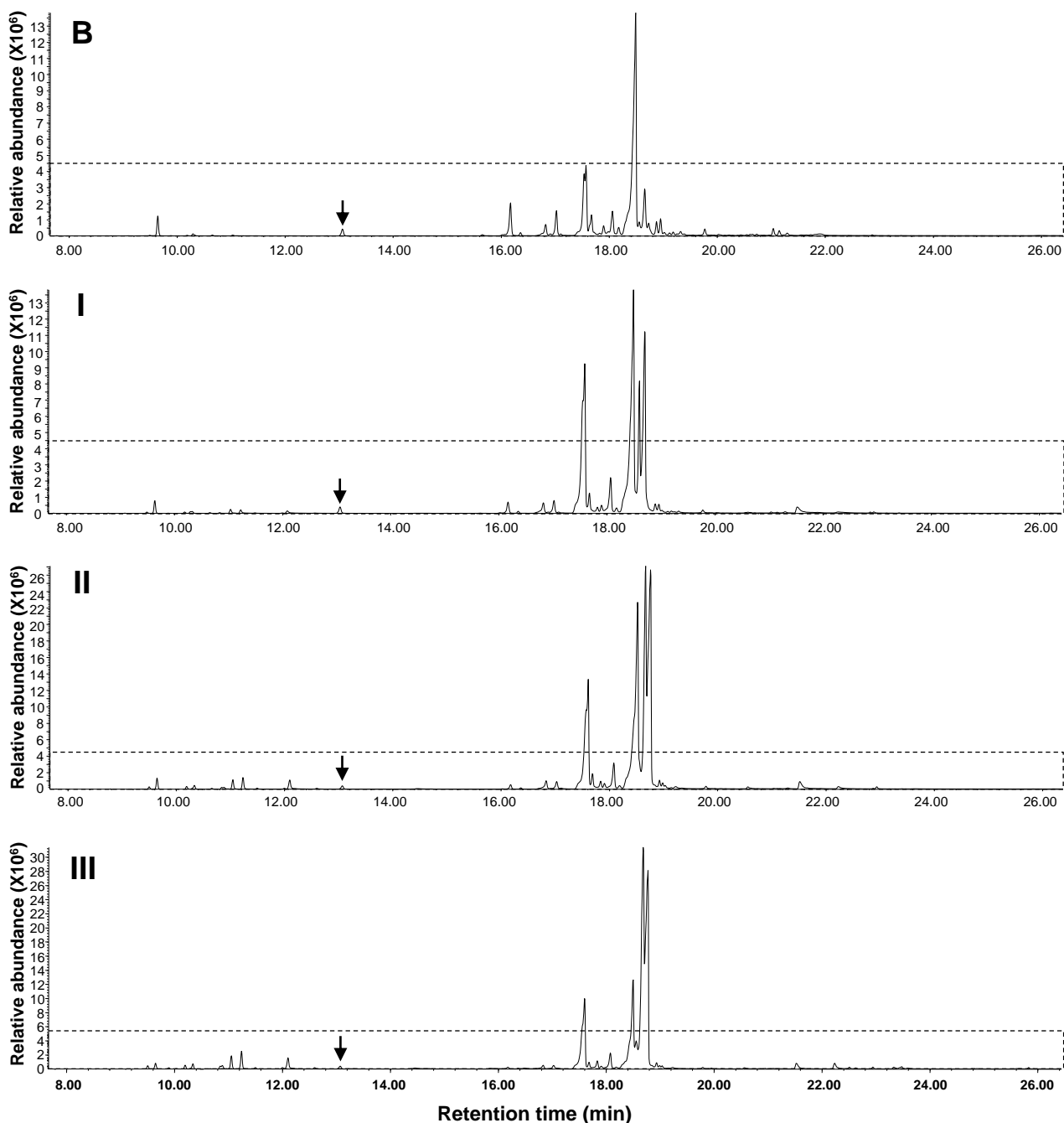
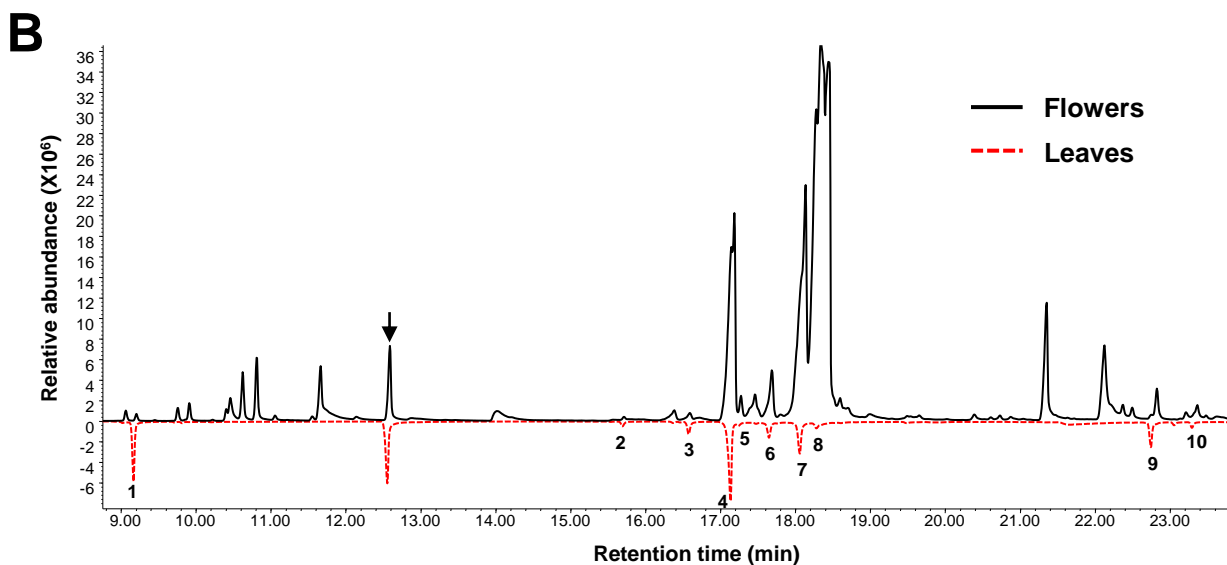
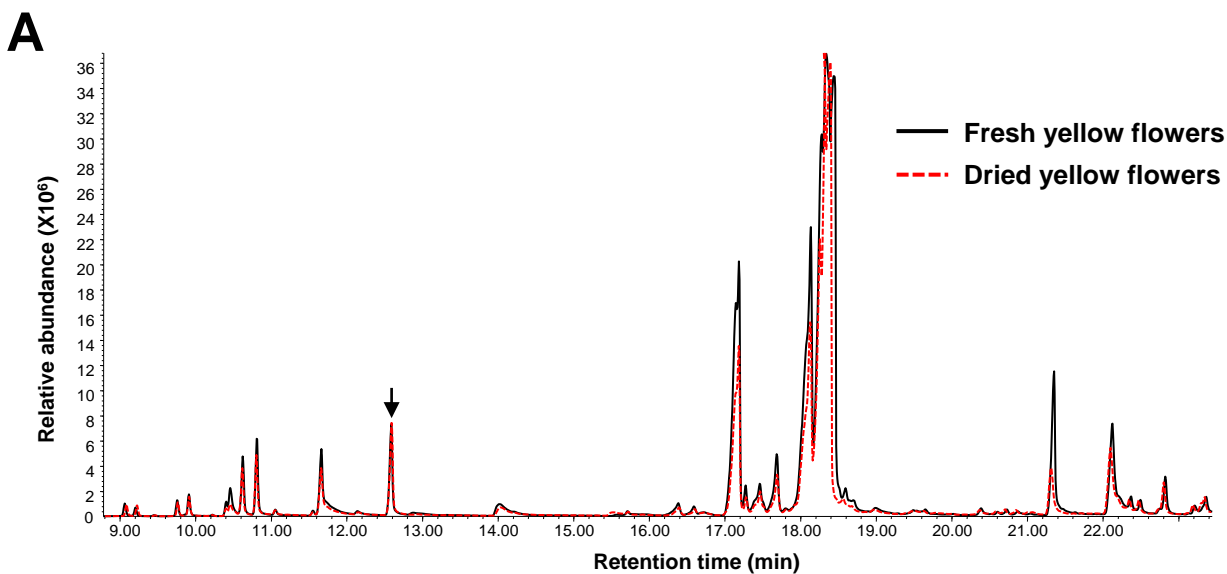


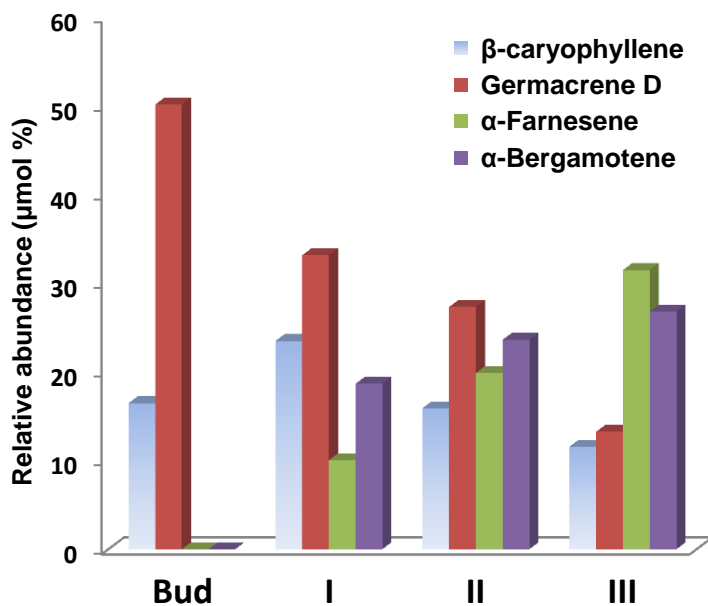
Supplementary Figures



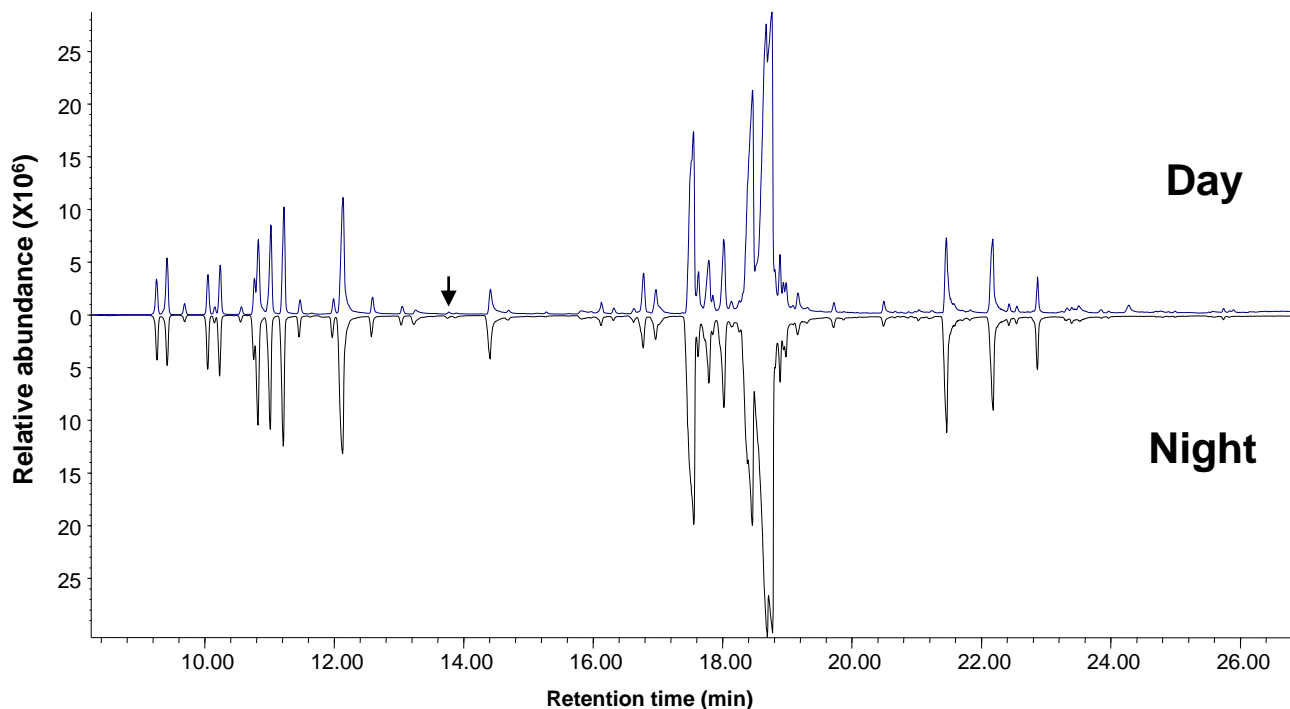
Supplementary Fig. S1. Compositional variation of dwarf ylang ylang essential oils during flower development. B, floral buds; I, undeveloped small flowers; II, mature green flowers; III, fully mature yellow flowers. The arrows indicate the peak of camphor, the internal standard used in the assay.



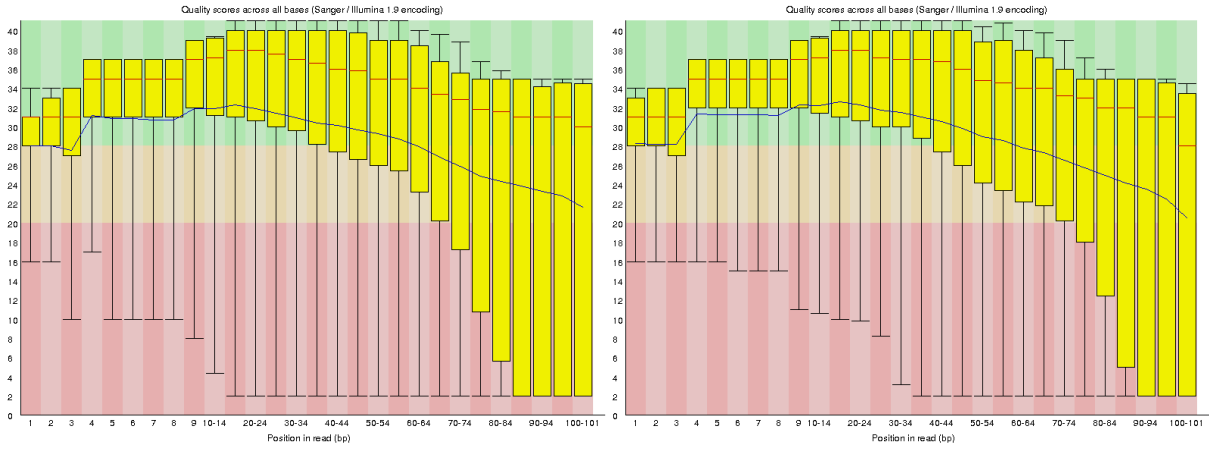
Supplementary Fig. S2. Total ion chromatograms of essential oils from dwarf ylang ylang flowers. (A) GC traces showing no difference between fresh and dried flowers, (B) GC traces of flowers and leaves from dwarf ylang ylang. 1, α -pinene; 2, Elixene; 3, β -elemene; 4, β -caryophyllene; 5, γ -muurolene; 6, Humulene; 7, Germacrene D; 8, γ -elemene; 9 and 10, phytol. The arrows indicate the peak of camphor, the internal standard used in the assay.



Supplementary Fig. S3. Variation of four major terpenes during flower development. Bud, floral buds; I, undeveloped small flowers; II, mature green flowers; III, fully mature yellow flowers.



Supplementary Fig. S4. Total ion chromatograms of essential oils from dwarf ylang ylang mature yellow flowers sampled at day and night. The arrows indicate the peak of camphor, the internal standard used in the assay.



Supplementary Fig. S5. Quality of deep sequencing. The sequence quality was evaluated by FastQC.

A

```

CoDXR 1 MPLSLCTAMALKPFSESPFISGISFKGSSRGN--LHKLQGGFAFKGKESRTSVRVGISCL-
RcDXR 1 -----MALNLLSPAEIKSVSLDSTRSSQ-LPKLPGSFLSKRKEFGR---KVQCS-
EuDXR 1 -----MALNLLPRTFNPVSPFHTSKSNRNLFPNLQGGFAFKRKRDIATNGLRVHCSA
ZmDXR 1 -----MAALKASFRGELSAASFLDSSRGP--LVQHKVDFTEQRKGRKRAISLRRTCCS-

CoDXR 58 ----AKAPPPAWPGRAVPEPFSKRTWVGPKPISIVGSTGSIQTQLDIVAENPDKFKVVA
RcDXR 47 ----AQAPPPAWPGRAVAEPGRKTDGPKPISIVGSTGSIQTQLDIVAENPDKFRVVA
EuDXR 53 EEVGVAVAPPPAWPGRAVVEPGRKSWDGPKPISIVGSTGSIQTQLDIVAENPDKFKVVA
ZmDXR 51 ----MQQAPPPAWPGRAVAEPGRRSWDGPKPISIVGSTGSIQTQLDIVAENPDKFRVVA

CoDXR 113 LAAGSNVTLLADQVKTRFPQLVSI RNESLIGELKEALADADYKPEIIPGEEGLIEVARHP
RcDXR 102 LAAGSNVTLLADQVKTRFPQLVSVRDESLVDELKEALADVDEKPEIIPGEEQGVIEVARHP
EuDXR 113 LAAGSNVTLLADQVKTRFPQLVAVRNESLVDLKEALADAEYTPPEIIPGEEQGVIEVARHP
ZmDXR 107 LAAGSNVTLLADQVKTRFKLVAVRNESLVDLKEALADCEEKPEIIPGEEQGVIEVARHP

CoDXR 173 DAMTVVTVGIVGCAGLKPTVAAIEAGKDIALANKETLIAGGPFVLP LAHKHKVKILPADSE
RcDXR 162 DAVSVTVGIVGCAGLKPTVAAIEAGKDICLANKETLIAGGPFVLP LAHKHYNVKILPADSE
EuDXR 173 DAVTVVTVGIVGCAGLKPTVAAIEAGKDIALANKETLIAGGPFVLP LAHKHNVKILPADSE
ZmDXR 167 DAVTVVTVGIVGCAGLKPTVAAIEAGKDIALANKETLIAGGPFVLP LAHKHKVKILPADSE

CoDXR 233 HSAIFQCIQGLPEGALRRIILTASGGAFRDLPVEKLDVVKVADALKHPNWSMGKKITVDS
RcDXR 222 HSAIFQCIQGLPEGALRRIILTASGGAFRDWPVEKLDVVKVADALKHPNWSMGKKITVDS
EuDXR 233 HSAIFQCIQGLPEGALRRIILTASGGAFRDLPVDKLDVVKVADALKHPNWSMGKKITVDS
ZmDXR 227 HSAIFQCIQGLSPEGALRRIILTASGGAFRDWPVDRDKLDVVKVADALKHPNWSMGKKITVDS

CoDXR 293 ATLFNKGLEVIEAHYLFGAEDYDNI EIVIHQSI IHSMIETQDSSVLAQLGLFDMRFLPILY
RcDXR 282 ATLFNKGLEVIEAHYLFGADYDNI EIVIHQSI IHSNVETEDSSVLAQLGWPMRIPILY
EuDXR 293 ATLFNKGLEVIEAHYLFGAEDYDIEIVIHQSI IHSMIETQDSSVLAQLGWPMRIPILY
ZmDXR 287 ATLFNKGLEVIEAHYLFGAEDYDIEIVIHQSI IHSNVETQDSSVLAQLGWPMRIPILY

CoDXR 353 TMSWPERIFCSEVTWPRDLCKLGLSTFRAPDNV KYPSEMLAYAAGRAGTMTGVLSAAN
RcDXR 342 TMSWPDRIYCS EITWPRDLCKLGLSTFKAPDNV KYPSEMLAYAAGRAGTMTGVLSAAN
EuDXR 353 TMSWPDRIYCS EITWPRDLCKLGLSTFKMPDNV KYPSEMLAYAAGRAGTMTGVLSAAN
ZmDXR 347 TMSWPDRIYCS EVTWPRDLCKLGLSTFRAPDNV KYPSEMDLAYAAGRAGTMTGVLSAAN

CoDXR 413 EKAVEMFIDEKISYLDIFKIVELTCDKHRAELVTSPLSEIIVHYDLWARVYAASLQPSDD
RcDXR 402 EKAVEMFIDEQISYLDIFKIVELTCDKHRAELVTSPLSEIIVHYDLWARVYAASLQPSGG
EuDXR 413 EKAVEMFIDEKISYLDIFKVVVELTCDKHRAELVTSPLSEIIVHYDLWARVYAASLQPSGG
ZmDXR 407 EKAVELFIDEKISYLDIFKVVVELTCAHARNELVTSPLSEIIVHYDLWARVYAASLQPSGG

CoCMK 1 MASTHFVCQHFLLSFGNGKNTPGSLSYKNSSLFTPLR-HRRASL GKSTQKORYPHVRA
CrCMK 1 MASSQSLCGYQLYTCS--SGKTQLNSFKKGSISVSSTATPHGFLSFGQNPQFQRALFVRA
SmCMK 1 MASS---SSQFLCAH--NPKPHFNSYTNATLPOFSSFKPNGSSSLRKKIQSRIHLIRA
LjCMK 1 MASSHFLCGQHLYSSS--HGRTKISL SFKKGLFQSSSCRPNNGSF SFDKKTQYQRTQLVKS

CoCMK 60 TAS-DSKASRKQVEIVYDFDERLNR LADEV DKNAGLQRLS LFS PCKINVFLRITRKREDG
CrCMK 59 TASSDKTGRKQVEIIVYDPEEELNKLAD EVDKNAGLSRLN LFS PCKINVFLRITGKREDG
SmCMK 55 TAS-DSTTGRKQLEVYVYDLENKLNKLAD EVDKAGISRLT LFS PCKINVFLRITGKRADG
LjCMK 59 MAA-DSKTGKQVEIVYDFDEKMSL LADEV DKNAGLSRLS LFS PCKINVFLRITSKREDG

CoCMK 119 FHDLASLFHVISLGDITKIFSLSPTRKKDQLSTNVPGVPLDDRNLIIKALNLYRKKTGTDN
CrCMK 119 YHDLASLFHVISLGDKIKFSLAPSKSKDRLSTNVSGVPLDDKLNIIKALNLYRKKTGTDN
SmCMK 114 FHDLASLFHVISLGDKIKFSLSPSKSTDR LSTNVPGVPLDERNLI I KALNLYFRKKTGVND
LjCMK 118 FHDLASLFHVISLGDKIKFSLAPSKSKDRLSTNVPGVPLDDSNLIIKALNLYRKKTGSDK

CoCMK 179 YFWIHLDKKVP TGAGLGGGSSNAATALWAANQFS GGLATEKELQEWSGEIGSDCPFFFSH
CrCMK 179 FFWVHLDKKVP TGAGLGGGSSNAATALWAANQFAGGIIASENDLQEWSGEIGSDIPFFFSH
SmCMK 174 YFWIHLDKKVP TGAGLGGGSSNAATALWAANQFS GCVASEKDLQEWSGEIGSDIPFFFSH
LjCMK 178 FFWIHLDKKVP TGAGLGGGSSNAATALWAANQFS GGLASEKELQEWSGEIGSDVPPFFSH

CoCMK 239 GAA YCTGRGEVVEDLPPPIPLDIPMVLIKPFAQACPTGEVYKRLRLDQTNVDP LTLLENI
CrCMK 239 GAA YCTGRGEVVEDIPSPDIPMVLIKPFAQACSTA EYVYKRLRLLEETSLIDPLILEKI
SmCMK 234 GAA YCTGRGEVVEDIPPVFLDIPMVLIKPFAQACPTGEVYKRLRMNQTSQIDPLV LLEKI
LjCMK 238 GAA YCTGRGEVVEDVTLPIGFDVPMVLIKPFAACSTA EYVYRRFRLDQTSNIDPQT LLEKI

CoCMK 299 TRTGISQDVCVNDLEPPAF EVLPSLKLKRLI LAAGRGEYDAVFMMSGSGSTIVGIGSEFP
CrCMK 299 SKAGISQDVCVNDLEPPEVFDVLP SLKRLKQVLAAGRQYV GAVFMMSGSGSTIVGVGSPDP
SmCMK 294 SKGGISQDVCVNDLEPPAF EVVPSLKRLKQRI AAGRQYDAVFMMSGSGSTIVGVGSPDP
LjCMK 298 SLNGISQDVCVNDLEPPAF EVLPSLKLKRLKQRI I AAGRGEYDAVFMMSGSGSTIVGIGSPDP

CoCMK 359 PQFVYDEDEHRDTEVFSEARFLTRGENQWYTEMSS TASSFESREDLASTVQ
CrCMK 359 PQFIYDDEEYKDVLSLSEASF LTRPPNQWYSEFGLSTACSSSDFQSSE
SmCMK 354 PQFVYDDEEYKNVFLSDAKFITR SAHQWYSEP-LSTDESTCDVE-----
LjCMK 358 PQFVYDDEYKDVFLSEASFITRAEN EYTEE-FSTNNASAPSGLSYATE

```

Supplementary Fig. S6. Alignment of deduced amino acid sequences of representative genes involved in biosynthetic pathways for VOCs. (A) Alignment of deduced amino acid sequences of two representative genes, DXR and CMK, involved in MEP pathway. DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; CMK, 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase. Co, *Cananga odorata*; Cr, *Catharanthus roseus*; Eu, *Eucommia ulmoides*; Lj, *Lonicera japonica*; Rc, *Ricinus communis*; Sm, *Salvia miltiorrhiz*; Zm, *Zea mays*. Accession numbers: RcDXR (XP_002511399), EuDXR (AFU93070), ZmDXR (NP_001105139), CrCMK (ABI35992), SmCMK (ABP96842), LjCMK (AGE10581). Completely conserved residues are shaded in dark gray, identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

B

| | | |
|--------|-----|--|
| CoHMGS | 1 | MDSQ-KDVGILAMDIYPPPTCVLQDALEHDHGASKGKYITIGLQDCMAFPCTEVEDVISMS |
| NtHMGS | 1 | MEFKANDVIGILAVDIYPPPTCVQQAELRAYDAGASKGKYITIGLQDCFLAPCTELEDVISMS |
| GsHMGS | 1 | MA---KNVIGILADIDYPPPTCIQQELLEAHADGASKGKYITIGLQDCMAFPCTEVEDVISMS |
| PnHMGS | 1 | MAEQ-KNVIGILAMEIYPPPTCIQQEVLRAHDGASKGKYITIGLQDCMGFPCTEVEDVISMS |
| CoHMGS | 60 | LTVVTSLLEKYGVDFKQIGRLEVGSSETVIDKSKSIKTLWLMQIPEKKGNTDIEGVDSTNAC |
| NtHMGS | 61 | LTVVTSLLEKYQIDPKMIGRLEVGTETVIDKSKSIKTLWLMQIPEHGNNTDIEGVDSTNAC |
| GsHMGS | 58 | LTVVSSLLEKYAIDPKQIGRLEVGSSETVIDKSKSIKTFIMQIPEKYGNNTDIEGVDSTNAC |
| PnHMGS | 50 | LTVVTSLLEKYKIDPKQIGRLEVGSSETVIDKSKSIKTFIMQIPEKCGNTDIEGVDSTNAC |
| CoHMGS | 120 | YGGTAALFNCVNVVNESSWDGRFGLVVCADSAYVAEGPARPTGGAAAAMLIQHPAPITVE |
| NtHMGS | 121 | YGGTAALFNCVNVVNESSWDGRYGLVVCADSAYVAEGPARPTGGAAAAMLIQHPAPITAF |
| GsHMGS | 118 | YGGTAALFNCVNVVNESSWDGRYGLVVCADSAYVAEGPARPTGGAAAAMLIQHPAPITSE |
| PnHMGS | 120 | YGGTAALFNCVNVVNESSWDGRYGLVVCADSAYVAEGPARPTGGAAATIAMIQITDAPITVE |
| CoHMGS | 180 | ENKYRGTTHMAHVDFYKPNLASEYPPVVDGKLSQTCYLMALDSCYKRFSSKYEKLEKPKPFS |
| NtHMGS | 181 | ESKYRGTTHMSHVDFYKPNLASEYPPVVDGKLSQTCYLMALDSCYKRFCAKYEKFEKQKQFS |
| GsHMGS | 178 | ESKLRGSHMAHAYDFYKPNLASEYPPVVDGQLSQTCLMALDSCYNHLSHKYEKQKQKQFS |
| PnHMGS | 180 | ESKFRGSHMSHAYDFYKPNLASEYPPVVDGKLSQTCYLMALDSCYKRYCKKYEKLEKQKQFS |
| CoHMGS | 240 | ISDADYFVPHSPYNKLVQKSPARLYYNDPLRNPFSVENDARIKLESFSSLSGDESYQNRD |
| NtHMGS | 241 | ISDADYFVPHSPYNKLVQKSPARLYYNDPLRNPSSVDKEAREKLEPFSSLSGNSYQSRD |
| GsHMGS | 238 | ISDAEYFVPHSPYNKLVQKSPARLYYNDPLRNPASVDEAAKLEKLEPFATLSGDESYQSRD |
| PnHMGS | 240 | MDDADYFVPHSPYNKLVQKSPARLYYNDPLRNPASVDESAKEKLEKLEPFATLSGDESYASRD |
| CoHMGS | 300 | LENVSQQVAKQLYDAKVQFSTLLFKQVGNMYTASLYAAAFATVLEHNKHSSTLEGKRVVMPFSY |
| NtHMGS | 301 | LENVSQQVAKNLYDAKVQFATLVKQVGNMYTASLYAAAFASVLEDKHSSTLAGQRIVMPFSY |
| GsHMGS | 298 | LENASQQVAKPQYDAKVQFTLLIPKQVGNMYTASLYAAAFISLHKNKHSSTLAGKRVVMPFSY |
| PnHMGS | 300 | LENATQQVAKSQYDAKVQFTLLIPKQVGNMYTASLYAAAFASLHKNKHSSTLAGKRVVMPFSY |
| CoHMGS | 360 | GSGLSSTMFSPFLQEGQHPFSLSNIVSLLDVYRKLESRHTFPPEKPFVETMKLMEHRYGKQ |
| NtHMGS | 361 | GSGLSSMFSLRLQDQHPFSLSNIDNVNNSGKLEARHVFPPEKPFVETMKLMEHRYGAK |
| GsHMGS | 358 | GSGLTSTMFSLLLREGQHPFSLSNIDKMDVAGKLSRHEFPPEKPFVETMKLMEHRYGKQ |
| PnHMGS | 360 | GSGLSATMFSLRLREGQHPFSLSNIANVMNVAELKLSRNEFPPEKPFVETMKLMEHRYGAK |
| CoHMGS | 420 | DFVIDIKGTSLSPGTFYLTQVDSMYRERYAKKAGEKPTISYENGSLPNGH |
| NtHMGS | 421 | DFVT-AKDTSLSPGTFYLTQVDSMYRERYSRK-GLNEKSSAVANGTLANGH |
| GsHMGS | 418 | EFVT-SKDTSLSPGTFYLTQVDSMYRERYAKK-TS-----ENGLVTNGH |
| PnHMGS | 420 | DFVT-SKDCSLSPGTYLTQVDSMYRERYAKK-AVDKTTIGTENGLTLANGH |
| CoPMK | 1 | MAEVVISAPGKVLVTGGYLVLEHSNPGIIVLSTARFYAIVKPLYEAVDFNSAWAWTDVK |
| PnPMK | 1 | -MAIVASAPGKVLMTGGYLILERPNEGVLVSTNARFYAIVKPLCDELKPDSSAWAWTDVK |
| ZmPMK | 1 | -MEVVASAPGKVLIIAGGYLVLERPNAGLVLSTARFYAVVRPLRDSLPAWTDVKT |
| AePMK | 1 | -MAVVASAPGKVLMTGGYLVLERPNAGIIVLSTNARFVSVVKEIYDEVKPDSSAWAWADVK |
| CoPMK | 61 | LSSPQLPRETSYKLSLRNLTLCISPRPDRNPFIEQAVQYSVAAAHSMCSDKGMKDGLHK |
| PnPMK | 60 | LTSPQMARRETYKMSLKHLLQLCASSNSRNPFEVEYAVQYSVAAAYATL-DNDKKNALHK |
| ZmPMK | 60 | VTSPQLSEVATYKLSLNKTTLLQTSRRESTNPFVEQAIQFSVAAAKAIIIDKERKDVVDK |
| AePMK | 60 | LTSPQMSREMTYKLSLKYLLTQSVLSLSRNPFEVEYAVQYVVAAYASRL-DSSGKDALTK |
| CoPMK | 121 | LLLQGLDITAIAGCNDFYSYRNQIEANCIPLAPDVLASIPPFSPINFNKENSSTIVREQS |
| PnPMK | 119 | LLLQGLDITILIGCNQFYSYRNQIEALGLPLSPESFATIKKXPTISITFNAGESNG---ENS |
| ZmPMK | 120 | LLLQGLNLTIIIGNDFYSYRQIEARGLPLTPEVLLSLPPFSSITFNSEVANGTMTGKRC |
| AePMK | 119 | LLLIRGLDITILGCNEFYSYRNQIEARGLPLTPESSLSLPPFSSITFNKEESG---QNS |
| CoPMK | 181 | KPEVAKTGLGSSAAMTTAVVAAVLYQLGVVDLSSSTAGNPHGTICNPDLDLVHVAQTAHC |
| PnPMK | 175 | KPEVAKTGLGSSAAMTTVVVAALLSYLQVNNLSLSLSDQNOEMDTADLDLVHVAQTAHC |
| ZmPMK | 180 | KPEVAKTGLGSSAAMTTSVVAALLHYLQAVNLSCPGQSSGDNASGRELDLVHTIAQSAHC |
| AePMK | 175 | KPEVAKTGLGSSAAMTTAVVASLLHYLQVNNLSVVKDN-----SELDLTVHVAQTAHC |
| CoPMK | 241 | IAQGKVGSGFDVSSAAVYGSQRYIRFSPVLSPAQVAVTGQPLEEVISHILKEKWDHEKIQ |
| PnPMK | 235 | IAQGKVGSGFDVSSAVYGSQRYVRFSPFVLSAQGAVGGQPLEVITDVLKKGWDHERTK |
| ZmPMK | 240 | IAQGKIGSGFDVSSAAVYGSQRYVRFSPRILSSAQ-AIGGTVLPDVSVDVLTQRWDHENKQ |
| AePMK | 229 | IAQGEVGSFGFDVSSAVYGSQRYVRFSPGVISSAQDAVKAAPLEEVINDVLKAEWDHEK-D |
| CoPMK | 301 | FSLPPLMTLLLEGPPTGGSSSTPSMVGAVKQWQRSEPOKSAETWTRLAKANSMPFIQIALAL |
| PnPMK | 295 | FSLPPLMMLLLEGPPTRGSSSTPSMVGAVKMQKSDPQRSRDTWTKLSNANSALTEQLNLL |
| ZmPMK | 299 | FSLPPLMTLLLEGPPTGGSSSTPSMVGAVKRWLKSDEPKSRDVTWSKLAIANSTLENQIRIL |
| AePMK | 288 | MSHAPLMTLSIR-EPGTGGSSSTPSMVGAVKQKADPQTSRDTWRKLSGNAALEMQLNLT |
| CoPMK | 361 | KKYRQEKMETYKIVIGSCSAHSEKWLQATDFCQEGIRSLIAVRDAMDLDIRFHMROMG |
| PnPMK | 355 | RKLAEEHWDAYKCVISSCNMCKSEEMGQASEPSQVQIVKALLGSRDATLEIRCOMRQMG |
| ZmPMK | 359 | NGISENHHEAVESMVRSCSHLYGKWAIEVATNQHQLIIRSLAARDACTLEIRLHMREMG |
| AePMK | 347 | SNLARMSFDVYKDVINNCSSTLPSSEKWLQVATEPSTRDQVKAALLGAKVACLEIRYQMRMG |
| CoPMK | 421 | QAAGAPIEPESQTLLLDAMNLEGVLFAGVPGAGGFDAIFVVLGDTN-NNVANVWCSSQG |
| PnPMK | 415 | DAAGVPIEPESQTLRLDAMKMEGVLLAGVPGAGGFDAIFAVTLGDASSTNLTAKWSSHN |
| ZmPMK | 419 | IAAGVPIEPDSQTRLLDAMNMEGVLLAGVPGAGGFDAVFSVVLGDAS-NAVAHWSSVVG |
| AePMK | 407 | EAAGVPIEPESQTLLESLTMNMEGVLLAGVPGAGGFDAVFAVTLGDAS-DKVIKWSRQNG |
| CoPMK | 480 | VLPMLVREDPRGLCLESGDPEPTEKISSAFSAIQV- |
| PnPMK | 475 | VLAMLVREDPRGVSLQSSDPEATEITSGISAVHIE |
| ZmPMK | 478 | VLPPLVREDCRGVSLLEDADPETREVSAAVWSIQIN |
| AePMK | 466 | VIALLVREDPNGVLENNSEAKEVTSGVSAIQIQ |

Supplementary Fig. S6 (Continued). (B) Alignment of deduced amino acid sequences of two representative genes, HMGS and PMK, involved in MVA pathway. HMGS, hydroxymethylglutaryl-CoA synthase; PMK, phosphomevalonate kinase. Ae, *Arnebia euchroma*; Co, *Cananga odorata*; Gs, *Glycine soja*; Nt, *Narcissus tazetta*; Pn, *Panax notoginseng*; Zm, *Zea mays*. Accession numbers: NtHMGS (AHF81872), GsHMGS (KHN14128), PnHMGS (AIK21781), PnPMK (AIK21784), ZmPMK (NP_001149345), AePMK (ABY27562). Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

C

```

CoDAHPS 1 MALSTSSSSLSIKTPLQSSIFPASKAHQHSLEPLLPTK---PEKPTSAVHAAEKSKNFV
VvDAHPS 1 MALSN-ASTLSSKSVYQTHSLFSSLSHQSSSLVPSKTHLESHPISAVHAAEPAKNFV
MtDAHPS 1 MSLSSSTSSLSLIP---TKSLLOQTKENPPSFFIIGLKPMPKPKGSLAVHAAEAKNFV
HbDAHPS 1 MSPTSTST---LIP---TKALITSNKPHQPSFVA--NKAPSRSLVQLISAVYSSERKNIIV

CoDAHPS 57 ADKPVKTSSEPPAAATAPAF-----GKNTLIGWSSKKALQLPEYFKREED
VvDAHPS 60 VDKPKTSSP--SVFAGS-----GKWTVDSWTKKALQLPEYFNESDLE
MtDAHPS 58 TEKPSKQPTTIPRNAST-----KWTIDSWSKKALQLPEYFQEDLE
HbDAHPS 53 SDMSGQQTTKTSAAATSAPAAAAAPTNNVVPKMSVDSWSKKALQLPEYFNKEE

CoDAHPS 101 AVLKLTLESFPIVFPAGEARHLEDRLEADAAMGKAFLLQGGDCAESFKEFNANNIRDTFRIL
VvDAHPS 102 SVLQTLLEAFPIVFPAGEARSLERLEGDAAMGNFLLQGGDCAESFKEFNANNIRDTFRIL
MtDAHPS 101 AVLKTLDAFPIVFPAGEARTLEEHLEGEAAMGNFLLQGGDCAESFKEFNANNIRDTFRIL
HbDAHPS 113 SVLRITLDAPPIVFPAGEARSLERLESEAMGNFLLQGGDCAESFKEFNANNIRDTFRIL

CoDAHPS 161 LQMGVVLMMFGGQMPVVKVGRMAGQPAKPRSDAFEEKNGVKLPSYRGDNINGDAFNEKSRI
VvDAHPS 162 LQMGAVLMMFGGQVPVVKVGRMAGQPAKPRSDPFEEKNGVKLPSYKGDNINGDAFNEKSRI
MtDAHPS 161 LQMSVVMFMFGGQMPVVKVGRMAGQPAKPRSDNPEEKNGVKLPSYRGDNINGDAFDEKSRV
HbDAHPS 173 LQMGAVLMMFGGQMPVVKVGRMAGQPAKPRSEPFEEKNGVKLPSYRGDNINGDTEKSRV

CoDAHPS 221 PDFERMIRAYCQSAATLNLLRAFATGGYAAQQRVNWNLDFTEHSEQGDYRELAAHRVDE
VvDAHPS 222 PDFORMIRAYCQAAATLNLLRAFATGGYAAQQRVTOWNLDFAEHSEQGDYRELAAHRVDE
MtDAHPS 221 PDFORMIRAYCQAAATLNLLRAFATGGYAAQQRVTOWNLDFTEHSEQGDYRELAAHRVDE
HbDAHPS 233 PDFORMIRAYCQSAATLNLLRAFATGGYAAQQRVNWNLDFTEHSEQGDYRELAAHRVDE

CoDAHPS 281 ALGFMAAAGLTEDHETINTDFWTSHECLLLPYEQATTEEDSTSGRYDCAHMLVWGER
VvDAHPS 282 ALGFMSAAGLTVDHPIMTTEFWTSHECLLLPYEQSLTRKDDSTGLYDCAHMLVWGER
MtDAHPS 281 ALGFMAAAGLTVDHPIMRTDFWTSHECLLLPYEQSLTRLDSTGLYDCAHMLVWGER
HbDAHPS 293 ALGFMSAAGLTVDHPIMTTEFWTSHECLLLPYEQSLARLDSTGLYDCAHMLVWGER

CoDAHPS 341 TRQLDGAHVEFLRGVANPLGIKVS DKMDPNELVKLIEILNPNKPGRITITIRMGAEANMR
VvDAHPS 342 TRQLDGAHVEFLRGVANPLGIKVS DKMDPNELVKLIEILNPNKPGRITITIRMGAEANMR
MtDAHPS 341 TRQLDGAHVEFLRGVANPLGIKVS DKMDPNELVKLIEILNPNKPGRITITIRMGAEANMR
HbDAHPS 353 TRQLDGAHVEFLRGVANPLGIKVS DKMDPNELVKLIEILNPNKPGRITITIRMGAEANMR

CoDAHPS 401 VKLPHLIRAVRRAGQIVTWVSDPMHGNTIKAPCGLKTRPFDSTLRAEVRAPFDVHQEGSH
VvDAHPS 402 VKLPHLIRAVRQAGQIVTWVCDPMHGNTIKAPCGLKTRPFDALRAEVRAPFDVHQEGSH
MtDAHPS 401 VKLPHLIRAVRRAGQIVTWVSDPMHGNTIKAPCGLKTRPFDALRAEVRAPFDVHQEGSH
HbDAHPS 413 VKLPHLIRAVRRAGQIVTWVSDPMHGNTIKAPCGLKTRPFDSTLRAEVRAPFDVHQEGSH

CoDAHPS 461 PGGVHLEMTGQNVTECIIGGSRVTVPDDLSSRYHTHCDPRLNASQSLAPLIIAERLRKR
VvDAHPS 462 PGGVHLEMTGQNVTECIIGGSRVTVPDDLSSRYHTHCDPRLNASQSLAPLIIAERLRKR
MtDAHPS 461 PGGVHLEMTGQNVTECIIGGSKIIVTDDLSSRYHTHCDPRLNASQSLAPLIIAERLRKR
HbDAHPS 473 PGGVHLEMTDQNVTECIIGGSRVTVPDDLSSRYHTHCDPRLNASQSLAPLIIAERLRKR

CoDAHPS 521 IGFQTASVALGL
VvDAHPS 522 MG--TQRLALGL
MtDAHPS 521 IR--SQPLESTGF
HbDAHPS 533 IN--SQPIASTSL

CoCS 1 --MASSSISSPFLKGTKEFPWGIGPHASDLRGLTYPSSVOISIRRAPFRKKLEVVHAAGS
RcCS 1 MASCSSSLVSKPPLGASRLNCS-----DNRKLSISTVRISFSFRAP--KKLQIHAAGS
HbCS 1 --MASSSLSSKPLGAYRLDGSS----SLNLRDLSTATVQISFRPRTP--KKLQIHAAGS
PtCS 1 --MASSTLTSSKPLGSSRIDGAS----ISSDLRQLSSISVQISFRSRIE--KKLQIHAAGS

CoCS 59 SFGNFFRITTYGESHGGGVGCIIDGCPPIPLSEADLQVLDLRRRPGQSRITTPRKETDT
RcCS 52 TFGTHFRVITTFGESHGGGVGCIIDGCPPIPLSEADLQVLDLRRRPGQSRITTPRKETDT
HbCS 54 TYGTYFRVITTFGESHGGGVGCIIDGCPPIPLSEADLQVLDLRRRPGQSRITTPRKETDT
PtCS 54 TFGTNFRVITTFGESHGGGVGCIIDGCPPIPLSEADLQVLDLRRRPGQSRITTPRKETDT

CoCS 119 CKIYSGVSEGLTGTPIHVSVPNTDQRGHDYSEMSIAYRPSHADATYDFKYGVRSVQGGG
RcCS 112 CKIYSGVSEGMTTGTPIHVFVPNTDQRGHDYSEMSVAYRPSHADATYDMKYGVRSVQGGG
HbCS 114 CKIYSGVSEGVTTGTPIHVFVPNTDQRGHDYSEMSVAYRPSHADATYDMKYGVRSVQGGG
PtCS 114 CKIYSGVSEGLTGTPIHVFVPNTDQRGLDYSEMSVAYRPSHADATYDMKYGVRSVQGGG

CoCS 179 RSSARETIQIRVIPGAVAKKILKMVSETEVLAYVSQVHKVILPEGVVDHNDVLTLEQIESNI
RcCS 172 RSSARETIQIRVASGAIKKILKQFSGTEVLAYVSQVHKVVLPEDMVDHETLLDQIESNI
HbCS 174 RSSARETIQIRVASGAIKKILRQFSGTEILAYVSQVHQQVLPEDQIDHQSLTIDQIESNI
PtCS 174 RSSARETIQIRVAAGVAKKILKLYAGTEILAYVSQVHKVVLPEGVVDHSDLLTDQIESNI

CoCS 239 VRCPPDPEYAEKMIASAIADAVRVRGDSVGGVVTICIVRNVPRGLGSPVFDKLEAELAKAVLSL
RcCS 232 VRCPPDPEYAEKMIAAIIRKIRVRGDSVGGVVTICIVRNVPRGLGSPVFDKLEAELAKAVMSL
HbCS 234 VRCPPDPEYAEKMIAAIITVRVRGDSVGGVVTICIVRNAFPRGLGSPVFDKLEAELAKAALS
PtCS 234 VRCPPDPEYAEKMIAAIADAVRVRGDSVGGVVTICIVRNAFPRGLGSPVFDKLEAELAKAAMS

CoCS 299 PATKGFPEISGGFGTFLTGSEHNDEFYTDKGRIRTRTNRSGGIQGGISNGETITMRFVAF
RcCS 292 PATKGFPEFGSGFAGTFLTGSEHNDEFYMEEHGKIRTRTNRSGGIQGGISNGETIYMRIAF
HbCS 294 PATKGFPEFGSGFAGTFLTGSEHNDEFYVDEHGKIRTRTNRSGGIQGGISNGETINRIAF
PtCS 294 PATKGFPEFGSGFAGTLLTGSEHNDEFYTDKHGKIRTRTNRSGGIQGGISNGEITNMRIAF

CoCS 359 KPTSTIISRKQHTVTRDNQIEILLARGRHDPVVPRAVPMVEAMVALVLDQLMAHYAQCY
RcCS 352 KPTSTIIGKQHTVTRDNVETELIARGRHDPVVPRAVPMVEAMVALVLDQLMAQYAOQY
HbCS 354 KPTSTIIGKQHTVTRDKKETELIARGRHDPVVPRAVPMVEAMVALVMDQLMAQYAOQY
PtCS 354 KPTSTIIGKQHTVTRDKKETELIARGRHDPVVPRAVPMVEAMVALVMDQLMAQYSQCY

CoCS 419 LLFINPSELQEPLEAF--KLEAAHVPL
RcCS 412 MFPINPELQEPKLLP--RAESVNMTI
HbCS 414 MFPVNPSELQEPKLLPSPSIEATNMST
PtCS 414 LLFINSELQEPHIME--RLEAANASV

```

Supplementary Fig. S6 (Continued). (C) Alignment of deduced amino acid sequences of two representative genes, DAHPS and CS, involved in shikimate pathway. DAHPS, 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; CS, chorismate synthase. Co, *Cananga odorata*; Hb, *Hevea brasiliensis*; Mt, *Medicago truncatula*; Pt, *Populus trichocarpa*; Rc, *Ricinus communis*; Vv, *Vitis vinifera*. Accession numbers : VvDAHPS (NP_001268127), MtDAHPS (XP_003615152), HbDAHPS (AFY09700), RcCS (XP_002529571), HbCS (ADR70879), PtCS (XP_002315301). Completely conserved residues are shaded in dark gray, and identical residues are shaded in light gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

D

COPAL 1 MEFPCQEAQNNGMLGCLIA-----DPLNWGVVREALKGSHLDEVKMVEEPRKFSIK
 PtPAL 1 MEFPCQDSRNNGSLGLEN-----TNDPLNWGMAAESLKGSHLDEVRKRIEYRKRKPVVK
 MbPAL 1 MEFAPKAGVVEGEAEC-----LKADPLNWIKAASSTLKGSHLDEVRKRVVEPRKFLVR
 RCoPAL 1 MEFSEDDNNNNHSSLESFCTATAGHDFPLNNGMLADSLKGSHLDEVRKRVYDTRRFFVVR

 COPAL 51 LGGALLTIAQVAIVYH-EAGVKVLSLSARAGVKASDQWVMEEMNKGTDSYGVVTFGGAT
 PtPAL 53 LGGFTLTIAGQVTAISERDVGVMVLSSEARAGVKASDQWVMDSMSKGTDSYGVVTFGGAT
 MbPAL 54 LGGFTLTIAGQVAIVAAARSPPRVLSSEARAGVKASDQWVMEEMNKGTDSYGVVTFGGAT
 RCoPAL 61 LGGFTLTIAGQVTAISNRDAGIKVLSSEARAGVKASDQWVMDSMSKGTDSYGVVTFGGAT

 COPAL 110 SHRRTKQGGALQKELIRFLNAGIFTPRESGNTLFPATRAAMLVRIINTLLOGYSGRIFE
 PtPAL 113 SHRRTKQGGELQKELIRFLNAGIFGNGTSSHTLRSATRAAMLVRIINTLLOGYSGRIFE
 MbPAL 114 SHRRTKQGGALQKELIRFLNAGIFGSPRESGNTLHSSAKAAMLVRIINTLLOGYSGRIFE
 RCoPAL 121 SHRRTKQGGALQKELIRFLNAGIFGNGTSSCHTLHSAATRAAMLVRIINTLLOGYSGRIFE

 COPAL 170 ILEAITSPLNHHITFCPLRGTITASGDVLVPLSYIAGLLTGRFNSKARTAEKEMDAARA
 PtPAL 173 ILEAITKLLNHNITFCPLRGTITASGDVLVPLSYIAGLLTGRFNSKAVGFNGEFLSPARA
 MbPAL 174 ILEAIASLLNNGITFCPLRGTITASGDVLVPLSYIAGLLTGRFNAAVGFNGKVI GAARA
 RCoPAL 181 ILEAITKPIHNHVTFSLPLRGTITASGDVLVPLSYIAGLLTGRFNSTCVGFNGEFLPTARA

 COPAL 230 FRAAGIESPFPELOPKEGLALVNGTAVGSGFASLVLPFANVLAVLSEVMSIACEVMOGK
 PtPAL 233 FTQAGIDGGPFPELOPKEGLALVNGTAVGSGGLASMLVPEINVLAILSEVLSIAFAEVMOGK
 MbPAL 234 FRLASADGPFPELOPKEGLALVNGTAVGSGGLASMLVPEANVLAVLSEVLSAVFAEVMOGK
 RCoPAL 241 FRLASQIEGPFPELOPKEGLALVNGTAVGSGGLASMLVPEANVIGLSEVLSAVFAEVMOGK

 COPAL 290 PFFTDLHTRKLNKHPGQIEAAAI MEHILDGSSYVMEKXKLEHLDPLQKPKQDRYALRTSP
 PtPAL 293 PFFTDLHTRKLNKHPGQIEAAAI MEHILDGSSYVMEKXKLEHLDPLQKPKQDRYALRTSP
 MbPAL 294 PFFTDLHTRKLNKHPGQIEAAAI MEHILDGSSYVMEKXKLEHLDPLQKPKQDRYALRTSP
 RCoPAL 301 PFFTDLHTRKLNKHPGQIEAAAI MEHILDGSSYVMEKXKLEHLDPLQKPKQDRYALRTSP

 COPAL 350 OWLGLLIEVIRTSTLSIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSMNDTRLAVA
 PtPAL 353 OWLGLLIEVIRTSTKMIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSMNDTRLATA
 MbPAL 354 OWLGLQIEVIRASTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSMNDTRLAVA
 RCoPAL 361 OWLGLQIEVIRASAKMIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGVSMNDTRLATA

 COPAL 410 STGKMLFAQPSSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAETAMAAYCSSELQPLANPVT
 PtPAL 413 STGKMLFAQPSSELVNDFYNNGLPSNLTGGRNPSLDYGFKGAETAMAAYCSSELQPLANPVT
 MbPAL 414 ATGKMLFAQPSSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAETAMAAYCSSELQPLANPVT
 RCoPAL 421 ATGKMLFAQPSSELVNDFYNNGLPSNLTGGRNPSLDYGFKGAETAMAAYCSSELQPLANPVT

 COPAL 470 NHVQSAEQHNQDVSLSGLISARKTAAVAEILKLMSTFLVALCQAVDLRHLLENLSTVK
 PtPAL 473 NHVQSAEQHNQDVSLSGLISARKTAAVAEILKLMSTFLVALCQAVDLRHLLENLSTVK
 MbPAL 474 NHVQSAEQHNQDVSLSGLISARKTAAVAEILKLMSTFLVALCQAVDLRHLLENLSTVK
 RCoPAL 481 NHVQSAEQHNQDVSLSGLISARKTAAVAEILKLMSTFLVALCQAVDLRHLLENLSTVK

 COPAL 530 NTVSQAVKRVLTMTNGELHPSRFCEKDLKVVVDREHVFAVYDDPCSATYPLMKQLRQVL
 PtPAL 533 NTVSQAVKRVLTMTNGELHPSRFCEKDLKVVVDREHVFAVYDDPCSATYPLMKQLRQVL
 MbPAL 534 STVSQAVKRVLTMTGANGELHPSRFCEKDLKVVVDREHVFAVYDDPCSATYPLMKQLRQVL
 RCoPAL 541 NTVSQAVKRVLTMTNGELHPSRFCEKDLKVVVDREHVFAVYDDPCSATYPLMKQLRQVL

 COPAL 590 VDAHLSNGEKEDASTSIFPKIGAPFEDLKTQLPKEVENARADPESGNLAIPNRKECRS
 PtPAL 593 VDAHLSNGERERNSTSIPOKIGSFEELKTLTPKEVESARLEVENGFAPNRIKECRS
 MbPAL 594 VDAHLENGDKEKADGSSIFPKIATFESEHTAQLPKEVEAARAAVRCGKAIPNRKECRS
 RCoPAL 601 VDAHLSNGEKENSTSIPOKIGAPFEEKTLTPKEVENARTYDNGNFAIPNRKECRS

 COPAL 650 YPLYLRLVREELDGLLTKGKVRSPGEPFDKVFDAICQGVVDLPLECKLKNWGNAPLPIG
 PtPAL 653 YPLYLRFVREELDGLLTKGKVRSPGEPFDKVFDAICQGVVDLPLECKLKNWGNAPLPIG
 MbPAL 654 YPLYLRLVREELDGLLTKGKVRSPGEPFDKVFDAICQGVVDLPLECKLKNWGNAPLPIG
 RCoPAL 661 YPLYLRFVREELDGLLTKGKVRSPGEPFDKVFDAICQGVVDLPLECKLKNWGNAPLPIG

 CoC4H 1 MDLLELQNTFALFESIVVATVYVKLRGKRFKLPDGLPIFVFGNWLQVGDLDNHRNLT
 GmC4H 1 MDLLELKTILGLFLAAVVAIAVTLRGRKFKLPDGLPIFVFGNWLQVGDLDNHRNLT
 TcC4H 1 MDLLELLEALVSLFIVVILAILISKLRKRFRLPDGLPIFVFGNWLQVGDLDNHRNLT
 GaC4H 1 MDLLELREKLVLSLFTTIFAILVSKLRGKRFKLPDGLPIFVFGNWLQVGDLDNHRNLT

 CoC4H 61 LAKKPGDIFLLRMGQRNLVVVSSPELAKEVLHTQGVFPGSRTRNVVDPDIFTGKGQDMVPT
 GmC4H 61 LAKKPGDIFLLRMGQRNLVVVSSPELAKEVLHTQGVFPGSRTRNVVDPDIFTGKGQDMVPT
 TcC4H 61 LAKKPGDIFLLRMGQRNLVVVSSPELAKEVLHTQGVFPGSRTRNVVDPDIFTGKGQDMVPT
 GaC4H 61 LAKKPGDIFLLRMGQRNLVVVSSPELAKEVLHTQGVFPGSRTRNVVDPDIFTGKGQDMVPT

 CoC4H 121 VYGDHWRKMRRIIMTVPPFTNKVVQQYRHGWEDSARAVVEDVKKNPDAATNGIHLRRLQL
 GmC4H 121 VYGEHWRKMRRIIMTVPPFTNKVVQQYRHGWEDSARAVVEDVKKNPDAAVSTIRRLRQL
 TcC4H 121 VYGEHWRKMRRIIMTVPPFTNKVVQQYRHGWEDSARAVVEDVVRKNPEATNGIHLRRLQL
 GaC4H 121 VYGEHWRKMRRIIMTVPPFTNKVVQQYRHGWEDSARAVVEDVKKNPDAATNGIHLRRLQL

 CoC4H 181 MMYNNMYRIMPDRRFESEDDPLFVKLKALNGERSRLAQSFPEYNYGDFPILRFLRGLYK
 GmC4H 181 MMYNNMYRIMPDRRFESEDDPLFVKLKALNGERSRLAQSFPEYNYGDFPILRFLRGLYK
 TcC4H 181 MMYNNMYRIMPDRRFESEDDPLFVKLKALNGERSRLAQSFPEYNYGDFPILRFLRGLYK
 GaC4H 181 MMYNNMYRIMPDRRFESEDDPLFVKLKALNGERSRLAQSFPEYNYGDFPILRFLRGLYK

 CoC4H 241 ICKEVKERLQLFKDYFLERKKLASTKSSASGACAIHDILDAQKKEGINEEDNVLYI
 GmC4H 241 ICKEVKERLQLFKDYFLERKKLASTKSSASGACAIHDILDAQKKEGINEEDNVLYI
 TcC4H 241 ICKEVKERLQLFKDYFLERKNLASHITSDNN-ALKCAIDHILDAQKKEGINEEDNVLYI
 GaC4H 241 ICKEVKERLQLFKDYFLERKKLASTKRIEDNN-ALKCAIDHILDAQKKEGINEEDNVLYI

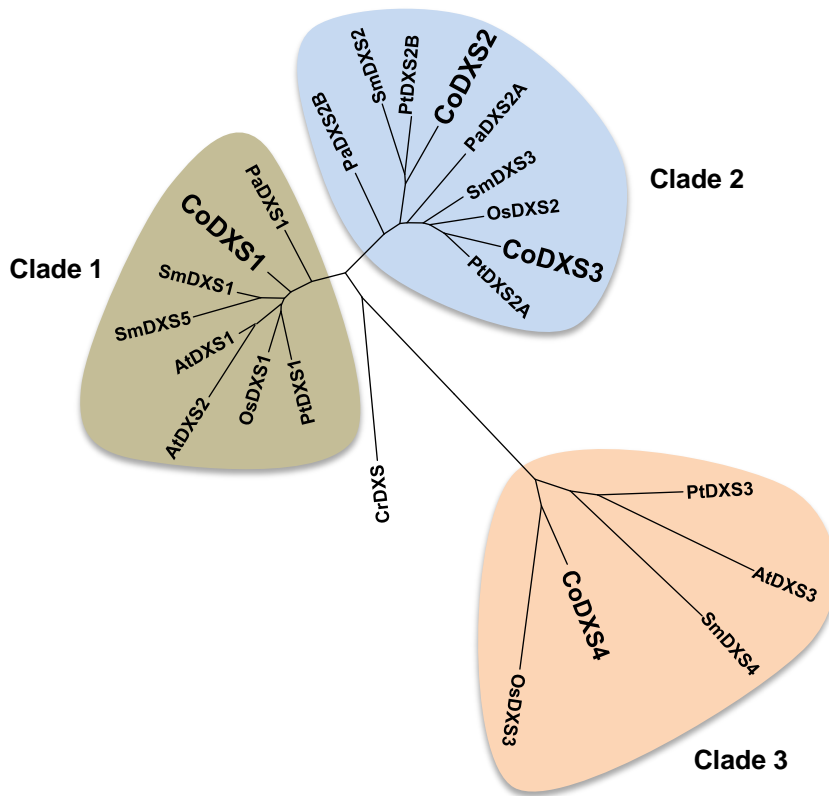
 CoC4H 301 VENINVAAIETTLLWSIEWGIAELVNHPEIQOKLRNEIDTVLGPVQVTEPDTHKLPYLQA
 GmC4H 301 VENINVAAIETTLLWSIEWGIAELVNHPEIQOKLRNEIDTVLGPVQVTEPDTHKLPYLQA
 TcC4H 300 VENINVAAIETTLLWSIEWGIAELVNHPEIQOKLRNEIDTVLGPVQVTEPDTHKLPYLQA
 GaC4H 300 VENINVAAIETTLLWSIEWGIAELVNHPEIQOKLRNEIDTVLGPVQVTEPDTHKLPYLQA

 CoC4H 361 VIKETLRLRMAIPLLPHMNLHDAKLGGYDIPAESKILVNAWLANNPAAHWKPEFRPE
 GmC4H 361 VIKETLRLRMAIPLLPHMNLHDAKLGGYDIPAESRILVNAWLANNPAAHWKPEFRPE
 TcC4H 360 VIKETLRLRMAIPLLPHMNLHDAKLGGYDIPAESKVLVNAWLANNPAAHWKPEFRPE
 GaC4H 360 VIKETLRLRMAIPLLPHMNLHDAKLGGYDIPAESKILVNAWLANNPAAHWKPEFRPE

 CoC4H 421 RPFEEESKVEANGNDPRYLPFGVGRSSCPGIIALPILGITLGRLVQNFELLPFGQSKL
 GmC4H 421 RPFEEESKVEANGNDPRYLPFGVGRSSCPGIIALPILGITLGRLVQNFELLPFGQSKL
 TcC4H 420 RPFEEESKVEANGNDPRYLPFGVGRSSCPGIIALPILGITLGRLVQNFELLPFGQSKL
 GaC4H 420 RPFEEESKVEANGNDPRYLPFGVGRSSCPGIIALPILGITLGRLVQNFELLPFGQSKL

 CoC4H 481 DTSEKGGQPSLHLKHSITIVKPRVAF
 GmC4H 481 DTSEKGGQPSLHLKHSITIVKPRVAF
 TcC4H 480 DTSEKGGQPSLHLKHSITIVKPRVAF
 GaC4H 480 DTSEKGGQPSLHLKHSITIVKPRVAF

Supplementary Fig. S6 (Continued). (D) Alignment of deduced amino acid sequences of two representative genes, PAL and C4H, involved in phenyl propanoid pathway. PAL, phenylalanine ammonia lyase; C4H, cinnamate-4-hydroxylase. Co, *Cananga odorata*; Ga, *Gossypium arboreum*; Gm, *Glycine max*; Mb, *Musa balbisiana*; Pt, *Populus trichocarpa*; Rc, *Ricinus communis*; Tc, *Theobroma cacao*. Accession numbers : PtPAL (ACC63889), MbPAL (BAG70992), RcPAL (XP_002531677), GmC4H (ACR44227), TcC4H (XP_007011365), GaC4H (AAG10197). Completely conserved residues are shaded in dark gray, and identical residues are shaded in light gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.



Supplementary Fig. S7. Phylogenetic analysis of DXSs from dwarf ylang ylang. The maximum likelihood tree was drawn by MEGA 6 program from an alignment of full-length CoDXSs (1-deoxy-D-xylulose 5- phosphate synthase) with other plant DXSs. At, *Arabidopsis thaliana*; Os, *Oryza sativa*; Pa, *Picea abies*; Pt, *Populus trichocarpa*; Sm, *Salvia miltiorrhiza*.

```

CoDXS1 1 ----MLMALAAFFHPCHFIG-----SAASILQKSTPHCSYLLGSEADLQLSYHHKPPK
CoDXS2 1 MAASAVFMHSSSLPLSDFTQEHSIRKLRSSISRPAAARKNTFFKYIAACSKDSDSYSTSAIS
CoDXS3 1 -----MLAKLPALCPFS-----PLPRPAATCKR-----QWCVKASASDDEGRLA
CoDXS4 1 -----MQVSALEPDL-----

CoDXS1 49 GKKVPWICASLSEETGEYYSQRPETPLLDLTINYPIHMKNLSTKELTQLADELRSDFVFNVS
CoDXS2 61 VNKDVPESQENSLSLNFTGKPKETPILNTVNYPIHMKNLTIKELARLADELREBIIVTVS
CoDXS3 40 IRK-----EKGGWKIDFTEEKPATPLLDLTINYPIHMKNLRVHDLEQLAAEIRADIVHTVS
CoDXS4 11 -----DFWEKAPTPLVLDVMVKNFMHLKNLSSKELKQLAAEIRSEISFSMS

CoDXS1 109 KTGGHLGSSLGVVELTVALHYVFNAPQDKILWDVGHQAYPHKILTGRFRDKMPTIRQTNGL
CoDXS2 121 KTGGHLSSSLGVARELTVALHHVFNTPEDKIVWDVGHQAYAHKILTGRRRSRMHTIRQTSGL
CoDXS3 95 KTGGHLSASLGVVELSIALHHVFNAPDDKIIWDVGHQAYPHKILTGRRRSKMHTIRKTSGL
CoDXS4 56 RTRRRPFKASLGVVELSIALIHYVFNAPMDKILWDVGHQAYAHKILTGRRRSLMHTLRQKNGL

CoDXS1 169 SGFTKRAESEYDCFGAGHSSTSSISAALGMAVGRDLKGRKNHVVAVIDGDGAMTAGQAYEAM
CoDXS2 181 AGFPKRDESKHDAFGVGHSSSTSSISAGLGMAGIRDLLRKNHVVAVIDGDGAMTAGQAYEAM
CoDXS3 155 AGFPKRDESIDAFGAGHSSTSSISAGLGMAVARDLLGKKNHVISVIGDGAMTAGQAYEAM
CoDXS4 116 SGFTSRFSEYDAFGAGHCNSISISAGLGMAVARDLKGEKERIVTVISNGTTMAGQAYEAM

CoDXS1 229 NNAGYLDNSNMIVILNDNKQVSLPTATLDGPPVPGALSRAISRLQSSRPLRELRVAKGI
CoDXS2 241 NNAGYLDNSNLIVLNDNKQVSLPTATIDGPAPPVPGALS KALTRLQSSRKLRLQREVAKI
CoDXS3 215 NNSGYLDNSNLIVILNDNKQVSLPTATLDGPATPVGALS SALSNLQASTEFKLRAAKSI
CoDXS4 176 SNAGFLDNSMVVILNDSRHCLHPKLD-EGSKMSISALSSTLSKIQSSKSPRQLREAAKVV

CoDXS1 289 TKQIGGPMHLEAAKVDEYARGMISGSGSTLFEELGLYYIGPVDGHNIDDLVSIKVEVKNIT
CoDXS2 301 TKSIGGQTHEIAAKVDEYARGLMGAPGATLFEELGLYYIGPVDGHNVEELVHIFEKVKAT
CoDXS3 275 TKQIGGEAHEAAKMDYARGMISPSKSCLFEEELGLYYIGPVDGHNMEGLLITLQKVKAM
CoDXS4 235 TKRIGRGMHLEAAKVDEFARGMMGPLGSTLFEELGLYYIGPVNNGHNIDDLICVLQEVASL

CoDXS1 349 KITGPVLVHVVTEKGRGYPYAERASDKYHGVTKFDPATGKQFKGSSPTQSYTMYFAEALI
CoDXS2 361 PATGPVLIIHII TEKKGKGYPPAEAAAADKMHGVVVFDPKTKGQTKVKAPTLSYTOYFAEGLI
CoDXS3 335 PAPGPVLIVHVTEKKGKGYPPAEAAAADKMHGVVVFDPATGKQFKPQSSTLSYTOYFAESLI
CoDXS4 295 DSSGPVLIVHVI TEDEG---SEEDQKSRLVCKHQGLTASYSRVMSSSLPRTYNDCEVVALV

CoDXS1 409 AEAEADKDIVGIHAAAMGGGTGMNLFLLRRFFTRCFDVGIAEQHAVTFAAGLACEGLKPFCA
CoDXS2 421 AEAKQDEKIVAIIHAAAMGGGTGLNVFQKQFFPERCFDVGIAEQHAVTFAAGLAAEGLKPFCA
CoDXS3 395 KEAEVDDKIVAIIHAAAMGGGTGLNVFQKQFFPERCFDVGIAEQHAVTFAAGLATEGLKPFCA
CoDXS4 352 AGAERDKDIVVHAGMGMDPSLELEFQETFPDNF GIGMAEQHAVTFAAGLSGGLKPFQV

CoDXS1 469 IYSSFLQRAYDQVIHDVDLQKLPVRFAMDRAGLVGADGPTHSGSFDVTYMACLPNMVVMA
CoDXS2 481 IYSSFLQRGYDQVVHDVDLQKLPVRFAMDRAGLVGADGPTHCGAFDITAYMACLPNMVVMA
CoDXS3 455 IYSSFLQRGYDQVVHDVDLQKLPVRFALDRAGLVGADGPTHCGAFDITYMACLPNMVVMA
CoDXS4 412 IPSTFLQRAYDQIVQDVDLQKIPVRFALDRAGLVGADGPTHCGAFDITFMSCLPNKICMA

CoDXS1 529 PSDEAELPHMVATAAAIINDRPSCFRYPFRNGIGVPLPFGNKGVPLEIGKGRVLIGGERVA
CoDXS2 541 PSDETELIIHMIATAAAIDDRPSCFRYPFRNGVGSVLEPDYKGTPLEIGKGRVLIGGERVA
CoDXS3 515 PSDEAELMHMVATAAATIDDQPCCFRFPFRNGVGVALE-----
CoDXS4 472 PADEDELVHVMVATAACINDRPPVCFRFPFRGAIIVGMNIE-LHSGLPLEIGKGRILAVQKDAE

CoDXS1 589 LLGYGTAVQSCLAASLVGQQGLQITVADARFCKPLDQDLIRSLAKSHEVLITVEEGSIG
CoDXS2 601 ILFGFTIVQNCMLAQQMLREMGSVATVADARFCKPLDGE LIRLANEHELIIITVEASIG
CoDXS3 552 -----LIARAKNRPKPAFLVANS SH-
CoDXS4 531 LLGYGIMVQNCILKARSLLANPGIHVTVADARFCKPLDIELVRKLCQEHFLLITVEEGTIG

CoDXS1 649 GFGSHVAQFMSLDGLLDGTTKWRPLILPDRYIEHGSFVDQMVEAGLMPSHVAATVFNVLG
CoDXS2 661 GFGSHVSHFLALNGLLDGKVKWRPMTLPDRYIDHGS PKDQIEEAGLTAKHIAATVLSLLG
CoDXS3 -----
CoDXS4 591 GFGSHVSHFISLDGQLDENVKWRPIVLPDNYIEQASPKEQLGLAGLTGHHIAATALNLLG

CoDXS1 709 KTREALNIMS
CoDXS2 721 ENKEAFNLLT
CoDXS3 -----
CoDXS4 651 RTRDALLLMR

```

Supplementary Fig. S8. Comparison of deduced amino acid sequences of four CoDXS small gene family. The thiamin diphosphate-binding site and the pyridine binding DRAG Domain are indicated by the open box and the horizontal line, respectively. Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

A

```

CoTPS1 1 ---MALNNTFLHFPDPCSLSSFSQCAVPKLPLAIFHK--TMAEQIRCPRASSQT-----SEP
Mg17 1 ---MALKLLFQCSPSPSSLAPLQPVLVLRPPSGAKARRNLRG-CASTQV-----TEL
SoSS 1 ----MSSISINIAMPLNSLHNFRKPKAWSTSCAPARLRA-SSSLQQ-----EKP
PsSS 1 MSVISIVPLASNSCLYKSLMSSTHELKALCRFIATLGMCRRGKSVMASMSTSLTTAVSDD

CoTPS1 50 ALARRSANFQETIWTNDFIQSLNSDYSSDVYVQRIEKLKKSVRQSLLEA-----
Mg17 51 MTARRSANYHPNIWDYDSVQSLTSDYKAYTYLERVEKLKEDVVRTLQEA-----
SoSS 49 HQIRRSQDYQPSLWDFNYIQSLNTPYKEQRHFNRQAE LIMQVRMLLKVK-----
PsSS 61 GVQRRIIGHHSNLWDDNFIQSLSSPYGASSYAESA KKLIGE VKEIFNLSLSMAAGGLMSPV

CoTPS1 99 DGPLAQLELIDDLQRLQVGRLEFEREINEMENGIYMDYKET-QAQW-----NLHFTSMYF
Mg17 100 VGLLDQLELVDCIHRLLGVGYHFDKRIKEIKTISTEPNNMGLIDG-----DLYAMALYF
SoSS 98 MEAIQQLELIDDLQYLGLSYFFQDEIKQILSSIHNEPR--YFHN-----DLYFTALGF
PsSS 121 DDLQHLQSMVENVERLGDIDRHFOETRIKVSLELDYVYSYWSKGGIGSGRDIVCTDLNNTALGF

CoTPS1 152 RLLRAREFDVSPERFSRF---MDEGTGNFQTSISNDPIGMLSLEYEASYL CMPGETTLDEAQ
Mg17 154 RLLRQHGVEVPQGVNRF---MDDSSSFKASLNCNDVKGMLSLEYEASYLALGETTLDEAK
SoSS 150 RILRQHGQFNVEDVDFCFK--IEKCSDFNANLAQDTKGMQLQLEYEASFLLRGEGEDTLELAR
PsSS 181 RILRLHGYTVFPDVFHEHFKDQMGRACSANHTERQISSINLFRASLIAFPGEKVMEEAE

CoTPS1 209 AFTCKHLKYWK---EKDVHPTIALQIEHALELPIHWRMPRLDSRWYIKLYEE-----KE
Mg17 211 AFTYRHLRGLK---G-NIDSNLKGLEHALELPLHWRVLRLEARWYIDTYER-----ME
SoSS 208 RFSSTRSLREKFDGEGDEIDEDLSSWIRHSLDLPLHWRVQGLEARWFLDAYAR-----RP
PsSS 241 IFSATYLKEAL---QTIPVSSLSQEMQYVLDYRWHSNLPRLETHTYIDILGETTINQMQD

CoTPS1 260 GTRPLLELAKLDFNMVQSAHQTELKRVSRWSEFGLAEKASFARDRLMEGYQWAIGTVF
Mg17 261 DMNPLLELAKLDFNIVQNVYQGGVVRKMSGWWKDLGLGQKLGFAARDRLMEGFLWTIGVKF
SoSS 262 DMNPLIFKLAKLNFNIVQATYQBELKDISRWNNSSCLAEKLPFVRDRIVECFFWAIAAFE
PsSS 298 VNIQKLELAKLEFNIFHSIQNELKCSIRWVKESGSPE-LTEIRHHEIFYTLAGSIDM

CoTPS1 320 EPEFGQCREVLAQLIAVIDDMYDVYGSPELELELFTDAVDRWNINTIEGLPDYMKLCF
Mg17 321 EPQFAQCREVLTAKINQLITITIDDVYDVYGSLEEELELFTKAVDRWDTNAMEELPEYMKICF
SoSS 322 PHQYSYQKMAAVIITFTITIIDVYDVYGTIELELELTD MIRRWDNKSIQLPYMQVCY
PsSS 357 EPKHSAPRSLSFVKMCHLITVLDIYDTFGTMDLRLFTSAVWRWRDSEIECLPEYMKGVY

CoTPS1 380 LSIYNTNQGGEYFLKDHGVDIIPHLRKAWADYCKALRTEARVNSKYTPTLDEYLNNA Y
Mg17 381 LALYNTVNEIAYDTLKEQGVDVIPYLRQSWADLCKAYLVEARWYVSGYTPTLDEYLNNAW
SoSS 382 LALYNFVNERAYDILKQHFNSIPYLQKSWVSLVEGYLKEAYWYNGYKPSLEEYLNNAK
PsSS 417 IILYETVNERAREARKDQGRDTLNYARLALEDYIGAYLKEAEWISYMLPTFEEYFKNGK

CoTPS1 440 TSASGPLILIHAFVFSGQEPWKEAIDCFVSSNKDIIRLSATIFRI TDDLETSABEIERGD
Mg17 441 ISIAGPVILVHAYVSMIQMITKEALLDCVGSYESIMQWSSMILRLADDLATSTDELERGD
SoSS 442 ISISAPTIISQLYFTLANSIDETAIESLY-QYHNILYLSGTILRLADDLGTSQHELERGD
PsSS 477 VSSGHRATLQPIILTDIPFPFHILQEI DFP-SKFNELACSLRLRGLTRCYQADDRGE

CoTPS1 500 VPKSIQCYMHEA-GASEAVSRAHIRGKISEVWRKMKNKYLTAPAT-RHKTFNAAAFNLART
Mg17 501 VPKSIQCYMHEN-TASEVVAEQMRARISDIWKKMKNKDVALSP--LEQPFKAAAVNLARM
SoSS 501 VPKAIQCYMNDT-NASEREAVEHVKFLIREAWKEMNTVTASDCPFTDDLVA AAAANLARA
PsSS 536 KASCISQYMKDNPGSTEDALNHINGMIEDTIKQLWELLRPNNVVIESSKKHHSFDISRA

CoTPS1 558 STCVYLYGDGYGVPNGKNKENITSLTVEPIVLE--
Mg17 558 AQCMYQHGDGHGNPHRESKDHLISLVVEPIQLMES
SoSS 560 AQFIYLDGDGHGVQHS EIHQQMGGLLFQPYV----
PsSS 596 FHHLRVRDGYTVSSNETKNLVVRTVLEPLPM---

```

Supplementary Fig. S9. Alignment of deduced amino acid sequences of CoTPSs and other plant TPSs. (A) Alignment of deduced amino acid sequences of CoTPS1 and other plant TPSs. PsSS, *Picea sitchensis* sabinene synthase (ADU85930); SoSS, *Salvia officinalis* sabinene synthase (AAC26018); Mg17, *Magnolia grandiflora* α -terpineol synthase (B3TPQ7). The conserved motifs DDXXD and NSE/DTE are marked. Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

B

```

CoTPS2      1 MALIFANGHSDVVESTQPPIGKQKKEIGRESVNYHPSVWGDRFAILTAHEIEVDEITKQRA
RcSeTPS1    1 MSLQVSAVPIKTSTQN-----ATSAVKRHSSTYHPTINGDHF LANLSHSKIIDGSIEQQF
Mg25        1 -----MDSETTQ----RPNMEIGRAFVNYHPSIWGEHFIAASPDVMRLDAH-KGRG
HaCS        1 MATTEANTMAQANSQT-----TIEPVRHHLANFPSPSINGDQELSFSLDNSQLEAYSK-AM

CoTPS2      61 EKLKHDV LKMLHNVS-VSLQDLNLIDIEIQRLGVGYHFEETIENAMKRIYNSRDNDDDD--
RcSeTPS1    56 EGLKQKVRKMIIDLNNEPCKKLG LIDAVQRLGVGYHFKSEIEDVLQKVYHDYSDEDD--
Mg25        47 EELKEVVRNMFSTVN-DPLLKMNLDIAIQRLGVAYHFEEMIDKALGQMYDDHIN GKDDGF
HaCS        54 EQPKENVRRMILNPAIDTNEKLG LLYCVYRLGLTYNFSKDI DGQLDELFKQLNLQSYNEA

CoTPS2      118 -LHAVALRFRLLRQHGYNVS SDVFKPKDEK-GEFKASIRENVRGLLSFYEAAYLG TADD
RcSeTPS1    114 -LNTVALRFRLLRQHGIKVS CAIFKPKDSE-GNFKTS LINDALGMLSLEYEATHLSIRGE
Mg25        106 DLQTLALQFRLLRQGYNVS SGVFAKPKDE-GNFSSILSKDTHGLLSLEYEAALG THGD
HaCS        114 DLYTISIHQVFRHFGRYRFS CDVENKPKDSSSGKFKEDMTRDVRGMISLYESAQLIRIGE

CoTPS2      176 TILDQAI DFTTDQLKSVLPNLPPISELVKLALDVP LHKRIERLQSFYFISIQE EKERN
RcSeTPS1    172 DVLDEALAF TTTNLQSVLPQLNTHLAAQISRALNRPIRKYLPRL EARNYIDYATEESYN
Mg25        165 DILDEAIF TTVHLKSTLPHVSAPITKLVLELALEIPLHRRMERLQTRFYIS IYEDRERN
HaCS        174 SILDEAGAF AESKIKTIEKTL DGTLAQQVKHVLERPFNRGHQMVEAR KYLFLPEEIEISRY

CoTPS2      236 EVLLEFAKLD FNVLQSLHKEELS QLSRWWKDNDFA RKLFPFIRDRLVECYFWILGVY YEPF
RcSeTPS1    232 TLLNFAKLD FNMLQELHQKELNVVTKWKS LDDVATKLPYARDRVVECYFWMVGVY FEPQ
Mg25        225 DVLLFEFS KLEFLRLQSLHQRELRDISLWKKEMD LLALEPFRDRVLEGEYFWTVGVY FEPH
HaCS        234 DSSLMLAKV HFNVLQLLQKEELRSVSKWKKDL DLPATLYVVRDRVPELVVWILAFFLEPY

CoTPS2      296 YSRGRMMT TRVISLTSIMDDTYD VYGLDELELLT TATERWEWAAMDEL PDMKHLHFSAL
RcSeTPS1    292 YSFARIMM TKIIAITSLLDDTYD NYATGEELEIL TEATERWDIKAKDALPEYMKII YTTL
Mg25        285 YSRARMIM TRMIAFATVMDDTYD VYGTLEELELLT TATERWNRGDMDQLPDMKVI FIAL
HaCS        294 YSEVRIIT TRKIVLLVLVLD DTYDANATIEESR LLTHAINRWEVSAMLQ LPEYMKPLYEIL

CoTPS2      356 LTAVENFE EELSKEGKA--YRISYFNAYTKLAKAYLE EARWASADYVPTLE EYMKHAQV
RcSeTPS1    352 LDIYNEYE ENIAKEBEKSLLSYVYAK EVMKRVVRAYLEVRWRDNCYTP TMEEYMQSALL
Mg25        345 LDGVDAT EDDLTGEGKS--YRIYLLK EAVKDLAKAYLEEARWVSSGYVPTSE EYMKVALI
HaCS        354 LNEYDGA ---FYKHGRT--NVIETSKKAFQDLARS YHQSEWRHAKAVEP SFREYMKIGTT

CoTPS2      414 SSAYPVV TLSLLGMGATA TKEAFEWAINMPNAINAIS VVCR LKDDITSAEL EQQRVHVA
RcSeTPS1    412 TTCSPMLA IASFLGLKEIATKEAYEWASEDPK IIRASSIVCRLMDDIVSHEFEQTRKHVA
Mg25        403 SAVYPM L FVAFLIGMDEVVTKVLEWAIHMP TMLR TCSIVARLMDDIPS NKLEQERKHVS
HaCS        408 TSAHNV LSKTALIGMGNIVTREALAWYESY PKIVQLSELIGRLEDDVVS VEFERERAPTA

CoTPS2      474 TAVECYIK ENGTTYEETCKL FQKADDAWKELNKEWMEAIQ VPKEIFKHVVNLARVIE FL
RcSeTPS1    472 SGVECYIK QYQASEEEV IKLFRKREVTN AWKDLNEECLNPTVPMPMLERV VNLTRAIDVI
Mg25        463 SSVECYMK EHGTSYHESI QKLRMVASGWK DINK ECLKPTVPTAVINVLNFRVLE I I
HaCS        468 TSVDA YMKTYQVSENVAVKILK KLVENGW KDLNEACLKPTESLDLLAPI IGLTNMTDVA

CoTPS2      534 YRYKDMYTESDGETKECITMLLLDFV V
RcSeTPS1    532 YKDDDG YTN-SHIMKDYVASVLKDFVPV
Mg25        523 YQHRDGY TDASVETKEHIASL FVDFIPL
HaCS        528 VRHNDGLTFPEKTLKEYITLLFCVQVPM

```

Supplementary Fig. S9 (Continued). (B) Alignment of deduced amino acid sequences of CoTPS2 and other plant TPSs. RcSeTPS1, *Ricinus communis* α -copaene synthase (B9S9Z3); Mg25, *Magnolia grandiflora* β -cubebene synthase (B3TPQ6); HaCS, *Helianthus annuus* α -copaene synthase (Q4U3F6). The conserved motifs DDXD and NSE/DTE are marked. Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

C

```

CoTPS3 1 -----MNPVSLLSLSGERRSANWKPSNWDS
LcTOS 1 MSLIIQSLPHWSRIPPRPQLSQFQNSSRPKPLIQAGQVQHNAIQIARRSANYHPSIWDP
LaBERS 1 -----MEARRSGNFESSIWD
LdTPS7 1 -----MEARRSGNFKASIWDD

CoTPS3 26 NQIHQSLKSDFNDLQEKWHTELKQAVEQMLEAVAEPQKLTLLIDDIQRLGVAYRFEKQID
LcTOS 61 QYIESLKSPLYGDECFGRLEKLFKFEAKRLLLEATIEPIESWDELVDSIQRLGVAYHFEDK
LaBERS 17 DYIQSLTSSYTGKMYVDKSEKLEIEVKMMMDEATDELEQLLELINDLQRLGISYHFKDGLA
LdTPS7 17 DFLQSLTSPYTAKEYLKQADKLLKQVKVVIKETKQRLDQLDLIDNIQRLGISHHFRDEIQ

CoTPS3 86 DALSSIIYSN----YAAEVSSKKDLAASLYFRLLRQHGCYVSPDIFIQFKDEAGQFKASL
LcTOS 121 EGLDGVYG-----VGAHAGDDLTYAALQFRLLRQHGYGVTPDIFNKFLEKERTFKACT
LaBERS 77 KMLNNIYK-----SDSKYMEKDLHLTALKFRLLRQHGYRVQDVFSSFMDDGEGFEAWV
LdTPS7 77 RVLQNIYEKMRVECPDRMLMEKDLYSTSLQFRLLRQHGYHVSQDVFCSFMDGAGNQA--

CoTPS3 142 GDDVEGLLSLYEASYLEGKIKETILDDAKAFSTSTLENLMPHV-EADIASRISHAHLPLH
LcTOS 174 SLDAKGLLSLYEASHTMIHGEEVLEDAKEFVSVKHLNYLMGNLQN-NLREQVQHALEMPLH
LaBERS 131 VEDVSVLVLSYEASHISVEGESILDMADKDFSSHHTEMVEQIGEACLAEQVKRTLPLH
LdTPS7 135 VDDLKGI LALYEASFLSREGENILGSA RDFS TRH LKQKLEETDPI LAEKIRRALPLH

CoTPS3 201 WNMRRMEARLYIDVYRENKKRRNDNLLEFARLDFNMLQVIHQDLKDVSFVWDFLDLFRK
LcTOS 233 WRMPRLEAKHYIDVN-GRSDERNMVLLELARLDFNFVQSKHQEELKEVSRWWRDLGLAKK
LaBERS 191 WRVGRLEARWFVQAY-ETFPNSNPTLVELAKLDFNMVQAKYQDELKRCRSRWYEETGLPEK
LdTPS7 195 WRLQKLEAIWFINIY-ESRF DANLILLQLAKLEPNMVQAQYQEDLKLWLSRWYKETGLPEK
DDXXD

CoTPS3 261 LGEIRDRLMESFIFSVGLNFEPQFSECRKAATKDILLITVLDIYDIYGSMDVEEINNA
LcTOS 292 LGEFSRDRLVENYLWAVGIAPEPKFSNCRKGLTKLISILTVIDDIYDVYGSLELELPTFA
LaBERS 250 MSFARHRLAECFLWSLGFIPDPHHGYSREIMTKIAVLITITDDIYDIYGALEELQEFTEA
LdTPS7 254 MNEARDRLAECFLWALGPIPEAHLGQARKILTIAVLIIVIMDDFYDIYGTLDIYKVFTE

CoTPS3 321 VNRWDLGAVDELPEYMQLCYLGLLNSVNELAVVTMKITGRNVLDLFLKKLWKRHFNAVKE
LcTOS 352 VKRWIDIEALETLPYMKICYLALFNFVHEVSYDTLKDYGWNILPFIREEWEERLCMSYLVE
LaBERS 310 FERWDINSLDLLPEYMQICFLAIFNSANELGYQILRDQGLNIIPNLKRSWAELS RAYLE
LdTPS7 314 LQRWDINALDNLPEYMRICFLAIFNTANEIAYDILRDQGINIISNLRLRLWAEELGRVYITE

CoTPS3 381 SRWFHRQYPTTLEEYMENAQISIGAPLVLTHAVKMLKYPNEDVNHVDKYLKL SMCY
LcTOS 412 AEWFGNGNKPALDEYLRNGWISVGGPVAMVHAYFLQGRPIRKDSINFLDHGSELTYWSSV
LaBERS 370 ARWFHNGFVPTTDDQYLN TAWISISGPLLLSYGYLTTNPIPNNKELKSLEKHPSIIRWPSM
LdTPS7 374 AKWYHSGYFPSTEEYLNVAWISITGPVLLFHAYFSIMNPIDMKELQYLEQYPGIIRWPSM
NSE/DTE

CoTPS3 441 VFRLYDDWGTSKAEIERGDVPKAIQCYMHEAKVSEETAREHIKNIINERWRELNEECLKA
LcTOS 472 ATRLNDDLGTSKAEMKRGDVPKAVECYMIQTGESYEDAREHIQGLVDRDCWKMMNEECLKC
LaBERS 430 VLRLADDLGTSSSEIKRGDVSISIQCYMNETGCCGEGDARHHVKSLEVALKRMNDEILME
LdTPS7 434 VLRLADDLGTASDEIKRGDVPKSIQCYMHETGCSEBEAREYVKQLIDTTLKMMKEILME

CoTPS3 501 TDLNRKFVAAVLDALFAAAFYHHRDGGFGEFDHKFKSQAMAFSQQV-----
LcTOS 532 CLPK-SYVETVLNMVRTAQCIYQHGDGIGTSTGVTQDRVISLICEPVPSQWP
LaBERS 490 KPFK-SFDTNAMNLRISLFCFYQYGDGFGKPHSDTIKNLVSLIVLPPHMP--
LdTPS7 494 KPTN-DFGATAMNLRISLFFYQYGDGFGVPHNQTKENLVSLIVKPICLT--

```

Supplementary Fig. S9 (Continued). (C) Alignment of deduced amino acid sequences of CoTPS3 and other plant TPSs. LcTOS, *Litsea cubeba* trans-ocimene synthase (AEJ91554); LaBERS, *Lavandula angustifolia* α -bergamotene synthase (Q2XSC4); LdTPS7, *Lippia dulcis* α -bergamotene synthase (J7LQ09). The conserved motifs DDXXD and NSE/DTE are marked. Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

D

```

CoTPS4 1 MAATE-NLSLLAQS-SQPWAGIYGSHGSPFPISSWLRRQSIAKTSYICMCTPLSMSQLIA
OeGES1 1 MDCTMTSISLFSQS-----SNGISGTARSPFQWPINHRFSSGQRDFICKSLPVSSESA
ObGES 1 MSCARITVTLPYRS-----AKTSIQRGITHYPALIRPRFSACTPLASAMPLS---S
CrGES 1 MAATISNLSFLAKSRALSRPSSSSLSWLERPKTSSTTCMSMPSSSSSSSSSS--SMSLPLA

CoTPS4 59 TPLITDIE-SLLKYLRQPQVLPHEIDDSTKRRR-LLERTRRELQT-TLEPLQAMKMITL
OeGES1 54 TPLIPAENGAMYNRIQPVIVTPEVDDGTRHSE-LVERTRRELQR-STKPVETLKLIDNL
ObGES 49 TPLINGDN-SQRKNTRQ----HMEESSKRREYLLLEETTRKLRNDTESVEKLLIDNI
CrGES 59 TPLIKDNE-SLIKFLRQPLVLPHEVDDSTKRRR-LLERTRKELELNAEKPLEALKMIDI

CoTPS4 116 QRLGLAYHFEDDINSILTG-FS--NGQPEDLLTASLRFRLLRHNHRINPNIQKPFMDK
OeGES1 112 QRLGIAYYFEDDINAILDQ-FS--DGLPDEDLFTALCFRLLRDQRLOTGSDVFLKFMEX
ObGES 103 QQLGIGYYPEDAINAVLRSPFS----TGEEDLFTAALRFRLLRHNGIEISPEIFLKFDE
CrGES 117 QRLGLSYHFEDDINSILTG-FSNISSQTHEDLLTASLCFRLLRHNHGKINPDIEQKFMEN

CoTPS4 173 QGKHIDSLKEDTRGLFSLYEASYL GANGEDILLQALEFTKAHLKESLPSLAPPKAKVQSQ
OeGES1 169 NMKFKHELAQDTIGLVSLYEASSMGANGEEILSEAKEFTEMLHRQSMPLAPQLRQVSS
ObGES 159 RGFHDE---SDTLGLLSLEYEASNLGVAAGEEILEEAMEFAEARLRRSLSEPAPLHGVEAQ
CrGES 176 NGKPKDSLKDDTLGMLSLEYEASYL GANGREILMEAQEFTKTHLKNSLPAMAPSLSKVQSQ

CoTPS4 233 ALELPRHRRMARLEARRYIEEYGGENGHSPDLELAKLDYNKVSLSHQLELSEISRWWKQ
OeGES1 229 ALELPRHLRMARLEARRYIEEYGNESDHFALLELARLDYNKVLQHQMELAEITRWKQ
ObGES 216 ALDVPRHRLMARLEARRFIEQYQKQSDHDGDLLELAILDYNQVQAQHQSELTEIRWWE
CrGES 236 ALHQPRHRRMLRLEARRFIEEYGAENDHNPDLLELAKLDYNKVSLSHQMELSEITRWKQ
DDXXD

CoTPS4 293 LGLVDKLTAFARDRPLECFLWTVGILPEPKYSSCRIELAKTIAILLVIDDIFDTHGTLDL
OeGES1 289 LGLVEKLSFARDRPLECFLWTVGILLPEPKYSSCRIELAKTIAILLVIDDIFDITYGKMEEL
ObGES 276 LGLVDKLSFGRDRPLECFLWTVGILLPEPKYSSVRIELAKAISILLVIDDIFDITYGEMDDL
CrGES 296 LGLVDKLTAFARDRPLECFLWTVGILLPEPKYSGCRIELAKTIAILLVIDDIFDTHGTLDL

CoTPS4 353 ILFTNAIRRDLEAMEDLPEYMRICYMALYNTTNEICYKILKQNGWSVLPYLKATWIDMI
OeGES1 349 VLFTAEAIQRWDLDELETLPYMRICYMALYNTTNEICYKILKEYGFCVLPYLKSTWIDMI
ObGES 336 ILFTDAIRRDLEAMEGLPEYMRICYMALYNTTNEVCYKVLKRDGTGRIVLLNKSTWIDMI
CrGES 356 LLFTNAIKRDLEAMEDLPEYMRICYMALYNTTNEICYKVLKENGWSVLPYLKATWIDMI

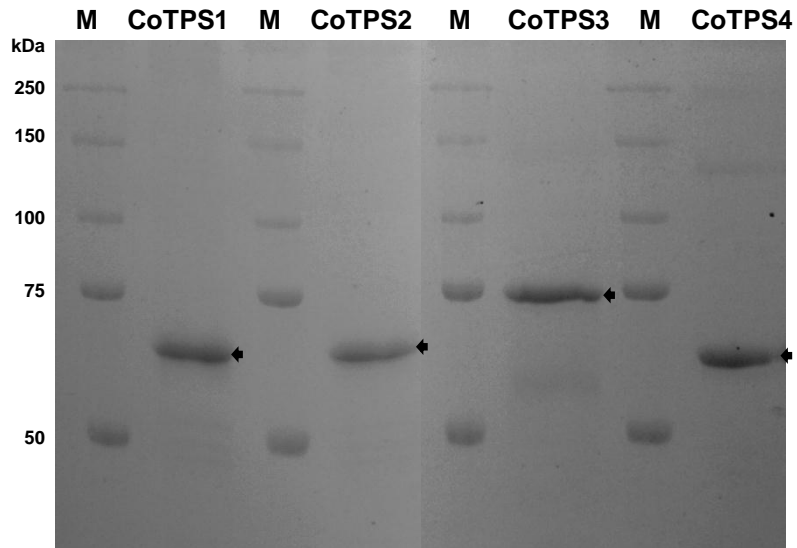
CoTPS4 413 EGFMLEASWLNNTGYVPMMEEYVENGVRTAGAYMALVHLFFLLIGQGVTEEN-VKLLVQKYP
OeGES1 409 EGFNVEANWFNGGHGPNLEEYIENGVSTAGAYMALVHLFFLLIGEGVTNENIAKLLRKPYP
ObGES 396 EGFMEEAQWFGGSAFKLEEYIENGVSTAGAYMAFAHIFLLIGEGVTHQNSQLFTQKYP
CrGES 416 EGFNVEAEWFNDSYVPMMEEYVENGVRTAGSYMALVHLFFLLIGQGVTEEN-VKLLIKYP
NSE/DTE

CoTPS4 472 KLFSYSGRILRLWDDLGTAKEEQERGLDASSIDLPMRENNITSDEEGRKCIILKIIDLNLK
OeGES1 469 KLFSAAGRILRLWDDLGTAKEEEREGDLASCMLQILMREKNIDCENEGRNYILKALNGLWR
ObGES 456 KVFSAAGRILRLWDDLGTAKEEQERGLDASCVQLPMKEKSLT-EEEARSRILEEIKGLWR
CrGES 475 KLFSSSGRILRLWDDLGTAKEEQERGLDASSIQLPMREKEIKSEEEGRKGIILEIENLWK

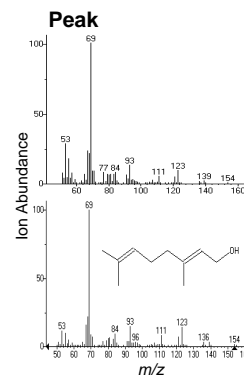
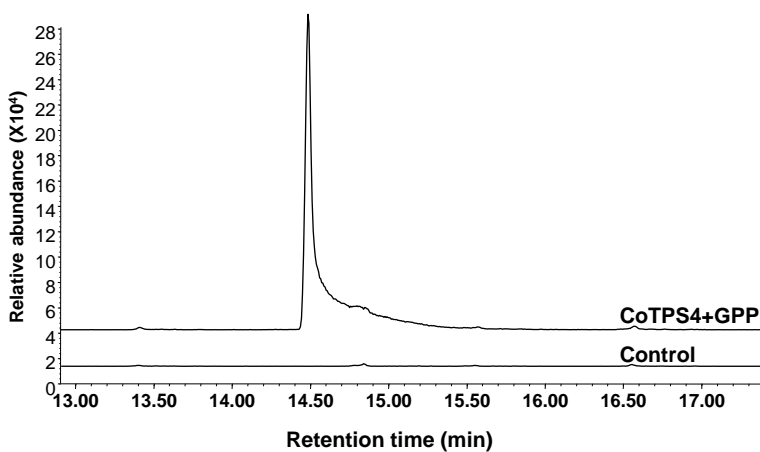
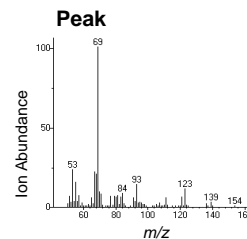
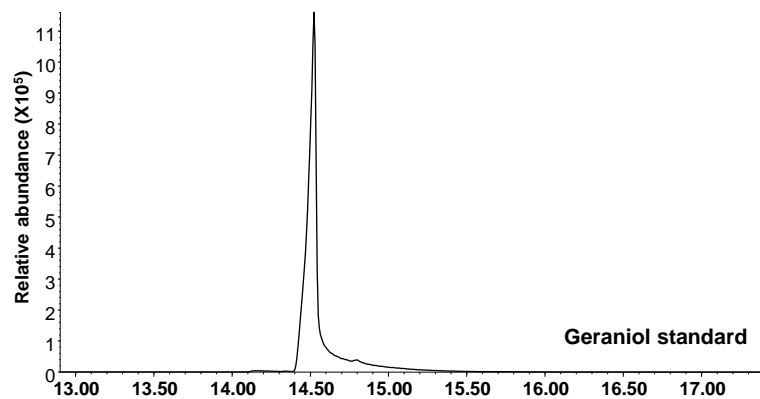
CoTPS4 532 ELNGELVSEHALPLAIIKAAFNMARASQVVYQHEEDSYFSSVDNYVQALFFTPFN
OeGES1 529 DLNDELISPAMPLAITKVALNMARAFEVVIKHEEDSYFSSVDNYVQALFFTPIN
ObGES 515 DLNGELVYNKNLPLSIIKVALNMARASQVVYKHDQDTYFSSVDNYVDALFFTP--
CrGES 535 ELNGELVYREEMPLAIIKTAFNMARASQVVYQHEEDTYFSSVDNYVKALFFTPCF

```

Supplementary Fig. S9 (Continued). (D) Alignment of deduced amino acid sequences of CoTPS4 and other plant TPSs. ObGES, *Ocimum basilicum* geraniol synthase (AAR11765); OeGES1, *Olea europaea* geraniol synthase 1 (AFI47926); CrGES, *Catharanthus roseus* geraniol synthase (AFD64744). The conserved motifs DDXXD and NSE/DTE are marked. Completely conserved residues are shaded in dark gray, and identical residues are shaded in gray, and similar residues are shaded in light gray. Dashes indicate gaps introduced to maximize sequence alignment.

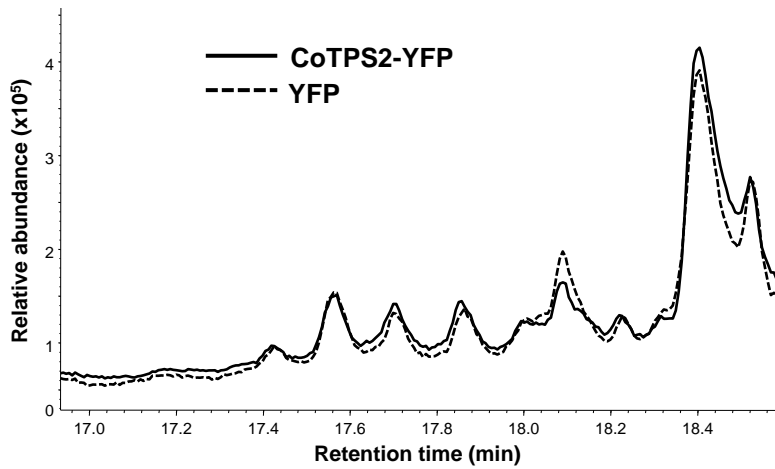


Supplementary Fig. S10. SDS-PAGE gel showing purified recombinant 6His-tagged CoTPS1, CoTPS2, CoTPS3 and CoTPS4 proteins. The position corresponding to the recombinant proteins are indicated by an arrow. M; Precision Plus Protein all blue standards (Bio-rad), kDa; kilodalton.



Geraniol

Supplementary Fig. S11. *In vitro* enzymatic assay of recombinant 6His-tagged CoTPS4s using GPP. The reaction products were analysed by GC-MS. The peak was identified by the mass spectra reference library and comparison of retention time with those of authentic standard (geraniol standard). Mass spectra for the geraniol standard and the peak formed with GPP are shown on the right. m/z, mass-to-charge ratio. Control, control assay using a heat-inactivated recombinant protein.



Supplementary Fig. S12. Transient expression of CoTPS2-YFP or YFP in *N. benthamiana*. CoTPS2-YFP or YFP was transiently expressed in *N. benthamiana* leaves by *Agrobacterium*-mediated infiltration. The chemical composition was analysed 3 dpi (days post-infiltration) by GC-MS.

Supplementary Table 1. Oligonucleotide primers used in this study.

| Name | Forward | Reverse |
|--------------------------------------|--|---------------------------------------|
| for qRT-PCR | | |
| DXS1 | TTTGATTGGGGGTGAGAGAG | GCCATCAAGGGACATGAACT |
| DXS2 | GGGGATCACAAAATCCATTG | TGGGTAGCCTTTTCCCTTTT |
| DXS3 | CTTCTGCAAGCCCCTAGATG | GTTGCAGCAATGTGCCTAGA |
| DXS4 | AAGGGGAGAATCCTTGCACT | TGTTCTTGGCAAAGCTTCT |
| DXR | GTTCTGAGTGCAGCCAATGA | CAAGAACAGGTCGCAGATCA |
| CMS | AGGAGCGTGCAGTGGTCTT | TGGCAAATTCAGGTCAACA |
| CMK | TCGGTCTGATTGTCCCTTC | CGAAGGCGCTTGTAACCTTC |
| MCS | ACGCAAGCAAAAACCTCTCGT | CGGTAGCGTTATCAGCATCA |
| HDS | GTGGTCCATTTGAGGAAAA | AATCCACAGTGGCAAGATCC |
| IDI1 | ACTGTACAAGGCCCATCTG | CCCACCATTCAGAAGGAAA |
| IDI2 | CATGAGTTGCTCCTTCAGCA | AAGCTTCTTTGTGCAGCAT |
| AACT | TGCAGTGAAGTCTGCTGGTC | CAACAGAGTTGCCACAGC |
| HMGS | CAAGACTTGGCTGATGCAAA | AACAACAAGCCAAATCGTC |
| HMGR | CAGGCTGATTTCCCTGACAT | CTGGCATGTGCATTGTATCC |
| MVK | GGTGTTCGGAGAGAGCTTG | ACCCCATGCACTGAAGTAG |
| PMK | TGCTCTCGATGCTACAATG | AGGGTCTTCTCTGACCAGCA |
| MPDC | CCCATCAATGACAGCATCAG | CGTATTCGGATCTCCCTCAA |
| DAHPS | ATCCTCCATCTTCCCTGCTT | TTCTTGGATTTCAGCCATC |
| DHQS | CCTTTCCTTCCGCTTCTCT | TTCCATGAACATGCCTTTGA |
| SDH | TGATCAGCTAGAGCGCAGAA | CCATCTGTTGACGGATTTT |
| SK | CACCGTCCGAAGATTGTGTT | CCAGACCCCATCATACCAAC |
| CS | GCTGCTGGAAGTCTTTTGG | TCTGTCCTGGCCTCCTTCTA |
| PAL | AATCTCTCAGGGGGAAGGAA | CAACCCACGGAGTTTACAT |
| SAMT | TTAACGGCAGGAAGAGCCTA | ATCACGGCCTTCATTTCTTG |
| CHS | TTCTGTTCTCAAGCGGAGT | CGGAAGTAGCCACATTCAT |
| CHI | CATGAGCATGGTCAGGAAGA | GCACAGCAACTCACTTTCCA |
| F3H | TATCTCTGGTGGGATCGAC | CTGCTTCTCCCTGAAGATGG |
| DFR | GCAATGGTGATTGGTCTCT | GAAGCGGAATTGTATGGGTAA |
| C4H | ACCCACAAGCTTCCCTACCT | AGGTACCAGGCATCACAGG |
| 4CL | GTTGCACCTGCAGAACTTGA | ACCTTCCGAGTCTCTTGAA |
| CCR | GGAGCGATCTGGACTACTGC | GAACATCGACGTAGGCCTGT |
| CVOMT | GCATGCTCACTCATTTCTGGA | CCCAAAGACTTTTTCCGTGA |
| C3H | ACGCTCCAGGAGAAGTACGA | CACACGGTCCAATTCCTTT |
| COMT | ACTCTCCAGTGCCTGTGCTT | AGTGCCCGTTGTCTAGTGT |
| CCOMT | TTGGATACCAGCTCTACCC | CTCACGGTTGATGCCATTG |
| CAD | AAAGTGGGAACAACGAATGC | ACCGACCTCTTGACCACAC |
| TPS1 | AGGGCCTTCCAGATTACATGAAGC | AGCGTTGGAGTGTACTTGCTATTG |
| TPS2 | CATCGATCATGGATGACACATATGATG | CTCAACAGCAGTCAGAAGGGCAC |
| TPS3 | AGGCTGTGGCTGAACCTCTA | AGCAGATCCTTCTTGCTGGA |
| TPS4 | TAGCCATCTTACTGGTGATTGACG | GTAGTGTGTACAATGCCATGTAAC |
| Actin | CTGGACGTGACCTCACAGATGCT | TCTTCTCAACAGAGGAGTGTCT |
| for TOPO cloning | | |
| TPS1 | CACCATGGCCCTGAATACGTTCTTGCAATTTCC | CTCAAGCAGATGGGTTCAACGGTAAAGTGA |
| TPS2 | CACCATGGCACTTATATTTGCAATGGCCACTCTGA | ATCAACAAAACGGTCCACTAGCACCATCGCA |
| TPS3 | CACCATGAATCCTGTTTCTCTTTTGAGCTTATCAGGAG | AACTTGTGGGAAAATAGAGCCATGGCTTGA |
| TPS4 | CACCATGGCTGCTACGAGAAACCTTCTTTACTT | ATTGAAAGCGGTGAAAACAAAGCTTGA |
| for E.coli expression vectors | | |
| TPS1 | aaGGATCCgATGGCCTTGAATACGTTCTTGCAATTTTC | aaCGGGCCGCTTACTCAAGCAGATGGGTTCAAC |
| TPS2 | aacGTCGACATGGCACTTATATTTGCAATGGCCACTCTGA | aaaCGGGCCGCTTAATCAACAAAACGGTCCACTAGCA |
| TPS3 | aacGGATCCgATGAATCCTGTTTCTCTTTTGAGCTTATCAGGAG | aacGTCGACAACTTGTGGGAAAATAGAGCCATGGCT |
| TPS4 | aaaGGATCCcATGGCTGCTACGAGAAACCTTCTTTACTT | aaaGTCGACATTGAAAGCGGTGAAAACAAAGCTTGA |

GGATCC, *Bam* HI; GCGGCCGC, *Not* I; GTCGAC, *Sal* I