
14	unknown			16.73	274.00	0.4 ± 0.01	n.d.	100
15	unknown			18.40	282.00	2.2 ± 0.04	0.04	100
16	unknown			18.61	281.00	6.4 ± 0.08	0.32	100
17	unknown			20.64	249.00	2.0 ± 0.06	n.d.	100
18	unknown			21.04	265.10	0.4 ± 0.01	n.d.	100
19	unknown			21.32	279.10	0.7 ± 0.02	n.d.	100
20	unknown			21.84	282.10	2.0 ± 0.05	n.d.	100
21	unkwown			21.90	296.00	0.9 ± 0.02	n.d.	100
22	unknown			22.34	265.10	3.2 ± 0.03	n.d.	100
23	unknown			17.16	284.10	3.7 ± 0.06	n.d.	100

n.d. not detected.

[a] Identified via retention time and EI-MS spectrum compared to reference standard.

[b] the proposition of metabolite content in periderm to the whole WT root containing periderm, cortex and xylem.

n.d. not detected.

Table S5. Summary of statistics for detection of positive selection for CPS group

Lineages	Branches	lnL (MA)	lnL (MA1)	2Δ/ (LRT)[a] Test1 (MA/M1a)	2Δ/ (LRT)[a] Test2 (MA/MA1)	ω_2[b]
CPS	*a*	-38972.35	-38974.20	3.95	3.69	350.86
	a_s	-38974.33	-38974.33	0.00	0. 00	1.00
	b	-38972.61	-38972.61	3.43	0.00	1.00
	c	-38961.74	-38966.66	25.18**	9.84**	5.64
	d	-38966.61	-38966.61	15.44**	0.00	1.00
	e	-38964.06	-38966.16	20.54**	4.21*	16.38
	e_s	-38962.98	-38965.51	22.69**	5.07*	7.09
	f	-38973.76	-38973.76	1.14	0.00	1.00
	g	-38971.57	-38971.57	5.52	0.00	1.00
	h	-38974.33	-38974.33	0.00	0.00	1.00
	i	-38974.33	-38974.33	0.00	0.00	1.00
	j	-38974.33	-38974.33	0.00	0.00	1.00
	k	-38974.33	-38974.33	0.00	0.00	1.00
	l	-38963.72	-38969.17	21.22**	10.90**	16.17
	l_s	-38974.33	-38974.33	0.00	0.00	1.00
	m	-38971.56	-38974.33	5.53	5.53*	∞
	n	-38971.38	-38973.99	5.89	5.23*	∞
	o	-38954.79	-38974.33	39.07**	39.07**	∞
	p	-38965.94	-38972.10	16.78**	12.33**	145.19
	p_s	-38946.54	-38962.91	55.58**	32.74**	15.46
	CfTPS1	-38971.89	-38974.33	4.88	4.88*	26.71
	SmCPS1	-38974.33	-38974.33	0.00	0.00	1.00
	MvCPS3	-38971.28	-38971.51	6.10*	0.47	3.04
	SmCPS2	-38972.57	-38972.57	3.53	0.00	1.00
	SsLPS	-38962.74	-38962.85	23.19**	0.22	1.78
	CfTPS2	-38967.34	-38970.40	13.97**	6.11*	10.37
	NtCPS	-38960.37	-38965.68	27.92**	10.63**	5.75
	LsCPS	-38974.33	-38974.33	0.00	0.00	1.00
	SrCPS	-38974.15	-38974.15	0.00	0.00	1.00
	GrTPS1	-38958.80	-38959.16	31.05**	0.72	1.15
	SmCPS4	-38956.95	-38958.97	34.75**	4.03*	3.43
	IeCPS2	-38963.46	-38968.50	21.74**	10.07**	∞
	SdCPS	-38974.33	-38974.33	0.00	0.00	1.00
	SmCPS5	-38974.33	-38974.33	0.00	0.00	1.00
	IeCPS1	-38971.02	-38973.36	6.61*	4.67*	8.06

SlTPS40	-38969.79	-38974.33	9.07*	9.07**	174.92
CmCPS1	-38970.85	-38972.23	6.95*	2.76	6.92
CmCPS2	-38970.85	-38972.23	6.95*	2.76	6.92
PsaCPS	-38964.99	-38972.42	18.67**	14.86**	34.26
AtCPS	-38972.75	-38974.33	3.15	3.15	26.09
ZmAN2	-38968.66	-38972.04	11.33**	6.75**	147.98
OsCPS2	-38962.97	-38963.00	22.72**	0.05	1.37
OsCPS4	-38957.11	-38957.11	34.43**	0.00	1.44

[a]2ΔL is twice the log-likelihood difference between MA and M1a, MA and MA1. Where under M1a (nearly neutral model), lnL was estimated to be -38974.33 for the CPS tree and -33070.77 for the KSL tree.

[b]Nonsynonymous: synonymous substitution (dN／dS) ratio of site classes 2a and 2b.

[∞]The dN/dS value was estimated to be 999.00.

[*] In Test 1, the difference at the significant level of $P < 0.01$ ([**]) and $P < 0.05$ ([*]) was compared to the chi-squared distribution with d.f. = 2 and in test 2 with d. f. = 1.

158 **Table S6.** Information about primers used in this study

Name	Sequence(5'to 3')	Description
RACE primer		
CPS2-RACE5	CACGAGCCATCCGCCATTTGGTGCTGCGC	3` primer for CPS2 5`-RACE
CPS2-RACE52	CACGTCCTTGATTAGGGCCACTATC	3` primer for CPS2 RACE
CPS3-RACE5	TTCGTCTCCCCATGAGCCATCAGGAAGC	3` primer for CPS3 5`-RACE
CPS3-RACE52	GGCCACACGTCATGTGCGCTTCAGAAG	3` primer for CPS3 5`-RACE
CPS4-RACE5N	CCAAGCTGAAGAGCAGTGATGTGGGCTC	3` primer for CPS4 5`-RACE
CPS4-RACE52	GAGCGTAGGCGGTGGACGACGGCGACGA	3` primer for CPS4 5`-RACE
CPS5-RACE53	TGTACCTTCCTCCGGTGGCGTGTTGGAA	3` primer for CPS5 5`-RACE
CPS5-RACE54	TCTTGTTGGTGTAACGCCTGGCAATTG	3` primer for CPS5 5`-RACE
CPS5-RACE56	TTGTACATTCCGGTGACCGCCTGCGTCGA	3` primer for CPS5 5`-RACE
CPS5-RACE57	GCCTGTCTACTGCCCATAGATGCTCGA	3` primer for CPS5 5`-RACE
KSL2-RACE52	CCCAGTACGTTTCCGTCTCCGACGAGTA	3` primer for KSL2 5`-RACE
KSL2-RACE53N	TTCCTGCCATGCTTGATTATGCCAGAGG	3` primer for KSL2 5`-RACE
Full length cloning primers for E. coli expression		
CPS1-EcoRV	GCGATATCATGGCCTCCTTATCCTCTACAATC	5`primer for full length CPS1
CPS1-NotI	GGAAAGCGGCCGCTCACGCGACTGGCTCGAAAAGC	3`primer for full length CPS1
CPS2+BamHI	CGCGGATCCATGACCTCTTTGTTCAGTCTA	5`primer for full length CPS2
CPS2+EagI	GCCGGCCGTCATACGACCGGTTCAAAAAGTA	3`primer for full length CPS2
CPS3+EcoRV	GCGATATCATGATCTCTCTTTGCTTTTTAGG	5`primer for full length CPS1
CPS3+NotI	TGCGGCCGCTTATACAATCGGCTCAAATAGTA	3`primer for full length CPS1
CPS4+EcoRV	GCGATATCATGTCATTTGCGTCCAACGC	5`primer for full length CPS4
CPS4+NotI	TGCGGCCGCCTAGACTATTTTTTCAAACAATAC	3`primer for full length CPS4
CPS5+EcoRV	GCGATATCATGCCTCTCGCTTCCAATCCCG	5`primer for full length CPS5
CPS5+XhoI	CCGCTCGAGCGGTTACGTACATGTACATGGTCATTG	3`primer for full length CPS5
KSL1+EcoRV	GCGATATCATGTCGCTCGCCTTCAACCCGGC	5`primer for full length KSL1
KSL1+NotI	AGCGGCCGCTCATTTCCCTCTCACATTATTAGC	3`primer for full length KSL1
KSL2+EcoRV	GCGATATCATGGCGCTTCCTCTCTCCACTTG	5`primer for full length KSL2
KSL2+NotI	GGAAAGCGGCCGCTCAACCATGAAGCTTTGATAATCC	3`primer for full length KSL2
At-CPS+EcoRI	CGGAATTCATGTCTCTTCAGTATCATGTTCTAA	5`primer for full length AtCPS
At-CPS+NotI	TTGCGGCCGCCTAGACTTTTTGAAACAAGACTTTG	3`primer for full length AtCPS
At-KSL+XhoI	CCCTCGAGTCAAGTTAAAGATTCTTCCTGTA	5`primer for full length AtKS
At-KSl+NcoI	CCCATGGCTATCAACCTTCGCTCCTC	3`primer for full length AtKS
qRT-PCR primers for gene expression		
CPS1-F	CCACATCGCCTTCAGGGAAGAAAT	5`primer for qRT-PCR *CPS1*
CPS1-R	TTTATGCTCGATTTCGCTGCGATCT	3`primer for qRT-PCR *CPS1*
CPS2-F	GGTCTCATCGCCTTCAACGAAGAT	5`primer for qRT-PCR *CPS2*
CPS2-R	TCCTTATCCTTTATGCTCCCATCCA	3`primer for qRT-PCR *CPS2*
CPS3-F	GGAGATGCCAATTCGAACATCAGA	5`primer for qRT-PCR *CPS3*
CPS3-R	TCAAATATAGTTGCGGCGGCCAAA	3`primer for qRT-PCR *CPS3*
CPS4-F	CGGCTGCCTTGGGCTACAACAATA	5`primer for qRT-PCR *CPS4*
CPS4-R	TCCCTGGTGACCTCCTCCTTCCCA	3`primer for qRT-PCR *CPS4*
CPS5-F	TAGAAGATGCAGCTACTTTCTCTGCT	5`primer for qRT-PCR *CPS5*

CPS5-R	CATCATCTTCACCGCCGTACTGTT	3`primer for qRT-PCR *CPS5*
KSL1-F	CTTCCCAAGACAATGCAAAGAT	5`primer for qRT-PCR *KSL1*
KSL1-R	ATTTCCCTCTCACATTATTAGC	3`primer for qRT-PCR *KSL1*
KSL2-F	TTAGTTTTGGAGGGCAAGAAGAGTGT	5`primer for qRT-PCR *KSL2*
KSL2-R	CTCCTGTTTGGTCGTTGAGAAGAATA	3`primer for qRT-PCR *KSL2*
actin-488	AGGAACCACCGATCCAGACA	5`primer for qRT-PCR *actin*
actin-222	GGTGCCCTGAGGTCCTGTT	3`primer for qRT-PCR *actin*
Pimers for GATEWAY RNA interference construction		
CPS1-RNAi-F	AGGAGATGGGAGTAGAGGGC	5`primer for RNAi *CPS1*
CPS1-RNAi-R	TACATTTCTTTTCCCGTGGC	3`primer for RNAi *CPS1*
CPS5-RNAi-F	GCTGCCACATTCATCAACCGTTGA	5`primer for RNAi *CPS5*
CPS5-RNAi-R	GATAGTTCCTGGATTACAATAAGCTGC	3`primer for RNAi *CPS5*
Primers for genomic sequence cloning		
CPS1g-3220F	TTGGAAGGCTCTGGGCGATCGA	5`primer for *CPS1*
CPS1g-4267R	TCGTACCATCTGCTCAAGCGACA	3`primer for *CPS1*
CPS1g-4116F	TTCGAGTGGCTCTACATGCAAGA	5`primer for *CPS1*
CPS1g-3620R	TTCCTCAGCACATCTGTGTGGA	3`primer for *CPS1*
CPS1g-2560F	TTACTATCGTTCAGCCGGACTA	5`primer for *CPS1*
CPS1g-2663R	TCTCCAGTGGAATCCTGTCGTGGA	3`primer for *CPS1*
CPS1g-956F	TATAGGAGTGACGTACATCAAGGA	5`primer for *CPS1*
CPS1g-1060R	TGCACCGCCGGAGAATCGTAAGGAA	3`primer for *CPS1*
CPS1g-2560R	TAGTCCGGCTGAACGATAGTAA	3`primer for *CPS1*
CPS3-1F	ATGATCTCTCTTTGCTTTTTAGG	5`primer for *CPS3*
CPS3-615R	GGCCACACGTCATGTGCGCTTCAGAAG	3`primer for *CPS3*
CPS3-431F	TCCTGATGGCTCATGGGGAGA	5`primer for *CPS3*
CPS3-855R	TAGTGTTGTTGTCACCTGGTGCA	3`primer for *CPS3*
CPS3-650F	TACCGCAAGGGATCATAAACTCAA	5`primer for *CPS3*
CPS3-3205R	TGCCATAGATGTGTCATCAACTTC	3`primer for *CPS3*
CPS3-3183F	GAAGTTGATGACACATCTATGGCA	5`primer for *CPS3*
CPS3-3508R	TTGGCGATCCACACTGTGGTGGA	3`primer for *CPS3*
CPS3-3428F	TGGTATGCCACGTTACCACGAGTA	5`primer for *CPS3*
CPS3-4764R	GCCATGCAGTTAGCTGAAGTGGCA	3`primer for *CPS3*
CPS3-4738F	TATGCCACTTCAGCTAACTGCA	5`primer for *CPS3*
CPS3-5633R	TTATACAATCGGCTCAAATAGTA	3`primer for *CPS3*
CPS4-1F	ATGTCATTTGCGTCCAACGCC	5`primer for CPS4
CPS4-782R	GACCCATCGTCCAACTGGTTATCCGA	3`primer for CPS4
CPS4-645F	GAGGTTGAATCGGAGAAGATGAAGGA	5`primer for CPS4
CPS4-2653R	TTGACAGGCTTGAGCAGATAGTCGA	3`primer for CPS4
CPS4-2630F	TCGACTATCTGCTCAAGCCTGTCAA	5`primer for CPS4
CPS4-3452R	TATGAGTACCTCGGCAGGAAGGTC	3`primer for CPS4
CPS4-3249F	TTCTGGTGTATGGCAGGGCAGA	5`primer for CPS4
CPS4O-R	CTAGACTATTTTTTCAAACAATAC	3`primer for CPS4