Supporting Information

1, 10/1, 11 Cyclization catalyzed by diverged plant sesquiterpene synthases is dependent on a single residue

Jin-Quan Huang^{‡a,b}, Dong-Mei Li^{‡a,c}, Jian-Xu Li^b, Jia-Ling Lin^{b,d}, Xiu Tian^b, Ling-Jian Wang^b,

Xiao-Ya Chen^b, and Xin Fang^{c*}

^a Yunnan University, Kunming, P.R. China

^b National Key Laboratory of Plant Molecular Genetics, Institute of Plant Physiology and Ecology/CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032, PR China.

^c State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences Kunming, Yunnan 650201, PR China ^d School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China

* To whom correspondence should be addressed:

E-mail: <u>xinfang@mail.kib.ac.cn</u>

¹ These authors contributed equally to this work.

Contents of Supporting Information

Experimental section	3
Fig. S1 Alignment of 1,10-cyclases.	5
Fig. S2 Alignment of 1,11-cyclases.	7
Fig. S3 Alignment of 1,6-cyclases.	8
Fig. S4 Mass spectra of products generated by enzymes used in the manuscript.	9
Fig. S5 Effect of substitution of Cys at preNSE/DTE position of SmSTPS3 with Ser.	14
Fig. S6 Effect of substitution of Ser at preNSE/DTE position of AaGAS with Cys.	14
Fig. S7 Effect of substitution of Cys at preNSE/DTE position of AaCES with Gly.	15
Table S1 Plant sesquiterpene synthases used for the phylogenetic analysis.	16
Table S2 Oligo-nucleotide primers used in this study.	19
Table S3 Product profiles of SmSTPS3, AaGAS and their mutants (%).	20
Table S4 Product profiles of AtTPS21 and its mutants (%).	20
Table S5 Product profiles of PTSiso and its mutants (%).	20
Table S6 Retention indices of new products generated by mutant enzymes.	21
Table S7 Kinetic parameters of SmSTPS3, AaGAS, AtTPS21, PTSiso and their mutants with FDP.	21

Experimental section Phylogenetic analysis

The protein sequences of 163 sesquiterpene synthases (Table S1) were obtained by searching the NCBI GenBank database. These sequences were aligned by the Clustal W program implemented in MEGA with default parameters. Sequences including the ^{pre}NSE/DTE residue and NSE/DTE motif were submitted to MEME web site (<u>http://meme-suite.org/</u>) to analyze amino acids frequency at the ^{pre}NSE/DTE site.

Plant materials

Salvia miltiorrhiza (Lamiaceae) and *Artemisia annua* cv. Qiute (Asteraceae) plants were grown in a greenhouse at 25°C under a light intensity of 150 µmol of photons·m⁻²·s⁻¹ and a 14-h light/10-h dark cycle. *Arabidopsis thaliana* (ecotype Columbia-0) (Brassicaceae) plants were grown at 22°C under a long-day photoperiod (16 h light/8 h dark) and *Pogostemon cablin* (Lamiaceae) plants were grown under similar conditions at 26°C. All plants were grown in a nursery at the Shanghai Institute of Plant Physiology and Ecology.

Gene isolation, expression, and sequence analysis

Total RNA was extracted using the Trizol reagent (Invitrogen, Carlsbad, CA), and 1-µg RNA was reverse transcribed into cDNA using an RNA PCR (AMV) kit (TaKaRa). The coding regions of *SmSTPS3*, *AaGAS*, *AtTPS21*, *PTSiso*, and *AaECS* were amplified from the cDNA library of each plant and cloned into the pET32a vector. *Escherichia. coli* BL21 (DE3) was used for protein expression.

Mutagenesis and protein preparation

Site-directed mutagenesis of sesquiterpene synthases was performed by polymerase chain reaction (PCR) as previously reported ^[1] using an overlap extension strategy with the respective primers. The primers used for gene cloning and mutagenesis are listed in Table S2.

After adding 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) to induce protein expression, *E. coli* BL21 (DE3) cells (50 mL) harboring the recombinant plasmids were incubated overnight at 20°C with shaking (200 rpm). The following day, the cells were collected, resuspended in 7 mL of cell lysis buffer (2.5 mM Tris-HCl, 30 mM NaCl, and 0.2 mM imidazole, pH 7.0), and disrupted using a constant cell disruption system at a pressure of 16.0 Kpsi. The supernatant was centrifuged,

passed through a Ni-NTA spin column (Qiagen) pre-equilibrated with lysis buffer, and washed with 10 mL wash buffer (2.5 mM Tris-HCl, 30 mM NaCl, and 0.4 mM imidazole, pH 7.0). Next, 2 mL of elution buffer (2.5 mM Tris-HCl, 30 mM NaCl, and 40 mM imidazole, pH 7.0) was added and the purified proteins were collected and determined for concentrations using the Bradford method with bovine serum albumin serving as a standard.

Enzyme assays

TPS activities were assayed by adding 10 µg of the purified proteins to 450 µL of reaction buffer [25 mM HEPES (pH 7.0), 5 mM MgCl₂, 5 mM DTT, and 40 µM FDP] in 2-mL screw-top glass vials (Agilent). The reaction mixture was overlaid with 500 µL of hexane and incubated at 37°C for 4 h. The organic phase was passed through a short column of Na₂SO₄, and then subjected to gas chromatography-mass spectrometry analysis (GC–MS: Agilent 6890 Series GC System coupled to an Agilent 5973 Network Mass Selective Detector) at an initial temperature of 60°C (5-min hold), which was subsequently increased to 240°C at 10°C/min. New products were identified by comparing their mass spectra (Fig. S4) and retention indices (Table S6) with those of reported known terpenes^[2-9, 11, 12] and the National Institute of Standards and Technology (NIST) libraries.

Enzyme kinetics

The kinetic parameters of SmSTPS3, AaGAS, AtTPS21, PTSiso, and the mutant enzymes were determined via a malachite green method by measuring the release of pyrophosphate ^[10]. Briefly, reactions were carried out in 50- μ L reaction systems containing 25 mM MES, 25 mM CAPS, 50 mM Tris, 2.5 mU of the coupling enzyme inorganic pyrophosphatase (*Saccharomyces cerevisiae*, Sigma), and 5 mM MgCl₂ at pH 7.5. The reaction was initiated by adding a fixed concentration (10–15 nM) of the purified enzyme. Each assay contained different concentrations of FDP in the range from 1.0 to 100 μ M. The reaction was allowed to proceed for 15 min at 30°C and then terminated by adding 12 μ L of a malachite green development solution, followed by a further 15-min incubation. A623 (UV absorbance at 623 nm) was measured using a NanoDrop 2000 microvolume spectrophotometer. One unit of enzyme activity was defined as the amount of enzyme required to produce 1 μ mol of product per minute under the aforementioned conditions. Values were obtained using GraphPad Prism 5 based on the fitting to Michaelis–Menten curves.

Fig. S1 Alignment of 1,10-cyclases. ^{pre}NSE/DTE residues are outlined in red.

	610	620	630	640	650
Dc (-)-germacrene D synthase	TKEVFDWVSTF	PKFTHQASLI	CRIRDDVVSQEF	EQKRNHVASTVQ	CYMKEYSAT
Vv (-)-germacrene D synthase	TKESFDWIFSK	PKIVRASAIV	CRLMDDMVFHKF	EQKRGHVASAVE	CYMKQHGAS
Vv germacrene D synthase	TKEAFDWVTSD	PKIMSSSNFI	ARLMDDISSHKF	EQKREHVASAIE	CYMKQYGVS
Le germacrene C synthase	SHETFEWLMNE	SVIVRASALI	ARAMNDIVGHED	EQERGHVASLIE	CYMKDYGAS
Ha germacrene A synthase 1	SEDALVWYESH	PQILQASELI	SRLQDDVMTYQF	ERERGQSATGVD	SYIKTYGVS
Ha (-)-germacrene D synthase	TDKSFEWVLTN	PPLIKASSQI	CRLMDDIVGHKE	EQKREHVASVVE	SYMKEHDVT
Sc germacrene A synthase	TQESFKWVLNN	PPLIKASSDI	SRIMNDIVGHKE	EQQRKHIASSVE	MYMKEYNLA
Ci germacrene A synthase	DEEALAWYETH	PKILKASELI	SRLQDDVMTFQF	ERKRGQSATGVD	AYIKEYNVS
Ls germacrene A synthase 1	SENALAWYESH	PKTLQASELI	SRLQDDVMTYQF	ERERGQSATGVD	AYIKTYGVS
Ls germacrene A synthase 2	SEDALAWYESH	PKTLQASELI	SRLQDDVMTYQF	ERERGQSATGVD	SYIKTYGVS
Ls germacrene A synthase 3	DEEALAWYETH	PKILKASELI	SRLQDDVMTFQF	ERKRGQSATGVD	AYIKEYNVS
Bs germacrene A synthase 1	NEEALAWYETH	PKILKASELI	SRLQDDVMTFQF	ERKRGQSATGVD	AYIKEYKVS
Bs germacrene A synthase 2	NEEALDWYESH	PKTLKASELI	SRLQDDVMTFQF	ERERGQSATGVD	AYIKTFGVS
Am germacrene A synthase	SEDALAWYESH	PKPLQAAELI	SRLQDDVMTYQF	ERERGQSATGVD	AYIKTYGVS
Ccvs germacrene A synthase	SKDALAWYESH	PRTLQAAELI	SRLQDDVMTYQF	ERERGQSATGVD	AYIKTFGVS
Cz germacrene B synthase	TKEVFDWYSSF	PKAVEAGCAI	GRILNDIESHER	EQAGDHFASTVE	SYMKEYNIN
Cs germacrene D synthase	TKEAFDWVFSD	PLIVKAASVI	CRLTDDMVGHEF	EHKREHIDSAVE	IYMKQYGVT
Ac germacrene A synthase	NEEAFAWHKSR	PLVLRSSEII	SRLQDDIVTSPG	EKEKGQVTSGID	YYMKAYGVS
Zp germacrene D synthase	TEAAFKWLFSN	PKMVKASEIV	CRFMDDIVSHKF	EQSRGHVASSIE	CYMKQNGAT
Tvvl germacrene D synthase	TREAFDWMSKD	PLIVKASSVI	CRLSDDIVGHEF	ELERPHIPTAVE	CFMKQYGVP
Xs germacrene A synthase	SADALAWYESH	PKILQASELI	SRLQDDVMTYQF	ERERGQSATGVD	SYIKTYGVS
Vo germacrene B/C/D synthase	PHQAFDWVSNN	PLMVKASTII	ARLDNDKVGHEH	EQDRGHVASGVE	CYMKQHGAT
Vo germacrene C synthase	PHQAFDWVSNN	PLMVKASTII	ARLDNDKVGHEH	EQDRGHVASGVE	CYMKQYGAT
Ga (-)-germacrene D synthase	TEETFNWASNN	PKIVRASSMI	SRIMDDIVSHKF	EQERGHCASAVE	CYMKQHGVS
Ha germacrene A synthase 2	SEDALAWYESH	PQILQASELI	SRLQDDVMTYQF	ERERGQSATGVD	AYIKTYGVS
Ha germacrene A synthase 1	SEDALVWYESH	PQILQASELI	SRLQDDVMTYQF	ERERGQSATGVD	SYIKTYGVS
Ha germacrene A synthase 3	SENALAWYESH	PKTLQASELI	SRLQDDVMTYQS	ERERGQSATGVD	AYIKTYGVS
Zo germacrene D synthase	TKEVFDWHSSF	PKAVEACCAI	GRILNDITSYER	EQGRGDNASTVE	SYMKDHGTN
Sl germacrene synthase	SHETFEWLMNE	SVIVRASALI	ARAMNDIVGHED	EQERGHVASLIE	CYMKDYGAS
Sl germacrene C synthase	SHETFEWLMNE	SVIVRASALI	ARAMNDIVGHED	EQERGHVASLIE	CYMKDYGAS
Sl epidermal germacrene C synthase	SHETFEWLMNE	SVIVRASALI	ARAMNDIVGHED	EQERGHVASLIE	CYMKDYGAS
Pc germacrene D synthase	TPQAFQWVLSE	PQIITASLAL	ARLSNDLVGIEF	EKERKYIATAVE	LYEEEHKVS
Pc germacrene A synthase	TKEDFEWFSKK	PRIAVATQII	TRVIDDIATYEV	EKEKGQSATGID	CYMKEHGVS
Pc germacrene D synthase	TKEAFDWVLSE	PDVMKAAITL	ARLTNDIVGIEL	EKERKHIATAVE	VYEDEHKLS
Sc (-)-germacrene D synthase	TQESFKWALNN	PPLIKASADV	SRIMDDIVGHKE	EQQRKHLPSRVI	EMYMKKYHLA
So (-)-germacrene D synthase	SKEEFDWVLSE	PPLLMASSII	TRLMDDLAGYGG	EEKLSAVE	HYYMLEYGVS
Sc (+)-germacrene D synthase	TQDSFKWALNN	PRLITASAYI	GRILNDIVGHKE	EQQRKHIPSIVE	EMYMKEHNLL
Ob germacrene D synthase	TKQDFDWITNE	PPILRAASVI	CRLMDDVVGHGI	EQKISSVI	DCYMKENGCS
Gs (-)-germacrene D synthase	TEEVLIWATSD	PIIIGAASII	CRLMDDIVGNEF	EQERRHVASSI	ESYMKQHNTS
Ls germacrene A synthase	SENALAWYESH	PKTLQASELI	SRLQDDVMTYQF	ERERGQSATGVI	DAYIKTYGVS
Ls germacrene A synthase 1	SEDALAWYESH	PKTLQASELI	SRLQDDVMTYQF	ERERGQSATGVI	DSYIKTYGVS
Cmfk germacrene D synthase 1	TEKAFDWLKTN	PKFVMDSSL	CKLVDDIHSNEF	EQKRPHIASAVI	DCYMKEHGVS
Aa germacrene D synthase 2	TEESFKWIDTF	PPLVKASSAV	CRIMDDIVGHKE	EQQREHVASSVI	SYMKQHDVT
Aa germacrene D synthase 3	TEESFKWIDTF	PPLVKASSAV	GRIMDDIVGHKE	EQQREHVASSVI	SYMKQHDVT

Fig. S1 continued Alignment of 1,10-cyclases. ^{pre}NSE/DTE residues are outlined in red.

SEDSFKWALTNPLLIKAACAIORTMDDIVGHKEEKERGHVASFVECYMKQHNVT TEESFKWIDTFPPVVKASSAVORIMDDIVGHKEEQQRGHVASSVESHMKQHDLT Aa (-)-germacrene D synthase Aa germacrene D synthase 4 SEDALAWYESHPKTLQASELI SRLQDDVMTYQFERERGQSATGVDAYIKTYGVS SEDALAWYESHPKTLQASELI SRLQDDVMTYQFERERGQSATGVDAYIKTYGVS TKETFEWVSNEPLIVRASSVICRLMDDMVGYGIESK----FTAVECYMNENDAL Aa germacrene A synthase 1 Aa germacrene A synthase 2 Hi (-)-germacrene D synthase VREDFEWLSKKPKMLVAGLLICRVIDDIATYEVEKDRGQIATGIESYMRDNGAT Lv germacrene A synthase SEDALAWYESHPKPLQASELISRLQDDVMTYQFERERGQSATGVDAYIKTYGVS Tp germacrene A synthase SEHALAWYESHPKTLQASELISRLQDDVMTYQFERERGQSATGVDSYIKTYGVS Cs germacrene A synthase DEEALAWFETHPKILKASELISRLQDDVMTFQFERKRGQSATGVDAYIKEYKVS To germacrene A synthase GEDALAWYESHPKTLQASELISRLQDDVMTYQFERERGQSATGVDSYIKTYGVS To germacrene A synthase KDEVFSWACSFPKIIEIAAINORLMDDIAGHENETLRSYTRRFFEMSSTMVNLT Do (-)-germacrene D synthase SKEAFDWLFSHPPVIEASESVORLMDDMRSHKFEQERGHVASGIECYMKQYGVT Sa germacrene D-4-ol synthase GEDALAWYESHPKPLQASELISRLQDDVMTYQFERERGQSATSVDAYIKTYGVS Cn germacrene A synthase TKEDFEWFSKKPRIAVATQIITRVIDDIATYEVEKEKGQSATGIDCYMKEHGVS Pc Germacrene A synthase TEEAFQWVLGSPDPVKAASVVORLQDDIVGYKFDKQRDHVASAIDCYMKDNINS Hi Bicyclogermacrene synthase SKADFDWIVNEPLIVRASCVICRLMDDLVGDEYEEK----PSSVLCYMKQYVVS TKEAFDWVTSDPKIMSSSNFITRLMDDIKSHKFEQKRGHVTSAVECYMKQYGVS Ov Bicyclogermacrene synthase Vv (+)-valencene synthase TKEVFEWISNNPKVVKAASVICRLMDDMQGHEFEQKRGHVASAIECYTKQHGVS Cs valencene synthase SNDVFKWLFNDPKIVRAAALIORLMDDIAGHKFEQERKHVASGVECYMEQYGVT Mn valencene synthase GYDVFNWAISFPKIIENAAMVSRLMDDISGHECEKERSNVATAVDCYIKEHGVT Tu Aristolochene synthase TEQDFEWLSKNPKILEASVIICRVIDDTATYEVEKSRGQIATGIECCMRDYGIS Nt 5-epi-aristolochene Synthase AEQDFEWLSKNPKILEASVIICRVIDDTATYEVEKSRGQIATGIECCMRDYGVS KEQQFEWLSKNPKILEGCVTICRVIDDIATYEVEKNRGQLSTGIECYMRDYSVS Na 5-epi-aristolochene synthase Cava 5-epi-aristolochene synthase TEODFEWLSKNPKILKANATIORVIDDIATYEVEKSRGEIATGIECYRREYGVS Ca 5-epi-aristolochene synthase VEEDFEWLANKPKMLVAGLLICRVIDDIATYEVEKERGQSATGIESYMRDNNAT Sm 5-epieremophilene synthase 1 VEEDFQWLAKKPKMLVAGLLICRVIDDIATYEVEKERGQSATGIESYMRDNNAT Sm 5-epieremophilene synthase 2 VEEDFQWLAKKPKMLVAGLLICRVIDDIATYEVEKERGQSATGIESYMRDNNAT Sm5-epieremophilene synthase 3 TKEHFEWLATNPKILEANATI Hm Vetispiradiene Synthase LKETLDWASTIPEMVKASSLIARYIDDLQTYKAEEERGETVSAVRCYMREYGVS LKETLDWASTIPEMVKASSLIARYIDDLQTYKAEEERGETVSAVRCYMREYDVS Am delta-guaiene synthase Ac Delta-guaiene synthase TVGTCKWATKKPKAVFAAERKGRLINDIVGYEEEHSRPHVATSIDCYMKEYGVS Tvvl guaiene synthase TTEAFEWILTKPKMVAASGAIGRLVDDVMSNDEEQERGHVATGLDCYMKQHGVS Cb hedycaryol synthase TKEAFDWVFSRPPFIEATLIIARLVNDITGHEFEKKREHVRTAVECYMEEHKVG Pc patchoulol synthase TKEAFDWVFSRPPFVEATLIIARLINDITGCEFENKREHVRTAVECYMEEHKVG Pc patchoulol synthase isoform TREAFEWADSAPDMVLASGEVGRFLNDMASYKLGKNKKDVANAHECYMKEYGAT Zmsm beta selinene synthase TKKDFDWIRNRPRLLQVAEVITRLMDDIAGHGTEKK----TTAVSCYMKEYECS Ob selinene synthase PDNILEQIHSPSKILDLLELTGRIADDLKDFEDEKERGEMASSLQCYMKENPES Ag Delta-selinene synthase At delta-cadinene synthase KLDAFEWLSSHPKIRVASEIISRFTDDISSYEFEHKREHVATGIDCYMQQFGVS TNEAFDWVTSDPKIMSSSNFIARLMDDIKSHKFEQKRGHVASAVECYMKQYGVS Vv gamma-cadinene synthase KREHFDWVLSDPDFVMASCIICRLADDIVGHEFEQTRDHIPSSVECYTQEHKTS Pd alpha-copaene/delta-cadinene synthase TKEAYEWASEDPKIIRASSIVCRLMDDIVSHEFEQTRKHVASGVECYIKQYGAS TPETFKWAASDPKIIKASTIICRFMDDIAEHKFNHRREDDCSAIECYMKQYGVT TPETFKWAANDPKIIQASTIICRFMDDVTEHKFKHRREDDCSAIECYMEEYGVT Rc (+)-delta-cadinene synthase Ga (+)-delta-cadinene synthase Gh (+)-delta-cadinene synthase TQEAIDWIKNEPNFAVKAGLIGRYWDDIGSHKRESKGGEMLTVMDCYMKQYSVS Ob Gamma-cadinene synthase TNEVLQWLSNGPQIIKASTIICRLMDDIASHKFEQEREHVASAVECYMKQYDCS Cm (+)-gamma-cadinene synthase TKEVLEWAIHMPTMLRTCSIVARLMDDIPSNKLEQERKHVSSSVECYMKEHGTS Mg Beta-cubebene synthase TKDVFEWAIKRPNIVVAASMICRNRDDIVGHKEEQERGDVPSGVECYTKDHGCT Pn copaene synthase

Fig. S2 Alignment of 1,11-cyclases. ^{pre}NSE/DTE residues are outlined in red.

	510	520		530	540	550
At hata, cantonhullana/aloha, humulana cunthasa	MTDUAKT D-AFEN	NCUDETDUAC	PT TE	PETDDICC	VEFEURDE	UNTOTOCYMOO
At beta-caryophyliene/aipha-numulene synthase	MTDVAKLD-AFEW	LNSHPKIKVAS	EIIS	REIDDISS	IEFERKRE	HVATGIDCIMQQ
Aa beta-caryophyliene synthase	MGDIVTDE-AFKW	ALTRPPIIKAS	SCATA	RLMDDIHS	QKEEKERI	HVASSVESIMKQ
Zmsp (E)-beta-caryophyllene synthase	MGDVATSE-ALEW	ASTYPKIVRAV	CIIA	RLANDIMS	YKREASNN	IMVSTVQTCAKE
Pn caryophyllene synthase	AGEITTKK-LLEW	IQSQPKYMKDT	CRLC	RIVDDIKT	YKFEEERG	HVASVVACYMEE
Zp (E)-beta-caryophyllene synthase	MGDVATSE-TLEW	ASTYPKIVRAV	CIIA	RLANDIMS	YKREASNN	TMVSTVQTCAKE
Gh (E)-beta-caryophyllene	TNHVTKEG-LQTF	EEYHPNIIRWS	STVL	RLANDLAT	SSYEIKRG	DIPKSIQCYMHE
Pl (E)-beta-caryophyllene synthase	MEDTTEEVLIW	ATNEPVIIAAS	TAIA	RIMDDIVG	DEFEQQRE	HVVSSIYCYMKQ
Ok beta-caryophyllene synthase	IPHHISQQ-EFDW	VLSEPPLLRAS	LTIT	RLMDDLAG	YGSEEK	LSAVHYYMSE
Li 9-epi-caryophyllene synthase	TTDPINNT-ELES	LEKHPGIIYWP	SMVL	RLADDLGT	SSDEMKRG	DVSKSVQCYMNE
Zl (E)-beta-caryophyllene synthase	MGDVATSE-ALEW	ASTYPKIVRAV	CIIA	RLANDIMS	YKREASNN	TMVSTVQTCAKE
Zmsm (E)-beta-caryophyllene synthase	MGDVATSE-ALEW	ASTYPKIVRAV	CIIA	RLANDIMS	YKREASNN	TMVSTVQTCAKE
SI caryophyllene/alpha-humulene synthase	VEEFISKE-TFEW	MINEPLIVRAS	SLIA	RAMDDIVG	HEVEQORE	HGASLIECYMKD
Vvcg (E)-beta-caryophyllene synthase 1	MGDVVTKE-TFDW	VFSEPKIVRAS	ATVS	RLMDDMVS	HKFEOKRG	HVASAVECYMKQ
Pd beta-carvophyllene synthase	LGNSLVMK-DFDW	VSCEPLMVKAS	AIIA	RLMDDMAG	HGFEKK	ISAVECYTNE
Vvcd (E)-beta-carvophyllene synthase	MGDALTKE-TLDW	IFSEPKIVRAS	AIVC	RLMDDMVP	HKFEOKRG	HVASAVECYMKO
Vvcg (E)-beta-carvophyllene synthase 2	MGE IMTKE-AFDW	VISDPKIITAS	TVIE	RLMDDITT	HKFEOKRG	HVASGIECYMKO
Mcvr (-)-beta-carvophyllene synthase	MGDIVTDE-AFEW	SLTKPPIIKAS	CATA	RLMDDTHS	OKDEKERI	HVASSVESYMKO
Mm beta-carvonhyllene synthase	MGDITTND-SINW	VISEPPLIKAT	MMTG	RLLNDTVS	SKKEOERV	HFOSTVOCYKKO
HI Alpha-humulene synthase	MKSSTVTKEVFEW	LSMDRKTTRAS	STIC	REMODIAE	HKFEOEKN	DEPTAVECYMKO
Ag gamma-humulana synthase	MGEHLPID-TLEO	TFL DSRFHHLT	FLAS	RIVDDARD	FOAFKDHG	DLSC-TECYLKD
Zz Alpha humulana cunthasa	LGEDATKE-TEEW	VASSORTIKSC	CTHC	TMDDTTS	HOPFOFPD	HEACTUREVMER
Dz slobe humulene synthese	LOLUANN-TVEO	TICDENTOFIC	FTTT	TODIES	FEDEVED	FINCTURCYMED
Pg aipna-numulene synthase	LOULLANN-IVEQ	THEREKTHELS	ELII	LIDDIKD	FEDERERG	EIASIVECIMAD
Pepg longitolene synthase	MGQLLPDD-VLEQ	THSPSKIHELV	ELIA	RLVDDSKD.	FETKKAGG	ELASGIECIVKD
Pa longitolene synthase 1	MGQLLPDD-VLEQ	IHSPSKIHELV	ELTA	RLVDDSKD	FETKKVGG	ELASGIECYVKD
Ps longifolene synthase	MGQLVPDN-ILEQ	IHLPSKIHELV	ELTA	RLVDDSKD	FQAKKDGG	EFASGTECYLKE
Ps alpha-longipinene synthase	MGQLLPDD-VLEQ	IHSPSKIHELV	ELTA	RLVDDSKD	FETNKVGG	ELASGIECYVKD
Pa Longifolene synthase 2	MGQLLPDD-VLEQ	IHSPSKIHELV	ELTA	RLVDDSKD	FETKKVGG	ELASGIECYVKD
Vo valerena-1,10-diene synthase	MGVTTVTKPAFDW	VTNNPLILIAS	CTIN	RLADDKVG	HELEQERG	HVASGVECYMKH

Fig. S3 Alignment of 1,6 cyclases. ^{pre}NSE/DTE residues are outlined in red.

Zm 7-epi-sesquithujene synthase Ag Alpha-bisabolene synthase Zm tps5-Del1 Zo(S)-beta-bisabolene synthase At (Z)-gamma-bisabolene synthase Gb alpha bisabolene synthase Pm(E)-gamma-bisabolene synthase Sa Beta-bisabolene synthase Ss Alpha-bisabolol synthase Pa E-alpha-bisabolene synthase Zo (S)-beta-bisabolene synthase Sb Beta-sesquiphellandrene synthase Ss a-bisabolol synthase Zo sesquiterpene synthase Ag E-alpha-bisabolene synthase Sa b-bisabolene synthase Ha (Z)-gamma-bisabolene synthase K7 Zm (S)-beta-macrocarpene synthase 1 Zm(S)-beta-macrocarpene synthase 2 Ee bisabolol synthase Pd(+)-epi-alpha-bisabolol synthase Aa bisabolol synthase Am (+)-alpha-bisabolol synthase Ak (+)-alpha-bisabolol synthase Mcvr a-bisabolol synthase Aabs koidzumiol synthase Aabs koidzumiol synthase Aa Amorpha-4,11-diene synthase Aa epi-cedrol synthase Mcvr sesquiterpene synthase 1 Mcvr sesquiterpene synthase 4 Vvcpn beta-curcumene synthase Pc Gamma-curcumene synthase Vvcg (E)-alpha-bergamotene synthase Pd Trans-alpha-bergamotene synthase Sa santalene bergamotene synthase 1 Sa santalene bergamotene synthase 2 Ss santalene synthase Sau santalene synthase Sb Zingiberene synthase Ob Alpha-zingiberene synthase At Alpha-barbatene synthase

720 730 740 TFEWALSFPQFIRTFGSFVRLSNDVVSTKREQTKDHSPSTVHCY KVDYPGR-RVLTELNSLISRLADDTKTYKAEKARGELASSIECY TFEWALSFPQFIRTFGSFVRLSNDVVSTKREQTKDHSPSTVHCY AFEWFASFPKIIEACATI AYEWVRSRPRLIKSLAAKGRLMDDITDFDSDMSNGFAANAINYY QVDLPGHPDKLIELNCIISRLSDDTKTFQAEKARGELASSIECY KIDYPERSRVLMEQICLISRLADDTQSYKAEKARGELASGIECY SAAYVNSVPPLVRYSGLLNRLYNDLGTSSAEIARGDTLKSIQCY SAAYVNSVPPLVRYSGLLNRLYNDLGTSSAEIARGDTLKSIQCY KVDYPGR-RVLTELNSLISRLADDTKTYKAEKARGELASSIECY AFEWFASFPKIIEACATIIRITNDITSMEREQKRAHVASTVDCY ILDWLLSYPKLLKSMTTFVRLSNDIASTKREQTGGHHASTVQCY SAAYVNSVPPLVRYSGLLNRLYNDLGTSSAEIARGDTLKSIQCY AFEWFASFPKIVEACAIIVRITNDITSKEREQKREHVASTVDCY KVDYPGR-RVLTELNSLISRLADDTKTYKAEKARGELASSIECY SAAYVNSVPPLVRYSGLINRLYNDLGTSSAEIARGDTLKSIQCY Ha (Z)-gamma-bisabolene synthase K11 IFKWLSNYPPIVKASSLILRYMNDLSTRKDEOERNHVASSVKCY IFKWLSNYPPIVKASSLILRYMNDLSTRKDEQERNHVASSVKCY IMKWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCY IMEWVMSDAELVKSFGIFVRLSNDIVSTKREQREKHCVSTVQCY TFKWVATHPPLVKASCLILRLMDDIATHKEEQERGHVASSIECY HILDLRTNPPVIKWVSILVRLADDLGTSTDELKRGDNPKSIQCH AFEWAVSEPPLLRYKGILGRRLNDLAGHKEEQERKHVSSSVESY AVEWAVSEPPLLRYKGILGRRLNDLAGHKEEQERKHVSSSVESY AVEWAVSEPPLLRYKGILGRRLNDLAGHKEEQERKHVSSSVESY AFKWVATHPPIVKAACKILRLMDDIATHKEEQERGHIASSIECY AVEWAVSEPPLFRYSGILGRRLNDLVSHKDEQERKHISSSVESY AVEWAVSEPPLFRYSGILGRRLNDLVSHKDEQERKHISSSVESY SVEWAVSAPPLFRYSGILGRRLNDLMTHKAEQERKHSSSSLESY SFEWVFTNPPLVNACCLLCRTMDDLGSHKGEQDRKHVASTIECY AFKWVATHPPIVKAACKILRLMDDIATHKEEQERGHIASSIECY PFKWALANPPLVRATCAICRFMDDIVGHEDEQQRNHVVSVVECY AFDWVLNGPKIVRACSTIIRLMDDMASHKFEQERGHIASSVECY NIDWLLSQPRLASSTAIVMRCCNDLGSNQRESKGGEVMTSLDCY VFDWLQQNPKIVRASSKVIRLMDDMATHKFEQERGHIASSIECY ELQYLEQYPGIIRWPSTVLRLADDLGTASDEIKRGDVPKSIQCY ALDKVHPLPDLLHYSSLLSRLINDMGTSSDELERGDNLKSIQCY ALDKVHPLPDLLHYSSLLSRLINDIGTSPDEMARGDNLKSIHCY ALDKVHPLPDLLHYSSLLSRLINDMGTSPDEMARGDNLKSIHCY ALDKVHPLPDLLHYSSLLSRLINDIGTSPDEMARGDNLKSIHCY VLDWLLTYPELLKCFTTFVRLSNDITSTKREQTGGHHASTVQCY ELQSLERHAHSL---SMILRLADDLGTSSDEMKRGDVPKAIQCF AFEWLISRPKLVRILGAKTRLMDDIADFEEDMEKGYTANALNYY **Fig. S4** Mass spectra of products generated by enzymes used in the manuscript. The most of sesquiterpenes are yielded by Wt enzymes ^[2-7], and by compared with known terpenes in reported enzyme assay ^[8, 9].





Fig. S4 continued Mass spectra of products generated by enzymes used in the manuscript.



Fig. S4 continued Mass spectra of products generated by enzymes used in the manuscript.



Fig. S4 continued Mass spectra of products generated by enzymes used in the manuscript.



Fig. S4 continued Mass spectra of products generated by enzymes used in the manuscript.



Fig. S5 Effect of substitution of Cys at preNSE/DTE position of SmSTPS3 with Ser.



Fig. S6 Effect of substitution of Ser at ^{pre}NSE/DTE position of AaGAS with Cys.



Fig. S7 Effect of substitution of Cys at ^{pre}NSE/DTE position of AaCES with Gly.

Table S1 Plant sesquiterpene synthases used for the phylogenetic analysis, their corresponding

protein ids, and	organism	names.
------------------	----------	--------

Sesquiterpene Synthases	Protein ID	Organism
Dc (-)-germacrene D synthese	PK1180372 1	Dendrohium catenatum (Orchidaceae)
Vy (-)-germacrene D synthase	ΔV561842.1	Vitis vinifora (Vitaceae)
Vy germacrene D synthase	HM807378 1	Vitis vinifera (Vitaceae)
I e germacrene C synthase	AF035631 1	Ivcopersicon esculentum cy VFNT
Le germaerene e synthase	11 055051.1	(Solanaceae)
Ha germacrene A synthase 1	A A V / 1 / 21 2	Holianthus annuus (Asteraceae)
Ha (_)-germacrene D synthase	OTG28618 1	Helianthus annuus (Asteraceae)
Sc germacrene Λ synthase	CAC36896 1	Solidago canadansis (Asteraceae)
Ci germacrene A synthase	A A M 21658 1	Cichorium introdus (Asteraceae)
Ls germacrene A synthase 1	AE480065 1	Lactuca sativa (Asteraceae)
Ls germacrene A synthase ?	AF489905.1	Lactuca sativa (Asteraceae)
Ls germacrene A synthase 3	AT 409904.1 AOT 80657 1	Lactuca sativa (Asteraceae)
Bs germacrene A synthase 1	AUX07167 1	Rarnadosia spinosa (Asteraceae)
Bs germacrene Λ synthase ?	AIX07168 1	Barnadosia spinosa (Asteraceae)
Am gormagrana A synthasa	AGD80125 1	Achillag millafolium (Astornoono)
Caus cormocrono A sunthase	AGD60155.1	Comara gardungulug vor soolumus
Covs germaciene A synthase	AE193043.1	(Asternagona)
Cz gormagrona P synthaga	DDE27240 1	(Asteraceae)
Cz germaerene D synthase	DDE32340.1 DDE32323 1	Changionanar seigdonhylloides
Cs germacrene D synthase	DDE52552.1	(A raliaaaaa)
A a company A grathage	DDE20101 1	(Araliaceae)
Ac germacrene A synthase	DDD99590 1	Aralla coradia (Alallaceae)
Zp germacrene D synthase	BBD88389.1	There is will be a serie la cinista (Amiceae)
Va compositione D synthase	AMU19318.1	Vanthium atum anium (Astarasasa)
As germacrene $P/C/D$ synthese	AJI00313.1	Valoriana officialia (Valorianaceae)
Vo germaciene B/C/D synthase	AFK4241/.1	Valeriana officinalis (Valerianaceae)
$C_{0}(x)$ as maximum and $C_{0}(x)$ as maximum and $C_{0}(x)$	AGB03010.1	<i>Consuming subanage</i> (Malyageae)
Ga (-)-germacrene D synthase	KHG04103.1	Gossyptum arboreum (Matvaceae)
Ha germacrene A synthase 2	EU32//83.1	Helianthus annuus (Asteraceae)
Ha germacrene A synthase 1	DQ010007.2	Heliantnus annuus (Asteraceae)
Ha germacrene A synthase 3	GU1/6380.1	Helianthus annuus cultivar HA300
7	A A V 40((5 1	(Asteraceae)
Zo germacrene D synthase	AAA40005.1	Salarana (Salarana)
Si germacrene synthase	AEMU5858.1	Solanum lycopersicum (Solanaceae)
SI germacrene C synthase	AAC39432.1	Solanum lycopersicum (Solanaceae)
SI epidermal germacrene C synthase	AAC39431.1	Solanum lycopersicum (Solanaceae)
Pc germacrene D synthase	AY 508/29.1	Pogostemon cablin (Lamiaceae)
Pc germacrene A synthase	AY508/28.1	Pogostemon cablin (Lamiaceae)
Pc germacrene D synthase	AY508/27.1	Pogostemon cablin (Lamiaceae)
Sc (-)-germacrene D synthase	AAR31145.1	Solidago canadensis (Asteraceae)
So (-)-germacrene D synthase	AQY54367.1	Salvia officinalis (Lamiaceae)
Sc (+)-germacrene D synthase	AAR31144.1	Solidago canadensis (Asteraceae)
Ob germacrene D synthase	AAV63/86.1	Ocimum basilicum (Labiatae)
Gs (-)-germacrene D synthase	RZB/5/4/.1	Glycine soja (Fabaceae)
Ls germacrene A synthase 2	AAM11627.1	Lactuca sativa (Asteraceae)
Ls germacrene A synthase l	AAM11626.1	Lactuca sativa (Asteraceae)
Cmtk germacrene D synthase 1	RWR76602.1	Cinnamomum micranthum f. kanehirae
Aa germacrene D synthase 2	PWA81675.1	Artemisia annua (Asteraceae)
Aa germacrene D synthase 3	PWA51042.1	Artemisia annua (Asteraceae)

Aa (-)-germacrene D synthase	PWA680101	Artemisia annua (Asteraceae)
Aa germacrene D synthase 4	PWA58326 1	Artemisia annua (Asteraceae)
Aa germacrene A synthase 1	DO447636 1	Artemisia annua (Asteraceae)
Aa germacrene A synthase 2	PWA48097 1	Artemisia annua (Asteraceae)
Hi (-)-germacrene D synthase	PIN14350 1	Handroanthus impetiginosus (Bignoniaceae)
Ly germacrene A synthase	AGN72806 1	Lavandula viridis (Lamiaceae)
Tn germacrene A synthase	AEH41844 1	Tanacetum parthenium (Asteraceae)
Cs germacrene A synthase	ABR00361 1	Crepidiastrum sonchifolium (Asteraceae)
To germacrene A synthase	ALY05869 1	Taraxacum officinale (Asteraceae)
To germacrene A synthase	ALY05868 1	Taraxacum officinale (Asteraceae)
Do (-)-germacrene D synthase	OEL12694 1	Dichanthelium oligosanthes (Poaceae)
Sa germacrene D-4-ol synthase	B5A435 1	Santalum album (Santalaceae)
Cn germacrene A synthase	AMN10097 1	Chamaemelum nobile (Asteraceae)
Pc germacrene A synthase	O49SP5 1	Pogostemon cablin (Lamiaceae)
Hi bicyclogermacrene synthase	PIN24805 1	Handroanthus impetiginosus (Bignoniaceae)
Ov bicyclogermacrene synthase	F2F2N7 1	Origanum vulgare (Lamiaceae)
$V_{V}(+)$ -valencene synthase	AV5618/3 1	Vitis vinifora (Vitaceae)
Cs valencene synthase	Δ F441124 1	Citrus singusis (Rutaceae)
Mn valencene synthase	XP 010002205 1	Morus notabilis
Tu aristolochene synthase	EMS62001 1	Triticum Urartu (Poaceae)
Nt 5 eni aristolochene Synthase	5EAU A	Nicotiana tabacum (Asteraceae)
Na 5 epi-aristolochene synthase	A A D05762 1	Nicotiana attonuata (Asteraceae)
Cava 5 ani aristologhene synthese	CAA06614.1	Cansieum annuum vor annuum
Cava 5-epi-ansiolochene synthase	CAA00014.1	(Solonococo)
Ca 5 ani aristologhang synthese	ND 001211620 1	(Solanaceae)
Sm 5 anioromonbilana synthasa 1	VV422512 1	Salvia miltionnhiza (Lomiocooo)
Sm 5 epieremophilene synthase 2	K1432312.1 KV/32513.1	Salvia militorrhiza (Lamiaceae)
Sm 5-epieremorphilene synthase 3	K1432313.1 KV/3251/ 1	Salvia militorrhiza (Lamiaceae)
Um votignizaciona Synthasa	5107 D	Huosavamus muticus (Solonococo)
Am dolta guaiano synthaso	JU7_D ALD94926 1	Aquilaria malagagensis (Thumalagagaga)
An delta guaione synthese	DOVMD9 1	Aquilaria arassna (Thymoloopoopo)
True quaiona synthese	A 71 40028 1	Thansia villoga var loginisto (Anigono)
Ch hadvaarvol synthase	AZL40020.1	Camellia brovistyla
De neteboulel synthese	DAU08090.1	Camella Drevisiyia Recogtomon achlin (Lomiacoco)
Pe patchoulol synthese variant	A I 306/30.1	Pogostemon cablin (Lamiaceae)
The patchould synthase variant	AIL24446.1	Zeg w gub
Ch salinana synthese	A 1 123330.1	<i>Deimum hasilisum</i> (Lehistee)
A g delte gelinene gynthage	AAV05/05.1	Abias grandis (Dinagogo)
At dalta and mana synthese	004404.1 A 1544020 1	Ables grundis (Filiaceae)
At denta-cadinene synthase	AJ344238.1	<i>Vitia vinifera (Vitegaga)</i>
vv gamma-cadinene synthase	HM80/380.1	Vins vinijera (Vinaceae)
Pu alpha-copaene/dena-	AFK23308.1	Phyla duicis (Verbenaceae)
Cadinene Synthase	D00072 1	Disiuus sourceuris (Eurobarthiosooo)
$C_{\alpha}(+)$ delta codinene synthese	D93923.1	Communis (Euphoroideeae)
Ga(+)-delta-cadinene synthase	Q43/14.1	Gossypium arboreum (Malvaceae)
Gh (+)-delta-cadinene synthase	U88318.1	Gossypium nirsutum (Malvaceae)
Cru (1) commence and in one courth and	QSSBPS.I	Current and Construction
Cm (+)-gamma-cadmene synthase	B2KSJ3.1	<i>Cucumis melo</i> (Cucurbitaceae)
Nig beta-cubebene synthase	B31PQ6.1	<i>Magnolia granaijiora</i> (Magnoliaceae)
Ph copaene synthase	AVY55520.1	Piper nigrum (Piperaceae)
At beta-caryopnyllene/alpha-humulene	AF49/491.1	Arabiaopsis inaliana (Brassicaceae)
syntnase	AE4702(1 1	(
Aa beta-caryopnyllene synthase	AF4/2301.1	Ariemisia annua (Asteraceae)
Zinsp (E)-beta-caryopnynene synthase	EU239042.1	<i>Lea mays</i> subsp. raivigiumis (roaceae)

Pn caryophyllene synthase	ARB08605.1	Piper nigrum (Piperaceae)
Zp (E)-beta-caryophyllene synthase	ABY /9214.1	Zea perennis (Poaceae)
Gh (E)-beta-caryophyllene	XP_016727799.1	Gossypium hirsutum (Malvaceae)
Pl (E)-beta-caryophyllene synthase	ARQ83254.1	Phaseolus lunatus (Fabaceae)
Ok beta-caryophyllene synthase	AKA94109.1	Ocimum kilimandscharicum (Labiatae)
Li 9-epi-caryophyllene synthase	AGU13712.1	Lavandula x intermedia (Lamiaceae)
Zl (E)-beta-caryophyllene synthase	ABY79211.1	Zea luxurians (Poaceae)
Zmsm (E)-beta-caryophyllene synthase	EU259641.1	Zea mays subsp. Mexicana (Poaceae)
SI caryophyllene/alpha-humulene synthase	GU647162.1	Solanum lycopersicum (Solanaceae)
Vvcg (E)-beta-caryophyllene synthase 1	HM807374.1	Vitis vinifera (Vitaceae)
Pd beta-caryophyllene synthase	AFR23370.1	Phyla dulcis (Verbenaceae)
Vvcd (E)-beta-caryophyllene synthase	JF808010.1	<i>Vitis vinifera</i> cultivar deying 84-1 (Vitaceae)
Vvcg (E)-beta-caryophyllene synthase 2	HM807373.1	<i>Vitis vinifera</i> isolate VvGwECar1 cultivar Gewurztraminer (Vitaceae)
Mcvr (-)-beta-carvophyllene synthase	I6RAO6 1	Matricaria chamomilla var recutita
Mm beta-carvophyllene synthase	ACN67535 1	Mikania micrantha (Asteraceae)
Hl alpha-humulene synthase	B6SCF5 1	Humulus lupulus (Cannabaceae)
Ag gamma-humulene synthase	AAC05728 1	Abies grandis (Pinaceae)
Zz alpha-humulene synthase	B1B1U3 1	<i>Tingiher zerumbet</i> (Tigiheraceae)
Pg alpha-humulene synthase	AD745513 1	Picea glauca (Pinaceae)
Peng longifolene synthase	ADZ45515.1	Picea engelmannii r Picea glauca (Pinaceae)
Pa longifolene synthase 1	ΔΔS47695 1	Picea abies (Pinaceae)
Ps longifolene synthase	ARV44454 1	Pinus sylvestris (Pinaceae)
Ps alpha-longininene synthase	AD7455161	Picea sitchensis (Pinaceae)
Pa longifolene synthase ?	06751.0.1	Picea abies (Pinaceae)
Vo valerena-1 10-diene synthase	AGB05610.1	Valeriana officinalis (Valerianaceae)
7m 7-eni-sesquithuiene synthase	O6ID73 1	Zag mays (Poaceae)
$\Delta \alpha$ alpha-bisabolene synthase	Q0JD75.1	Abias grandis (Pinaceae)
Zm tns5-Dell	O6ID70 1	Zag mays (Poaceae)
Zo(S) beta hisabolene synthase	Q0JD70.1 D2V7D0	Zingihar officingla (Zigiharacaaa)
Δt (7) gamma bisabolene synthase	$O_{0}T_{0}I_{0}$	Arabidonsis thaliana (Brassicaceae)
Ch alpha bisabolana synthasa	ATT 10/200 1	Ginkgo hiloha (Ginkgooceae)
Dm(E) gamma bisabolona synthese	AIU94290.1	Dinkgo biloba (Olikgoaceae)
Sa bata bigabalana synthasa	Q4Q5IN4 E2W205	Santalum austrocaledonicum (Sontolocono)
Sa alpha bisabolol synthese	E3 W 203	Santalum austrocaleaonicum (Santalaceae)
Da E alpha bisabolono synthaso	L3 W 200 A A S 47680	Diaga abias (Dinagaga)
Ta E-alpha-bisabolene synthese	AA34/009	Ticeu ubles (Filiaceae)
Sh hata aagguinhallan drana gynthaga	DAI0/954.1	Sanghum hission (Decesso)
So biashalal authors	L J I III2	Sorgnum Dicolor (Poaceae)
Ze acquitemene synthese	ADU6/004.1	Zingihan officingle (Zigihanaceae)
Lo sesquiterpene synthase	BAI0/933.1	Lingiber Officinate (Zigiberaceae)
Ag E-alpha-disabolene synthase	AAC24192.1	Ables granais (Pinaceae)
Sa b-disadolene synthase U_{1}	ADU8/005.1	Santaium austrocaleaonicum (Santaiaceae)
Ha (Z)-gamma-bisabolene synthase K11 Us (Z) samma bisabalana synthase $K7$	AME10498.1	Helianthus annuus (Asteraceae)
Ha (Z)-gamma-bisabolene synthase K/	AME10497.1	<i>Helianthus annuus</i> (Asteraceae)
Zm(S) beta macrocarpene synthase 1 Zm(S) beta macrocarpene synthase 2	AA588570.1	Zea mays (Poaceae)
Zm(S)-beta-macrocarpene synthase 2	QIEG/2.1	Zea mays (Poaceae)
Ee bisabolol synthase	AYJ/1561.1	Eremanthus erythropappus
Pa(+)-epi-alpha-bisabolol synthase	J/LHII	Pnyla dulcis (Verbenaceae)
Aa bisabolol synthase	AF V40969.1	Artemisia annua (Asteraceae)
Am (+)-aipna-bisaboloi synthase	ВАW 54954.1	Artemisia maritima (Asteraceae)
AK (+)-alpha-bisabolol synthase	ва <i>w 3</i> 4955.1	Artemisia kurramensis (Asteraceae)

Mcvr a-bisabolol synthase	AIW00681.1	Matricaria chamomilla var. recutita
Aabs koidzumiol synthase	BAW34956.1	Artemisia absinthium (Asteraceae)
Aabs koidzumiol synthase	BAW34957.1	Artemisia absinthium (Asteraceae)
Aa amorpha-4,11-diene synthase	Q9AR04.2	Artemisia annua (Asteraceae)
Aa epi-cedrol synthase	CAC08805.1	Artemisia annua (Asteraceae)
Mcvr sesquiterpene synthase 1	AIG92846.1	Matricaria chamomilla var. recutita
		(Asteraceae)
Mcvr sesquiterpene synthase 4	AIG92848.1	Matricaria chamomilla var. recutita
		(Asteraceae)
Vvcpn beta-curcumene synthase	HM807381.1	Vitis vinifera isolate VvPNbCur cultivar
		Pinot Noir (Vitaceae)
Pc gamma-curcumene synthase	AY508726.1	Pogostemon cablin (Lamiaceae)
Vvcg (E)-alpha-bergamotene synthase	HM807376.2	Vitis vinifera isolate VvGwaBer cultivar
		Gewurztraminer (Vitaceae)
Pd trans-alpha-bergamotene synthase	J7LQ09.1	Phyla dulcis (Verbenaceae)
La exo-alpha-bergamotene synthase	Q2XSC4.1	Lavandula angustifolia (Lamiaceae)
Sa santalene bergamotene synthase 1	ADP30866.1	Santalum album (Santalaceae)
Sa santalene bergamotene synthase 2	ADP30867.1	Santalum album (Santalaceae)
Ss santalene synthase	ADO87002.1	Santalum spicatum (Santalaceae)
Sau santalene synthase	ADO87001.1	Santalum austrocaledonicum (Santalaceae)
Sb zingiberene synthase	C5YHH7	Sorghum bicolor (Poaceae)
Ob alpha-zingiberene synthase	Q5SBP4.1	Ocimum basilicum (Labiatae)
At alpha-barbatene synthase	Q4KSH9.2	Arabidopsis thaliana (Brassicaceae)

 Table S2 Oligo-nucleotide primers used in this study.

Name	Sequence (5'-3')	Purpose
pET SmSTPS3-F:	TTTCCATGGCTATGGCTACAACACAAGTAGAAAT	Protein expression
pET SmSTPS3-R:	ACGCGTCGACCTACATCTTGATAGGATCCACGAA	Protein expression
	С	
pET AaGAS-F:	CGGGAGCTCATGGCAGCGGTTCAAGCTAATG	Protein expression
pET AaGAS-R:	CGGCTCGAGTTACACGGGGTAGAGAACCCAC	Protein expression
pET AtTPS21-F:	CGGGATCCATGGGGAGTGAAGTCAACCG	Protein expression
pET AtTPS21-R:	CGGAGCTCTCAAATGGGTATAGTTTCAATGAGCA	Protein expression
pET PTSiso-F:	TTTGGATCCATGGAGTTGTATGCCCAAAGTGTTG	Protein expression
pET PTSiso-R:	TTTGAGCTCATATGGAACAGGGTGAAGGTACAAC	Protein expression
	Т	
pET AaCES-F:	CGCGGATCCATGTCTCTTATAGTAGAGGATGTC	Protein expression
pET AaCES-R:	AAAGAGCTCTCATGTAATGATGGCATCAAC	Protein expression
SmSTPS3-C438G-F:	GGTCTCCTCATTGGCCGAGTCAT	mutagenesis
SmSTPS3-C438G-R:	ATGACTCGGCCAATGAGGAGACC	mutagenesis
SmSTPS3-C438S-F:	GGTCTCCTCATTAGCCGAGTCAT	mutagenesis
SmSTPS3-C438S-R:	ATGACTCGGCTAATGAGGAGACC	mutagenesis
AaGAS-S454G-F:	CTGAGTTGATTGGAAGACTTCAAG	mutagenesis
AaGAS-S454G-R:	CTTGAAGTCTTCCAATCAACTCAG	mutagenesis
AaGAS-S454C-F:	CTGAGTTGATTTGTAGACTTCAAG	mutagenesis
AaGAS-S454C-R:	CTTGAAGTCTACAAATCAACTCAG	mutagenesis
AtTPS21-S441G-F:	CTGAGATAATAGGTCGATTCACA	mutagenesis
AtTPS21- S441G-R:	TGTGAATCGACCTATTATCTCAG	mutagenesis
AtTPS21- S441C-F:	CTGAGATAATATGTCGATTCACA	mutagenesis
AtTPS21- S441C-R:	TGTGAATCGACATATTATCTCAG	mutagenesis
PTSiso-A445G-F:	CGTTAATCATTGGTAGGCTCATC	mutagenesis
PTSiso-A445G-R:	GATGAGCCTACCAATGATTAACG	mutagenesis
PTSiso-A445C-F:	CGTTAATCATTTGTAGGCTCATC	mutagenesis
PTSiso-A445C-R:	GATGAGCCTACAAATGATTAACG	mutagenesis
PTSiso-A445S-F:	CGTTAATCATTAGTAGGCTCATC	mutagenesis
PTSiso-A445S-R:	GATGAGCCTACTAATGATTAACG	mutagenesis

mutagenesis mutagenesis

Table S3 Product profiles of SmSTPS3, AaGAS and their mutants (%).

Compounds	SmSTPS3 Wt	SmSTPS:3C438G	SmSTPS:3C438S	AaGAS Wt	AaGAS:S454G	AaGAS:S454C
(-)-5-epieremophilene	100.0	39.4	92.2	_	_	_
β-elemene	a	5.9	5.3	93.5	19.5	93.8
elemol	_	54.7	2.5	_	74.4	_
unknown sesquiterpene	_	_	_	6.5	0.8	6.2
β-caryophyllene	—		_	_	4.3	_
α-humulene	—		—	—	1.0	—

 Table S4 Product profiles of AtTPS21 and its mutants (%).

Compounds	AtTPS21 Wt	AtTPS21:S441G	AtTPS21:S441C
β-caryophyllene	92.9	24.5	52.9
α-humulene	7.1	3.4	10.0
elemol	a	69.2	—
guaiol		2.9	
α-copaene		—	6.2
β-copaene			3.3
γ-muurolene			1.9
germacrene D			1.3
cubebol			2.4
δ-cadinene			1.6
germacrene D-4-ol		—	6.7
copaborneol			5.1
torreyol			8.6

Table S5 Product profiles of PTSiso and its mutants (%).

Compounds	PTSiso	PTSiso:	PTSiso:	PTSiso:
	Wt	A445S	A445G	A445C
β-patchoulene	1.8	1.9	1.2	3.2
β-elemene	6.4	11.3	6.4	37.5
β-caryophelene	3.2	2.0	1.8	4.6
α-guaiene	14.6	7.0	9.6	7.2
seychellene	6.4	3.4	2.7	1.2
α-patchoulene	3.4	3.0	1.8	1.8
γ-patchoulene	1.4	2.4	0.8	1.6
guai-4,11-diene	4.9	6.0	3.6	4.2
α-bulnesene	26.4	15.4	30.3	11.4
unknown sesquiterpene	1.6	1.7	1.0	5.6
elemol	a		4.8	1.4
unknown sesquiterpene				5.3
alcohol				
pogostol	2.6	2.6	3.4	2.6
patchoulol	27.2	43.3	12.4	11.1

Compounds	RI recorded	RI of reference
elemol	1550	1550 ^[11]
guaiol	1597	1597 ^[12]
α-copaene	1377	1377 ^[9]
β-copaene	1430	1430 ^[9]
γ-muurolene	1478	1478 ^[9]
germacrene D	1481	1481 ^[9]
cubebol	1517	1517 ^[9]
δ-cadinene	1525	1525 ^[9]
germacrene D-4-ol	1577	1577 ^[9]
copaborneol	1608	1608 ^[9]
torreyol	1647	1647 ^[9]
farnesol	1700	1700 ^[11]

Table S6 Retention indices of new products generated by mutant enzymes.

Table S7 Kinetic parameters of SmSTPS3, AaGAS, AtTPS21, PTSiso and their mutants with FDP.

	<i>Km</i> (µM)	kcat (s ⁻¹)	kcat/Km (s ⁻¹ mM ⁻¹)
SmSTPS3 WT	6.56 ± 0.43	$(0.44 \pm 0.02) \times 10^{-1}$	6.71 ± 0.45
C486G	5.87 ± 0.61	$(0.34 \pm 0.03) \times 10^{-1}$	5.79 ± 0.56
AaGAS WT	5.39 ± 0.37	$(0.56 \pm 0.04) \times 10^{-1}$	10.39 ± 0.72
S454G	5.57 ± 0.51	$(0.48 \pm 0.03) \times 10^{-1}$	8.62 ± 0.49
AtTPS21 WT	3.43 ± 0.32	$(0.17 \pm 0.02) \times 10^{-1}$	4.91 ± 0.41
S439G	3.26 ± 0.30	$(0.12 \pm 0.02) \times 10^{-1}$	3.68 ±0. 37
S439C	5.13 ± 0.45	$(0.06 \pm 0.01) \times 10^{-1}$	1.17 ± 0.09
PTSiso WT	6.83 ± 0.52	$(0.32 \pm 0.02) \times 10^{-2}$	0.46 ± 0.03
A455G	6.46 ± 0.48	$(0.43 \pm 0.04) \times 10^{-2}$	0.66 ± 0.05
A455S	4.14 ± 0.58	$(0.22 \pm 0.01) \times 10^{-2}$	0.53 ± 0.04
A455C	9.85 ± 0.49	$(0.34 \pm 0.03) \times 10^{-2}$	0.34 ± 0.03

References

- [1] Fang X, Li JX, Huang JQ, et al., Systematic identification of functional residues of Artemisia annua amorpha-4,11-diene synthase, Biochem J 2017; 474: 2191–2202.
- [2] Fang X, Li CY, Yang Y, et al., Identification of a novel (-)-5-epieremophilene synthase from Salvia miltiorrhiza via transcriptome mining. Front Plant Sci 2017; 8: 627.

20.0

- [3] Bertea CM, Voster A, Verstappen FW, et al., Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. Arch Biochem Biophys 2006; 448: 3–12.
- [4] Tholl D, Chen F, Petri J, et al., Two esequiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from *Arabidopsis* flowers. Plant J 2005; 42: 757–771.
- [5] Hartwig S, Frister T, Alemdar S, et al., Expression, purification and activity assay of a pathchoulol synthase cDNA variant fused to thioredoxin in *Escherichia coli*. Protein Expres Purif 2014; 97: 61–71.
- [6] Deguerry F, Pastore L, Wu S, et al., The diverse sesquiterpene profile of patchouli, *Pogostemon cablin*, in correlated with a limited number of sesquiterpene synthases. Arch Biochem Biophys 2006; 448: 123–136.
- [7] Hua L, Matsuda SPT, The molecular cloning of 8-epicedrol synthases from *Artemisia annua*. Arch Biochem Biophys 1999; 369: 208–212.
- [8] Liang J, Liu J, Brown R, et al., Direct production of dihydroxylated sesquiterpenoids by a maize terpene synthase. Plant J 2018; 94: 847–856.
- [9] Garms S, Köllner TG, Boland W, A multiproduct terpene synthase from *Medicago truncatula* generates cadalane sesquiterpenes via two different mechanisms. J Org Chem 2010; 75: 5590–5600.
- [10] Vardakou M, Salmon M, Faraldos JA, et al., Comparative analysis and validation of the malachite green assay for the high throughput biochemical characterization of terpene synthases, MethodsX 2014; 1: 187-196.
- [11] Xie Y, Wang J, Yang F, et al., Comparative analysis of essential oil components of two *Cryptomeria* species from China. Ind Crop Prod 2011; 34: 1226–1230.
- [12] Dural H, Bagci Y, Ertugrul K, et al., Essential oil composition of two endemic *Centaurea* species from Turkey, *Centaurea mucronifera* and *Centaurea chrysantha*, collected in the same habitat. Biochem Syst Ecol 2003; 31: 1417–1425.