**Uncovering a miltiradiene biosynthetic gene cluster in the Lamiaceae reveals a dynamic evolutionary trajectory**

Abigail E. Bryson, Emily R. Lanier, Kin H. Lau, Davis Mathieu, Alan E. Yocca, Garret P. Miller, Patrick P. Edger, C. Robin Buell, Bjӧrn Hamberger.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Species** | **Publication** | **Estimated Genome size (Mb)** | **Assembly coverage (%)** | **BUSCO**  **(% complete)** | **N50 contig** | **N50 scaffold** |
| *Callicarpa americana* | Hamilton *et al,* 2020 1 | 538 | 94.0% | 93.8% | 7.5 Mb | 29 Mb |
| *Hyssopus officinalis* | Lichman *et al,* 2020 2 | 341 | -- | 91.4% | n.r. | 29 Mb |
| *Lavendula angustifolia* | Malli *et al,* 2019 3 | 870 | 79.1% | 89.7% | n.r. | 96 Kb |
| Li *et al,* 2021 4 | 1016 | 90.0% | 91.4% | 1.2 Mb | 36 Mb |
| *Mentha longifolia* | Vining *et al,* 2017 5 | 400 | 88.0% | 71.0% | n.r. | 4 Kb |
| *Nepta cataria* | Lichman *et al,* 2020 2 | 493 | -- | 91.4% | n.r. | 296 Kb |
| *Nepta mussini* | Lichman *et al,* 2020 2 | 232 | -- | 90.2% | n.r. | 503 Kb |
| *Ocimum basilicum* | Bornowski *et al,* 2020 6 | 2360 | 87.6% | 86.7% | 48 Kb | 1.5 Mb |
| *Ocimum sanctum* | Rastogi *et al,* 2015 7 | 386 | -- | n.r. | 12 Kb | 303 Kb |
| *Ocimum tenuiflorum* | Upadhyay *et al,* 2015 8 | 612 | 61.0% | n.r. | n.r. | 27.1 Kb |
| *Origanum majorana* | Bornowski *et al,* 2020 6 | 858 | 88.8% | 89.5% | 35 Kb | 1.3 Mb |
| *Origanum vulgare* | Bornowski *et al,* 2020 6 | 705 | 89.4% | 90.1% | 26 Kb | 157 Kb |
| *Perilla citriodora* | Y. Zhang et al, 2021 9 | -- | -- | 92.6% | 68 Kb | 1.2 Mb |
| *Perilla frutescens* | Y. Zhang et al, 2021 9 | 1379 | 97.5% | 92.7% | 3.21 Mb | 62.64 Mb |
| *Pogostemon cablin* | He *et al,* 2018 Nature 10 | 2080 | 91.8% | 90.6% | 34 Kb | 699 Kb |
| He *et al,* 2016 Nature 11 | 1576 | 73.0% | n.r. | 416 bp | 1 Kb |
| *Rosmarinus officinalis* | Bornowski *et al,* 2020 6 | 1180 | 85.8% | 90.1% | 21 Kb | 368 Kb |
| *Saliva miltiorrhiza* | H. Xu *et al,* 2016 12 | 615 | 87.5% | n.r. | 12 kb | 51 Kb |
| G. Zhang *et al,* 2015 13 | 645 | 99.2% | n.r. | 82 kb | 1.2 Mb |
| Song *et al,* 2020 14 | 600 | 99.1% | 92.5% | 2.7 Mb | 69 Mb |
| Ma *et al,* 2021 15 | 557 | 91.5% | 91.1% | 500 Kb | 1.2 Mb |
| *Salvia splendens* | Dong *et al,* 2018 16 | 711 | 92.2% | 92.1% | 2.2 Mb | 3.1 Mb |
| Jia *et al,* 2021 17 | 807 | 95.0% | 92.0% | 3.8 Mb | 35.1 Mb |
| *Scutellaria baicalensis* | Q. Zhao *et al,* 2019 18 | 408 | 93.0% | 94.0% | 1.3 Mb | 33.2 Mb |
| Z. Xu *et al,* 2020 19 | 440 | 85.3% | 91.5% | 2.1 Mb | 40.8 Mb |
| *Scutellaria barbata* | Z. Xu *et al,* 2020 19 | 405 | 87.2% | 93.0% | 2.5 Mb | 23.7 Mb |
| *Tectona grandis* | Yasodha *et al,* 2018 20 | 465 | 79.8% | n.r. | 277 Kb | 357 Kb |
| D. Zhao *et al,* 2019 21 | 325 | 104.0% | 92.3% | 3 Mb | 16 Mb |

**Supplemental Table 1. A collection of Lamiaceae genomes** **published to date.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Number of Scaffolds** | **Total Size of Scaffolds (bp)** | **Total Size of Scaffolds (Mb)** | **N50 Contig Length (bp)** | **Number of Ns (Percent Ns)** | **Totals Gaps (Consecutive Ns)** | **Largest Scaffold (bp)** | **Largest Scaffold (Mb)** |
| *Leonotis leonurus* | 23,651 | 585,264,293 | 585.26 | 1,094,942 | 40,883,810 (7.0%) | 15,483 | 11,593,990 | 11.59 |
| *Plectranthus barbatus* | 62,959 | 1,249,907,925 | 1,249.91 | 258,138 | 70,313,430 (5.6%) | 30,507 | 3,093,914 | 3.09 |
| *Prunella vulgaris* | 46,736 | 820,275,670 | 820.28 | 444,240 | 38,970,920 (4.8%) | 20,293 | 5,268,047 | 5.27 |

**Supplemental Table 2. Assembly statistics for *Leonotis leonurus*, *Plectranthus barbatus*, and *Prunella vulgaris*.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Complete BUSCOs (C)** | **Complete Single-Copy BUSCOs (S)** | **Complete Duplicate BUSCOs (D)** | **Fragmented BUSCOs (F)** | **Missing BUSCOs (M)** |
| *Leonotis leonurus* | 98.5% | 80.0% | 18.5% | 0.5% | 1.0% |
| *Plectranthus barbatus* | 97.8% | 84.4% | 13.4% | 1.0% | 1.2% |
| *Prunella vulgaris* | 97.1% | 91.8% | 5.3% | 1.5% | 1.4% |

**Supplemental Table 3. BUSCO results for *Leonotis leonurus*, *Plectranthus barbatus*, and *Prunella vulgaris*.** BUSCO run in genome mode with lineage dataset 'embryophyta\_odb10.'

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Enzyme Family** | **Enzyme Name** | **GenBank Accession** | **Species Name** | **Enzyme Function** |
| TPS-c (Class II) | CamTPS2 | QMW69082.1 | *Callicarpa americana* | (-)-kolavenyl diphosphate synthase |
| TPS-c (Class II) | CamTPS1 | QMW69081.1 | *Callicarpa americana* | *ent*-copalyl diphosphate synthase |
| TPS-c (Class II) | CamTPS3 | QMW69080.1 | *Callicarpa americana* | *ent*-copalyl diposphate synthase |
| TPS-c (Class II) | PcTPS1 | AZB50385.1 | *Pogostemon cablin* | 10R-lambda-8,13E-dienyl diphosphate synthase |
| TPS-c (Class II) | ArTPS2 | AZB50378.1 | *Ajuga reptans* | cleroda-4(18),13E-dienyl diphosphate synthase |
| TPS-c (Class II) | LlTPS1 | AZB50381.1 | *Leonotis leonurus* | peregrinol diphosphate synthase |
| TPS-c (Class II) | ArTPS1 | AZB50377.1 | *Ajuga reptans* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | PbTPS16 | AZB50379.1 | *Plectranthus barbatus* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | SyTPS1 | AZB50384.1 | *Salvia yangii* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | NrTPS1 | AZB50382.1 | *Nepeta racemosa* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | OmTPS1 | AZB50383.1 | *Ocimum basilicum* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | MsTPS1 | AZB50380.1 | *Mesosphaerum suaveolens* | lambda-7,13E-dienyl diphosphate synthase |
| TPS-c (Class II) | MvCPS3 | A0A075FA51.1 | *Marrubium vulgare* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | PbTPS1 | AHW04046.1 | *Plectranthus barbatus* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | SdCPS1 | APH81399.1 | *Salvia divinorum* | *ent*-copalyl diphosphate synthase |
| TPS-c (Class II) | SmCPS1 | AHJ59321.1 | *Salvia miltiorrhiza* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | SmCPS2 | AHJ59322.1 | *Salvia miltiorrhiza* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | SmCPS4 | AKN91186.1 | *Salvia miltiorrhiza* | copal-8-ol diphosphate synthase |
| TPS-c (Class II) | SmCPS5 | AHJ59324.1 | *Salvia miltiorrhiza* | *ent*-copalyl diphosphate synthase |
| TPS-c (Class II) | PbTPS2 | AHW04047.1 | *Plectranthus barbatus* | hydroxy-labdane diphosphate synthase |
| TPS-c (Class II) | SdKPS2 | A0A1S5RW73.1 | *Salvia divinorum* | (-)-kolavenyl diphosphate synthase |
| TPS-c (Class II) | IrCPS1 | A0A1X9IRQ7.2 | *Isodon rubescens* | (+)-copalyl diphosphate synthase |
| TPS-c (Class II) | VacTPS3 | A0A2K9RFZ8.1 | *Vitex agnus-castus* | *Syn*-copalyl diphosphate synthase |
| TPS-c (Class II) | IeCPS2 | G3E4M4.1 | *Isodon eriocalyx* | *ent*-copalyl diphosphate synthase |
| TPS-c (Class II) | VacTPS5 | A0A2K9RG07.1 | *Vitex agnus-castus* | (-)-kolavenyl diphosphate synthase |
| TPS-c (Class II) | IeCPS1 | G3E4M6.1 | *Isodon eriocalyx* | *ent*-copalyl diphosphate synthase |
| TPS-c (Class II) | SsLPS | AET21247.1 | *Salvia sclarea* | 13-labden-8,15-diol diphosphate synthase |
| TPS-c (Class II) | MvCPS1 | AIE77090.1 | *Marrubium vulgare* | peregrinol diphosphate synthase |
| TPS-e (Class I) | SmKSL1 | ABV08817.1 | *Salvia miltiorrhiza* | miltiradiene synthase |
| TPS-e (Class I) | SmKSL2 | AHJ59325.1 | *Salvia miltiorrhiza* | *ent*-kaurene synthase |
| TPS-e (Class I) | LlTPS4 | AZB50368.1 | *Leonotis leonurus* | sandaracopimaradiene synthase |
| TPS-e (Class I) | SdKSL1 | AOZ15897.1 | *Salvia divinorum* | *ent*-pimara-8(14),15-diene synthase |
| TPS-e (Class I) | VacTPS2 | A0A2K9RFZ2.1 | *Vitex agnus-castus* | (13S)-vitexifolin A synthase; 9,13-epoxylabda-14-ene synthase; Viteagnusin D synthase |
| TPS-e (Class I) | IrKSL6 | A0A1Z3GCD1.1 | *Isodon rubescens* | isopimaradiene synthase |
| TPS-e (Class I) | MsTPS1 | AZB50369.1 | *Mentha spicata* | sandaracopimaradiene/ labd-13(16),14-diene-9-ol/ ent-13-epi-manoyl oxide synthase |
| TPS-e (Class I) | NrTPS2 | AZB50370.1 | *Nepeta racemosa* | *ent-*kaurene synthase |
| TPS-e (Class I) | IrKSL3 | A0A1X9ISH5.2 | *Isodon rubescens* | nezukol synthase; kaurene synthase, miltiradiene synthase |
| TPS-e (Class I) | IrKSL4 | A0A1Z3GBK8.1 | *Isodon rubescens* | *ent*-atiserene synthase; kaurene synthase |
| TPS-e (Class I) | PbTPS3 | AHW04048.1 | *Plectranthus barbatus* | miltiradiene synthase |
| TPS-e (Class I) | PbTPS4 | AHW04049.1 | *Plectranthus barbatus* | miltiradiene synthase |
| TPS-e (Class I) | SsScS | G8GJ94.1 | *Salvia sclarea* | sclareol synthase |
| TPS-e (Class I) | MvELS | AIE77094.1 | *Marrubium vulgare* | 9,13-epoxy-labd-14-en synthase |
| TPS-e (Class I) | OmTPS3 | AZB50371.1 | *Ocimum basilicum* | trans-abienol synthase |
| TPS-e (Class I) | OmTPS4 | AZB50372.1 | *Ocimum basilicum* | manool synthase |
| TPS-e (Class I) | OmTPS5 | AZB50373.1 | *Ocimum basilicum* | palustradiene synthase |
| TPS-e (Class I) | SyTPS3 | AZB50374.1 | *Salvia yangii* | miltiradiene synthase |
| TPS-e (Class I) | PvTPS1 | AZB50375.1 | *Prunella vulgaris* | miltiradiene synthase |
| TPS-e (Class I) | SoTPS1 | AZB50376.1 | *Salvia officinalis* | miltiradiene synthase |
| TPS-e (Class I) | ArTPS3 | AZB50367.1 | *Ajuga reptans* | miltiradiene synthase |
| TPS-e (Class I) | PbEKS | AGN70881.1 | *Plectranthus barbatus* | *ent*-kaurene synthase |
| CYP76AK/AH | SmCYP76AH1 | AGN04215.1 | *Salvia miltiorrhiza* | ferruginol synthase |
| CYP76AK/AH | SmCYP76AK1 | A0A125QZE2.1 | *Salvia miltiorrhiza* | 11-hydroxysugiol 20-monooxygenase |
| CYP76AK/AH | SmCYP76AK2 | AJD25182.1 | *Salvia miltiorrhiza* | Involved in terpenoid biosynthesis |
| CYP76AK/AH | SmCYP76AK3 | AJD25183.1 | *Salvia miltiorrhiza* | Involved in terpenoid biosynthesis |
| CYP76AK/AH | SmCYP76AH3 | AMB36496.1 | *Salvia miltiorrhiza* | ferruginol oxidase |
| CYP76AK/AH | RoCYP76AK7 | A0A1D8QMD2.1 | *Rosmarinus officinalis* | carnosic acid synthase |
| CYP76AK/AH | RoCYP76AK8 | A0A1D8QMG4.1 | *Rosmarinus officinalis* | carnosic acid synthase |
| CYP76AK/AH | RoCYP76AH23 | KP091844.1 | *Rosmarinus officinalis* | ferruginol synthase |
| CYP76AK/AH | RoCYP76AH22 | A0A0C5Q4Y6 | *Rosmarinus officinalis* | ferruginol synthase |
| CYP76AK/AH | SpCYP76AK6 | A0A0S1TPC7.1 | *Salvia pomifera* | carnosic acid synthase |
| CYP76AK/AH | SfCYP76AK6 | A0A1D8QMD1.1 | *Salvia fruticosa* | carnosic acid synthase |
| CYP76AK/AH | SfCYP76AH24 | A0A0C5QRZ2 | *Salvia fruticosa* | ferruginol synthase |
| CYP76AK/AH | SdCYP76AH39 | QMS79243.1 | *Salvia divinorum* | crotonolide G synthase |
| CYP76AK/AH | PbCYP76AH15 | AMZ03402.1 | *Plectranthus barbatus* | 11-oxo-13R-manoyl oxide synthase |
| CYP76AK/AH | PbCYP76AH17 | AMZ03404.1 | *Plectranthus barbatus* | 11-oxo-13R-manoyl oxide synthase; 1,11-dihydroxy-13R-manoyl oxide synthase |
| CYP76AK/AH | PbCYP76AH9 | AMZ03391.1 | *Plectranthus barbatus* | keto-13R- manoyl oxide (8,13-epoxy-labd-14-ene-1 1 -one) synthase or hydroxy- 13R-manoyl oxide synthase |
| CYP76AK/AH | PbCYP76AH8 | AMZ03392.1 | *Plectranthus barbatus* | 11-oxo-13R-manoyl oxide synthase; 1,11-dihydroxy-13R-manoyl oxide synthase |
| CYP76AK/AH | PbCYP76AH11 | AMZ03393.1 | *Plectranthus barbatus* | 6,7-dihydroxy-11-keto-manoyl oxide synthase; 11-oxo-13R-manoyl oxide synthase |
| CYP71D | SmCYP71D375 | AWD93836.1 | *Salvia miltiorrhiza* | miltirone/4-methylenemiltirone 16-hydroxylase |
| CYP71D | SmCYP71D411 | AJD25164.1 | *Salvia miltiorrhiza* | 11-hydroxysugiol 20-monooxygenase |
| CYP71D | MgCYP71D18 | Q6WKZ1.1 | *Mentha x gracilis* | (-)-(4S)-Limonene-6-hydroxylase |
| CYP71D | MpCYP71D13 | Q9XHE7.1 | *Mentha x piperita* | (-)-(4S)-Limonene-6-hydroxylase |
| CYP71D | PbCYP71D381 | KT382342.1 | *Plectranthus barbatus* | 13R-manoyl oxide oxidase |
| CYP71D | AtCYP71B2 | OAP13205.1 | *Arabidopsis thaliana* | nf |
| CYP71D | AtCYP71B3 | OAP05595.1 | *Arabidopsis thaliana* | nf |
| CYP71D | ZzCYP71BA1 | BAJ39893.1 | *Zingiber zerumbet* | 8-hydroxy-alpha-humulene synthase |
| CYP71D | VvCYP71BC1 |  | *Vitis vinifera* |  |
| CYP71D | VvCYP71BE1 | BAT70339.1 | *Vitis vinifera* | Sesquiterpene oxidase |
| CYP71D | NtCYP71D16 | AAD47832.1 | *Nicotiana tabacum* | Cembratriene‐diol hydroxylase |
| CYP71D | SpCYP71BE52 | ALM25794.1 | *Salvia pomifera* | Salviol synthase |
| CYP71D | SmCYP71D373 | Ma et al 2021 | *Salvia miltiorrhiza* | 2-isopropyl-8-methylphenanthrene-3,4-dione |
| CYP71D | SmCYP71D464 | Ma et al 2021 | *Salvia miltiorrhiza* | nf |
| CYP71D | SmCYP71D410 | AJD25163.1 | *Salvia miltiorrhiza* | nf |

**Supplemental Table 4. Species, enzyme names, GenBank accessions, and listed functions for reference enzymes used in the phylogenies in this study.**

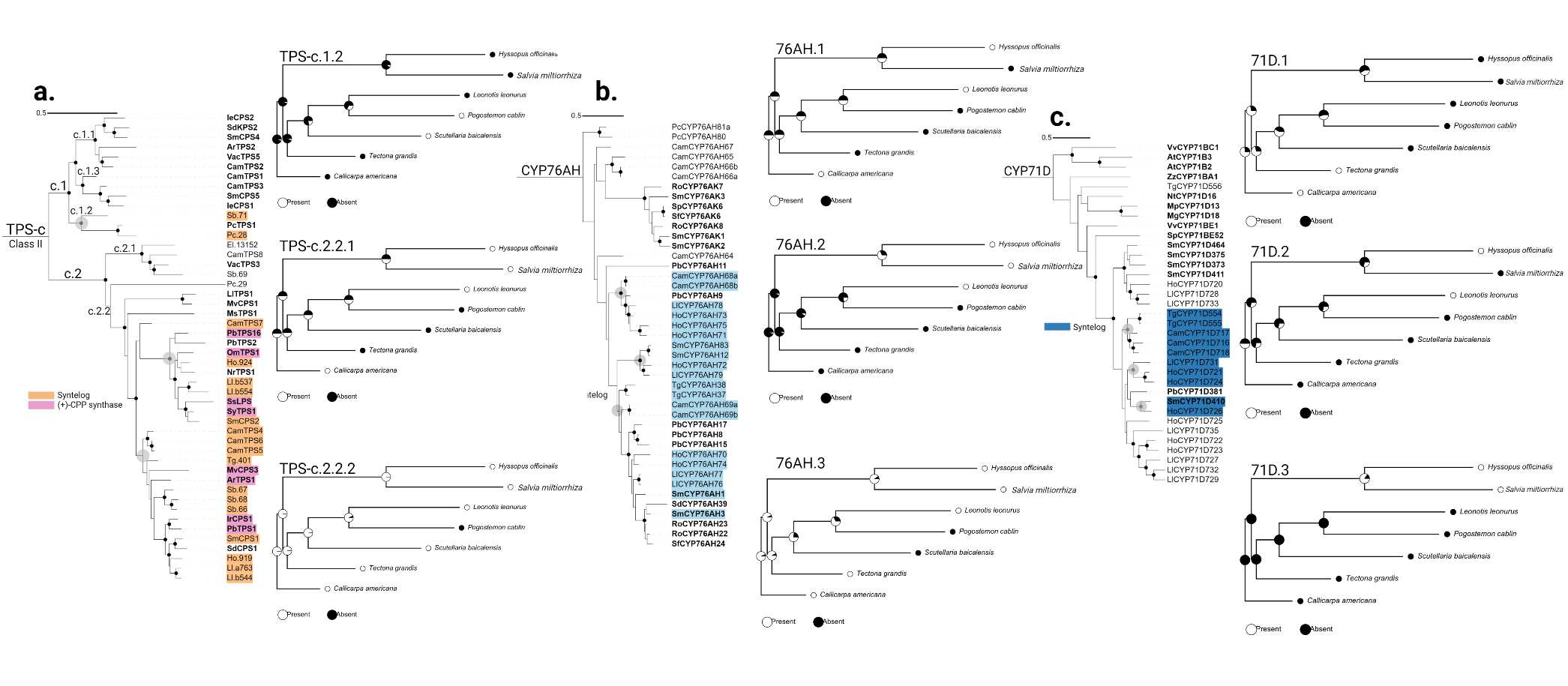
|  |  |  |
| --- | --- | --- |
| **Gene** | **Direction** | **Sequence, 5’-3’** |
| *CamCYP76AH64* | F | TCTAAATGGATTTCTTTGCTCTT |
| *CamCYP76AH64* | R | GGAGAGAATGTCTATGGTTCAC |
| *CamCYP76AH65* | F | ATGGATCTCTTTGAGCTCCTCGTT |
| *CamCYP76AH65* | R | TCAGAGCTTGATCGGAGTAGCCTT |
| *CamCYP76AH66* | F | ATGGATTTCTTTGCATTTTTTATTG |
| *CamCYP76AH66* | R | TCATGGTTTGATAGGAATTGC |
| *CamCYP76AH67* | F | GAACAAATTGAGATGGATCT |
| *CamCYP76AH67* | R | CATATTTCATAAGATCAAGGCT |
| *CamCYP76AH68* | F | ATGGATCTCTATGCCCTTTT |
| *CamCYP76AH68* | R | ATATTTCACAAAATCATCAGGG |
| *CamCYP76AH69* | F | CACTAGTCTCATGGAGTTT |
| *CamCYP76AH69* | R | ATAATCATTTGACCTGCTTGAAA |
| *CamCYP71D716* | F | TTTGTCAACCTTATGGAGTTT |
| *CamCYP71D716* | R | AAAAGGTCTCCTAACAGTG |
| *CamCYP71D717* | F | ATGGAGTTTGAGTTCCCATC |
| *CamCYP71D717* | R | TCATTTGACCTGCTTGAAAG |
| *CamTPS7* | F | ATGACCTCTATATCCACTGTTCT |
| *CamTPS7* | R | TCATACAACTGGTTCAAATAGTACT |
| *CamTPS9* | F | ATGTCACTCAGTTTCAACCTCA |
| *CamTPS9* | R | GCTTCATGAGGAGGAAGGAAGT |
| *CamTPS10* | F | ACCAATCCCATGAAACTTTCAT |
| *CamTPS10* | R | TATACATAAACCAAAATACCCAAATCG |

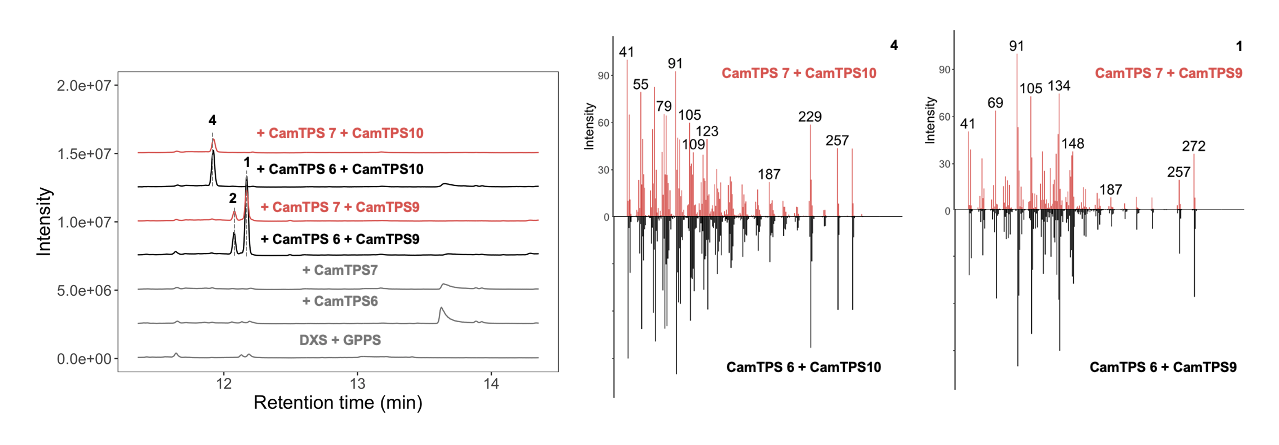
**Supplemental Table 5. Primers used in this study.**

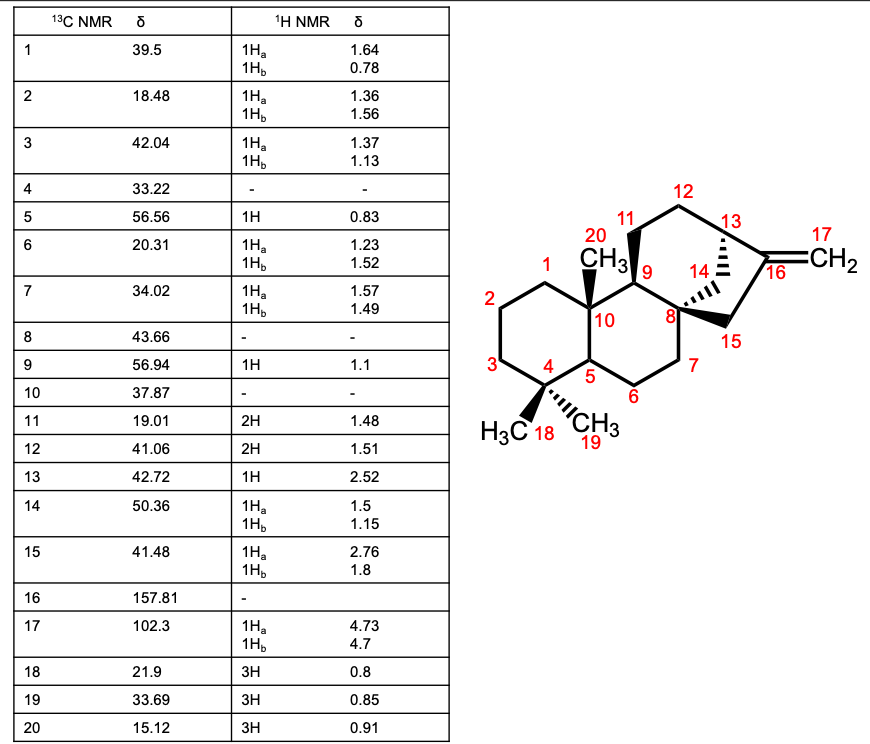
Diagram

Description automatically generated

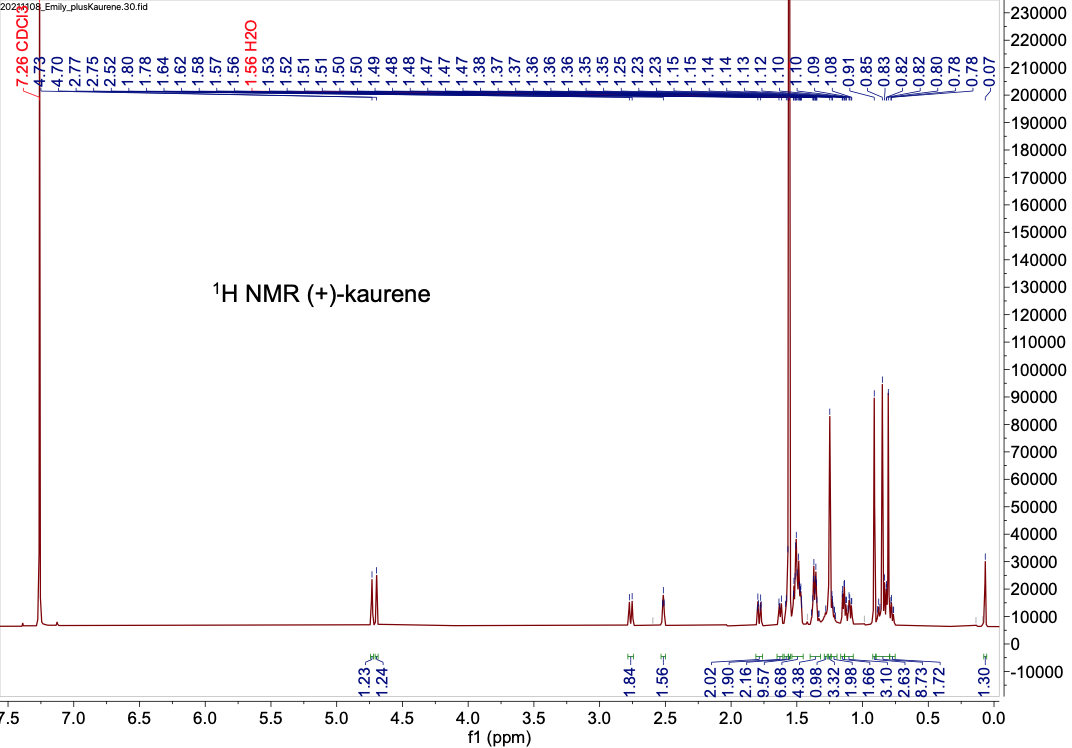
**Supplemental Fig. 1.** **Ancestral reconstruction of syntenic genes present across all examined species.** Circles at each node represent presence/absence (white/black, respectively) of each gene. Analysis was performed using the R package phytools.

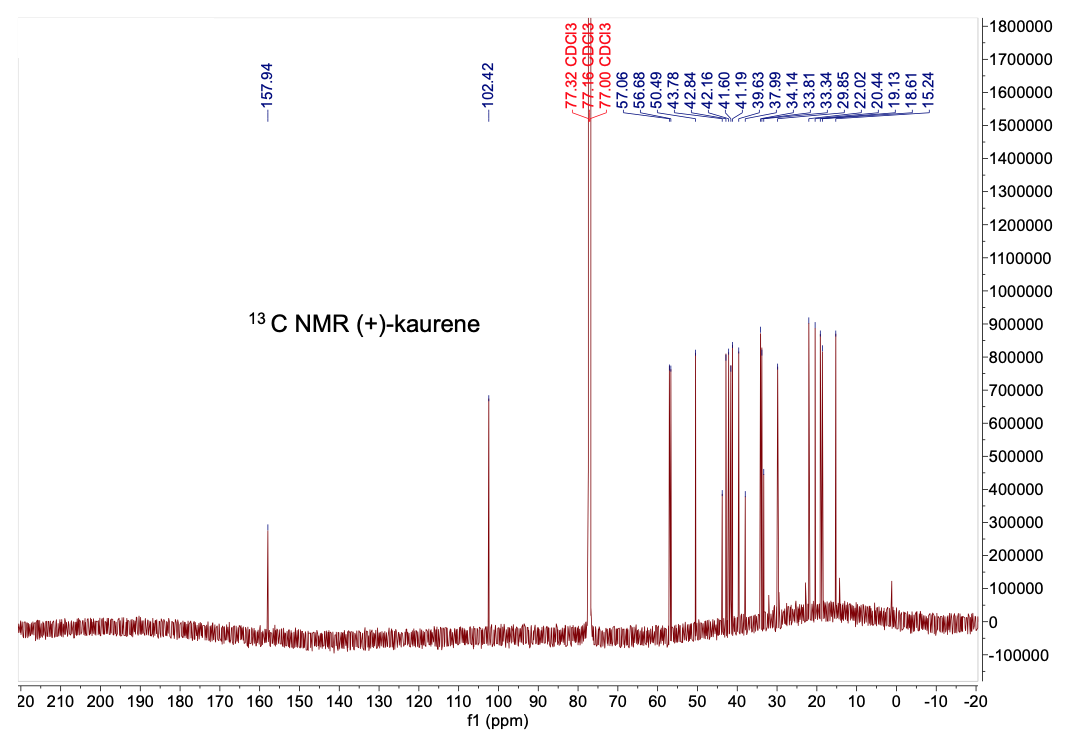
**Supplemental Fig. 2.** **Ancestral reconstruction of syntenic genes present across all examined species in a lineage-central manner.** Circles at each node represent presence/absence (white/black, respectively) of each gene. Each ancestral reconstruction is represented by gray nodes on the main enzyme phylogeny. Analysis was performed using the R package phytools.

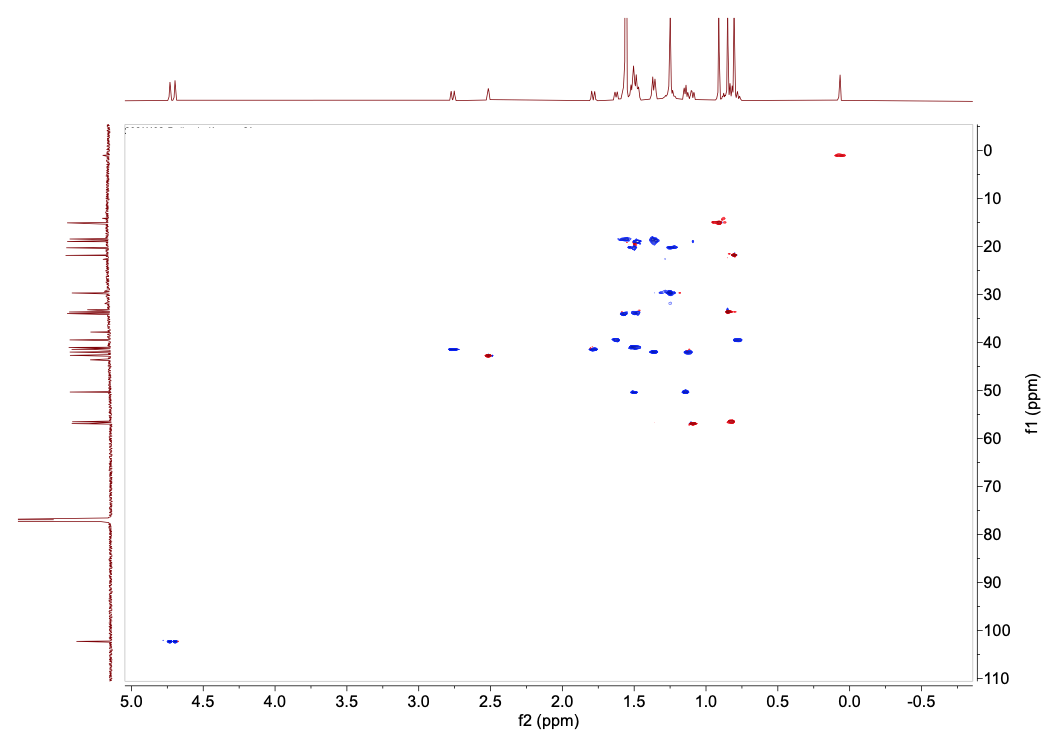
**Supplemental Fig. 3. Characterization of CamTPS7 as a (+)-CPP synthase by comparison with CamTPS6**. Both produce **1** & **2** when paired with CamTPS9 (miltiradiene synthase) and **4** when paired with CamTPS10 ((+)-kaurene synthase).

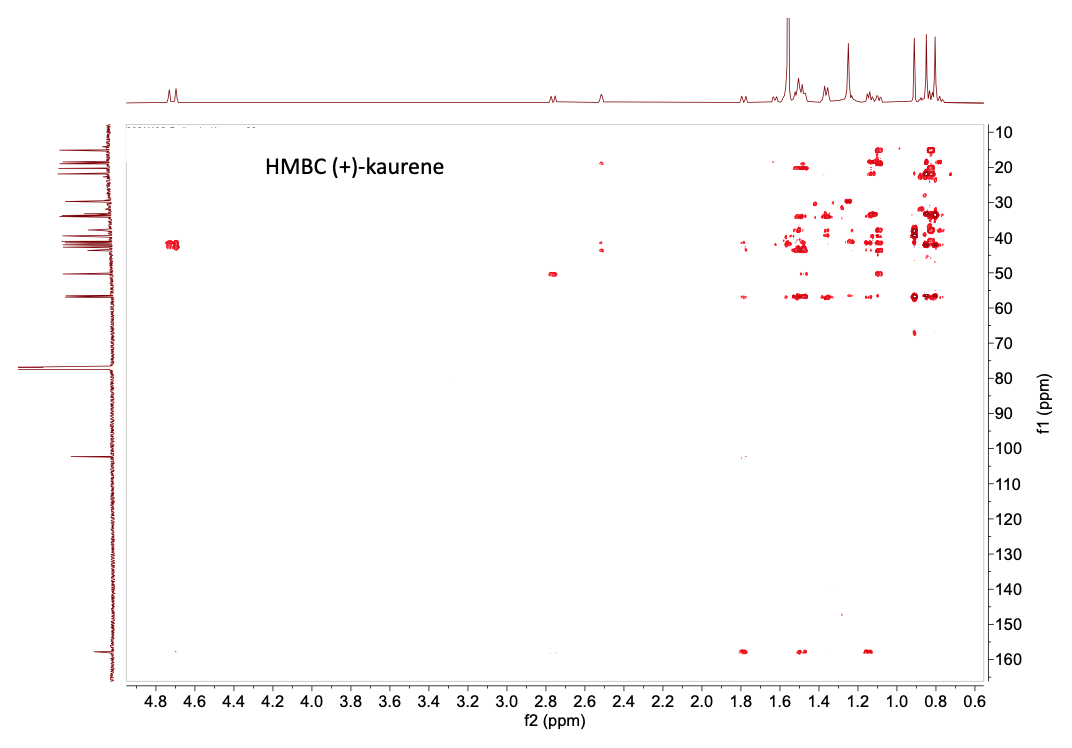


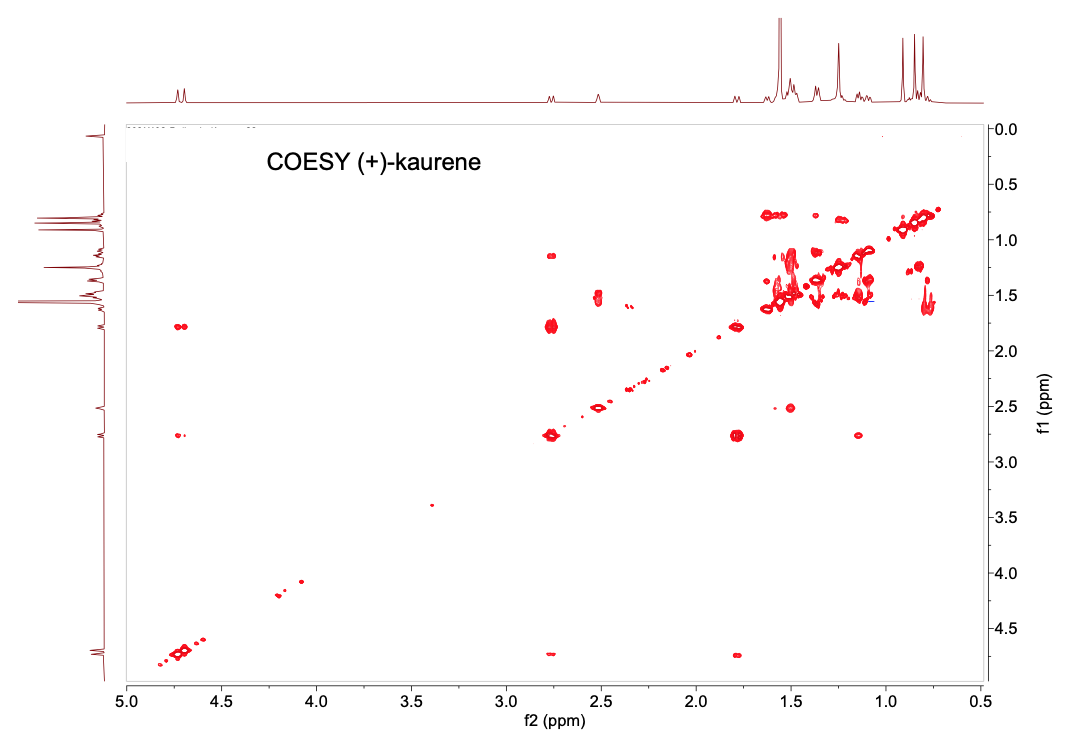
**Supplemental Fig. 4.** **NMR analysis confirms that 4 is (+)-kaurene.** Connectivity was deduced from 1 H, 13C, HSQC, HMBC, and COESY correlations. CDCl3 was used as the solvent, and CDCl3 peaks were referenced to 7.26 and 77.00 ppm for 1H and 13C spectra, respectively. Relative stereochemistry was assigned based on NOESY correlations. Selected NOESY correlations are shown on a 3D representation of (+)-kaurene generated with the Cahn–Ingold–Prelog CIP Tool 22. Absolute stereochemistry was assigned based on the configuration of (+)-CPP, the precursor of **4**. Continued on following pages.

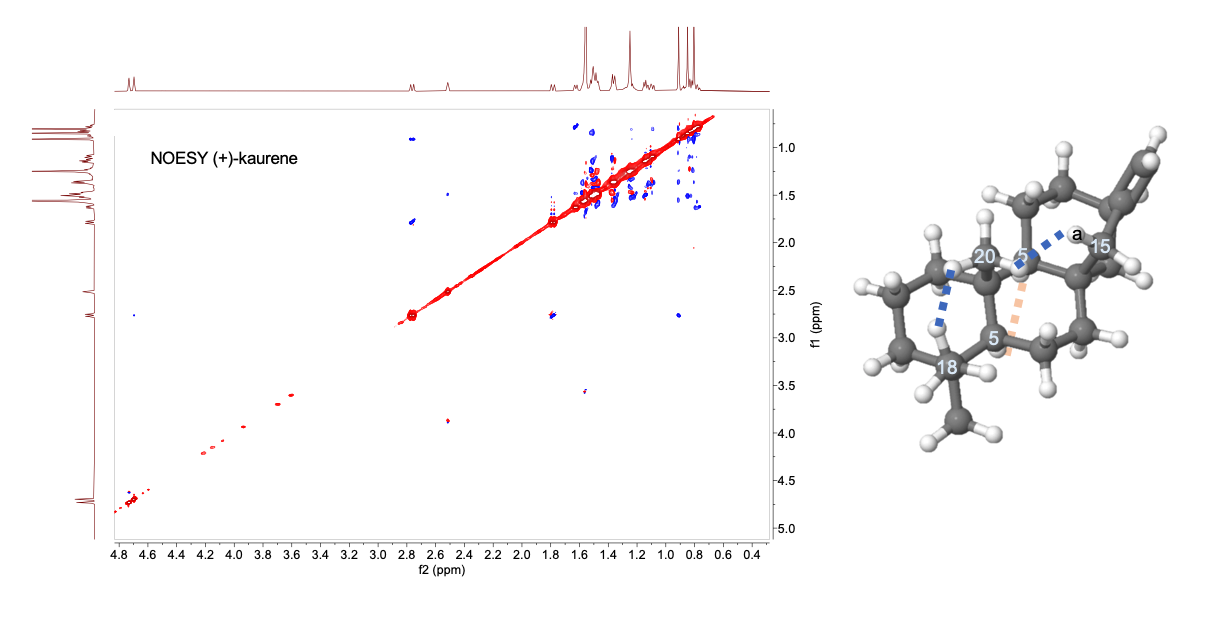


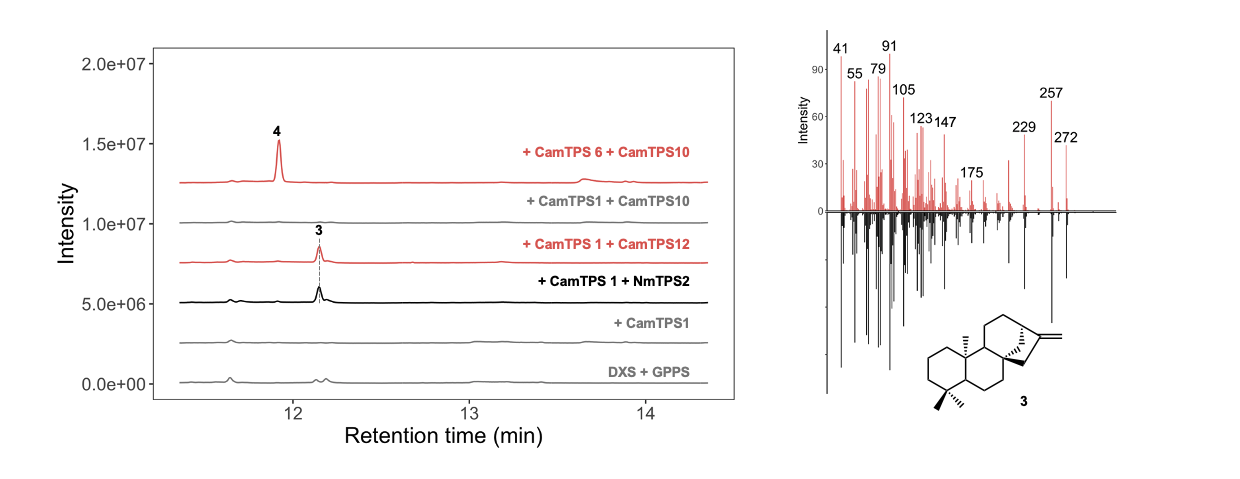




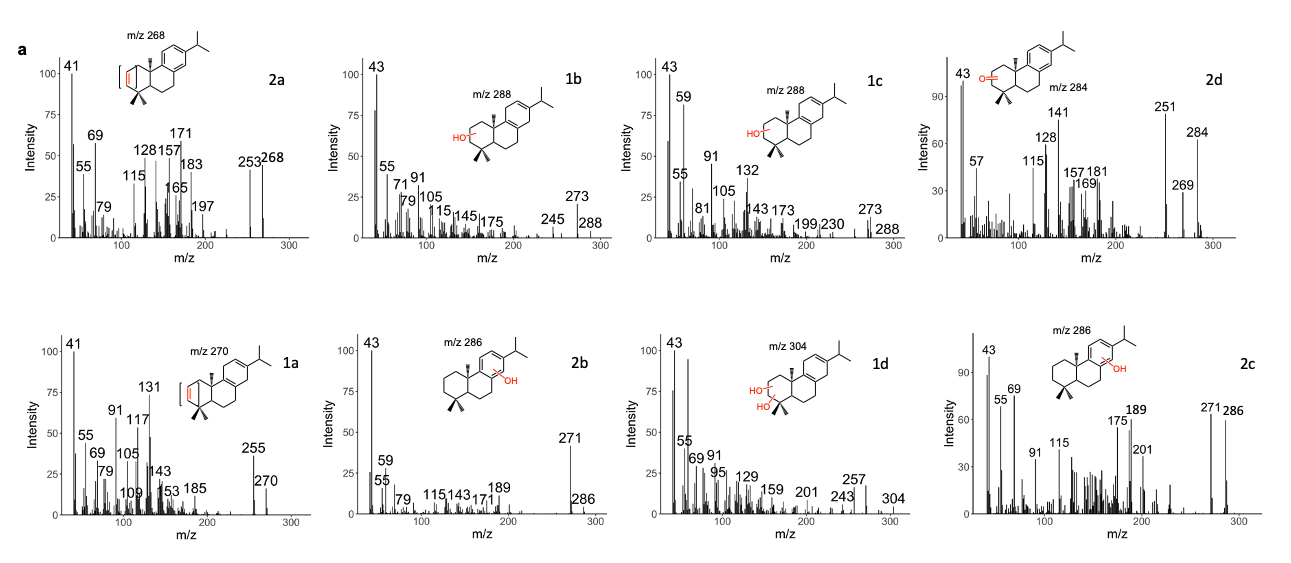








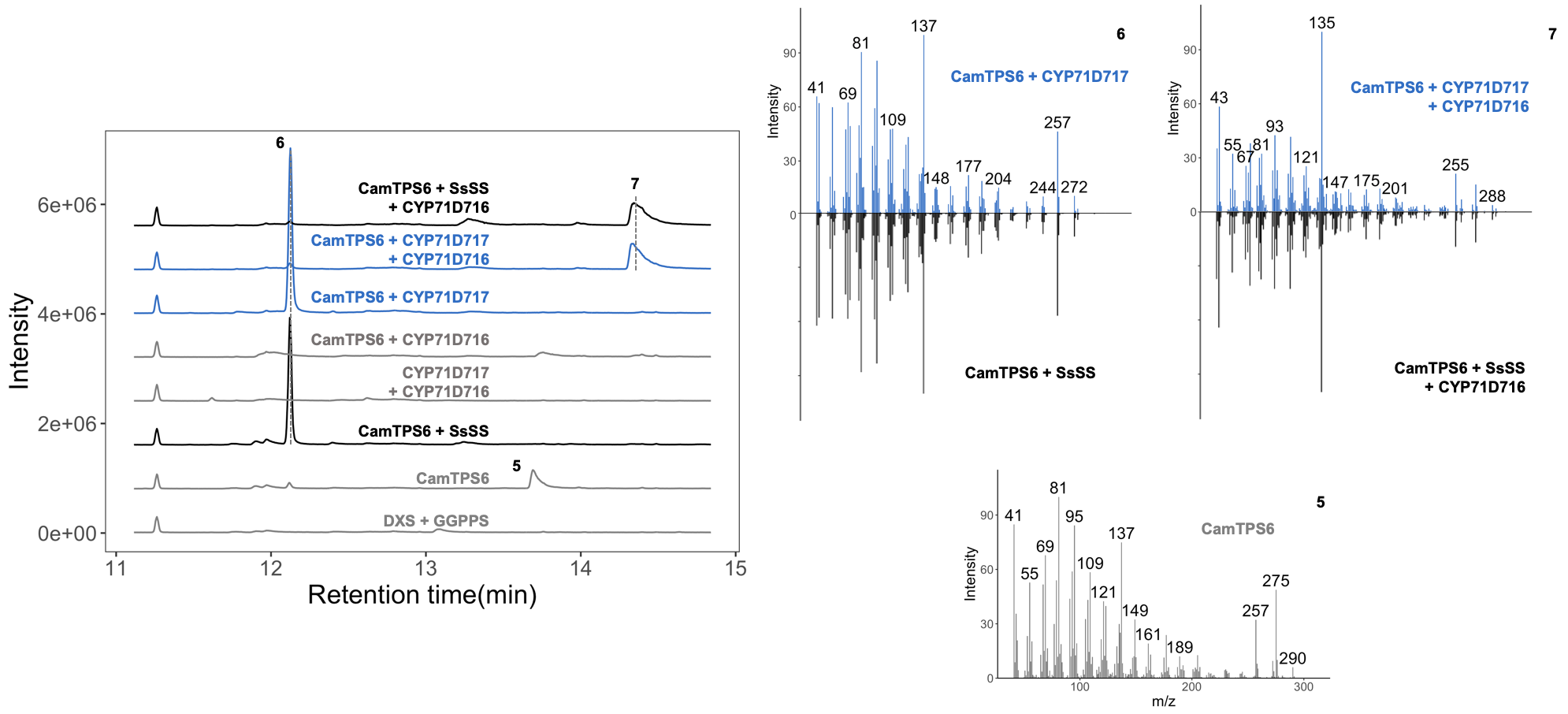
**Supplemental Fig. 5.** **Characterization of CamTPS12 as an *ent*-kaurene synthase by comparison with the product of NmTPS223 when combined with the *ent*-CPP synthase CamTPS11.**



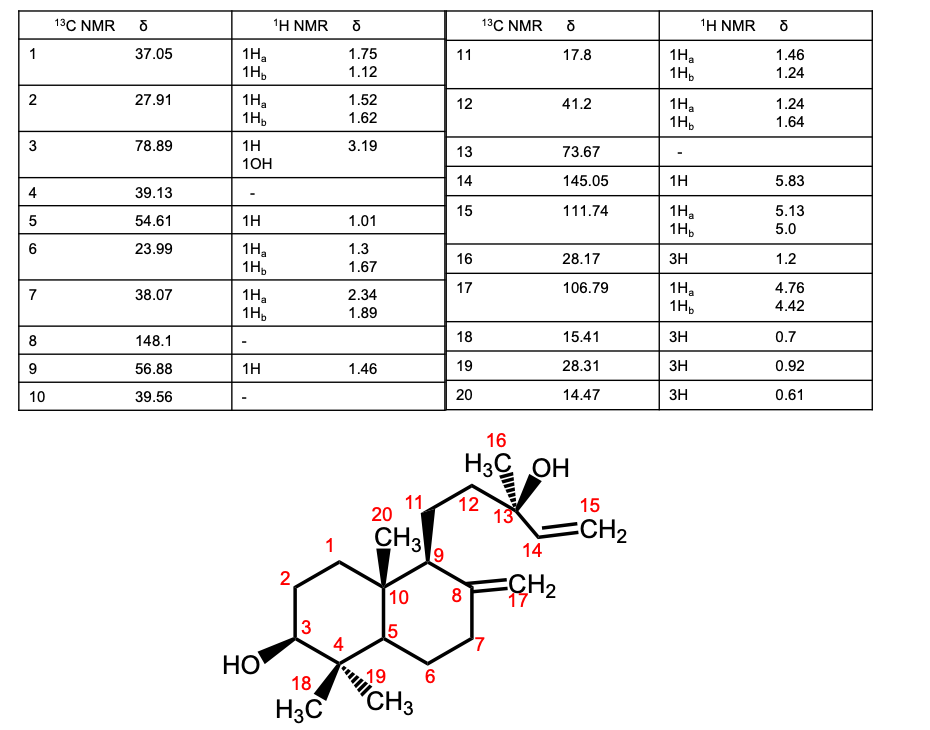
Chart, histogram

Description automatically generated

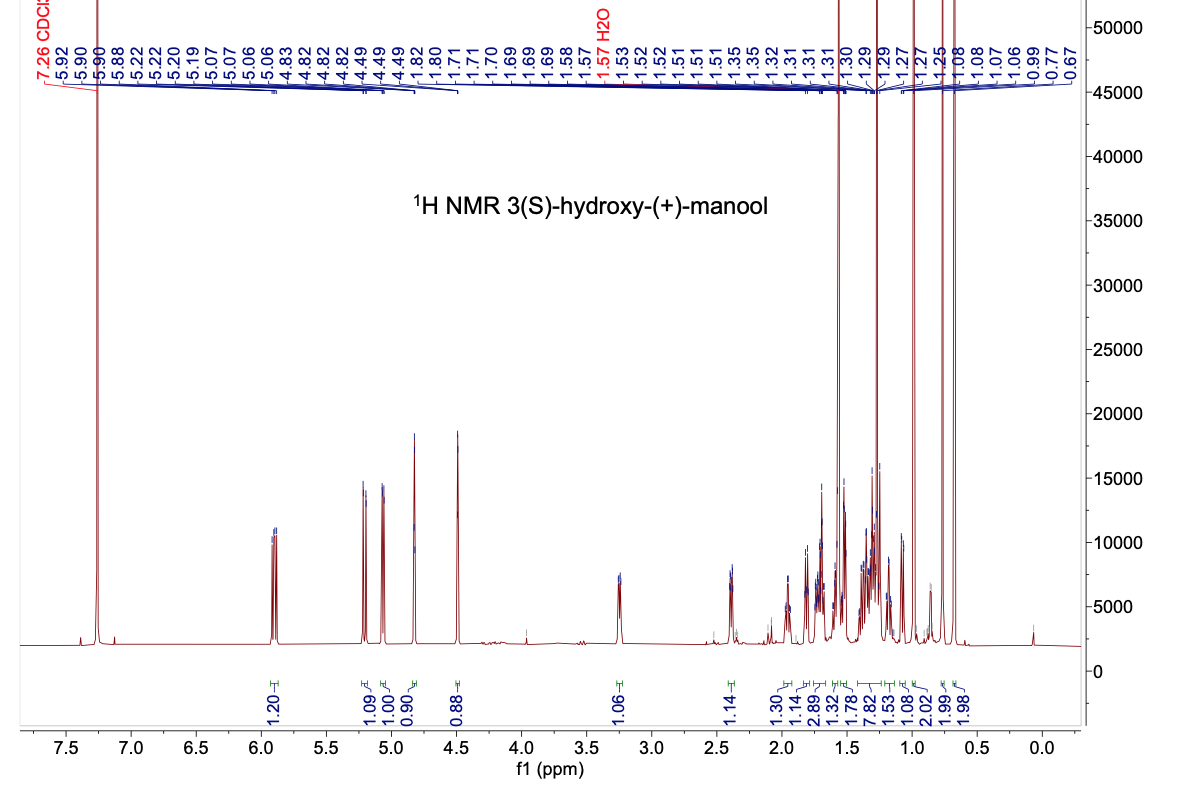
**Supplemental Fig. 6. Mass spectra of the CYP76AH oxidation products (a) and information for relevant NIST database hits (b**). Analysis of the mass spectra led to tentative assignment of 1a and 2a as miltiradiene and abietatriene respectively with an additional desaturation, possibly due to water loss, based on observed loss of 2 m/z compared with the parent spectrum. 1b and 1c display molecular ions consistent with a single oxidation of miltiradiene while 1d is apparently a doubly hydroxylated analog. 2b and 2c have characteristic m/z values of the aromatic abietatriene backbone (high intensity 271 and 286) and are good matches to single hydroxylation on the c-ring. 2d has a molecular ion that would correspond to keto-abietatriene and shows a similar spectrum to abietatrien-3-one from the NIST database.

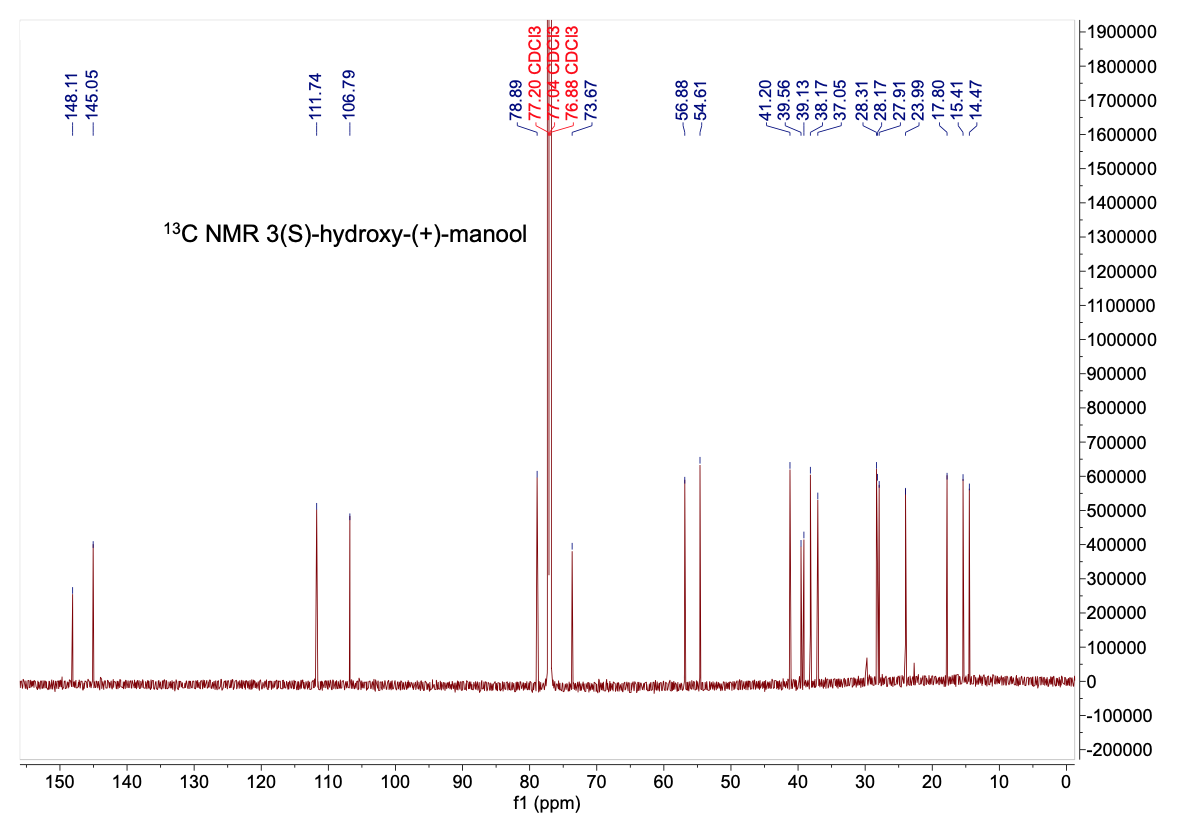


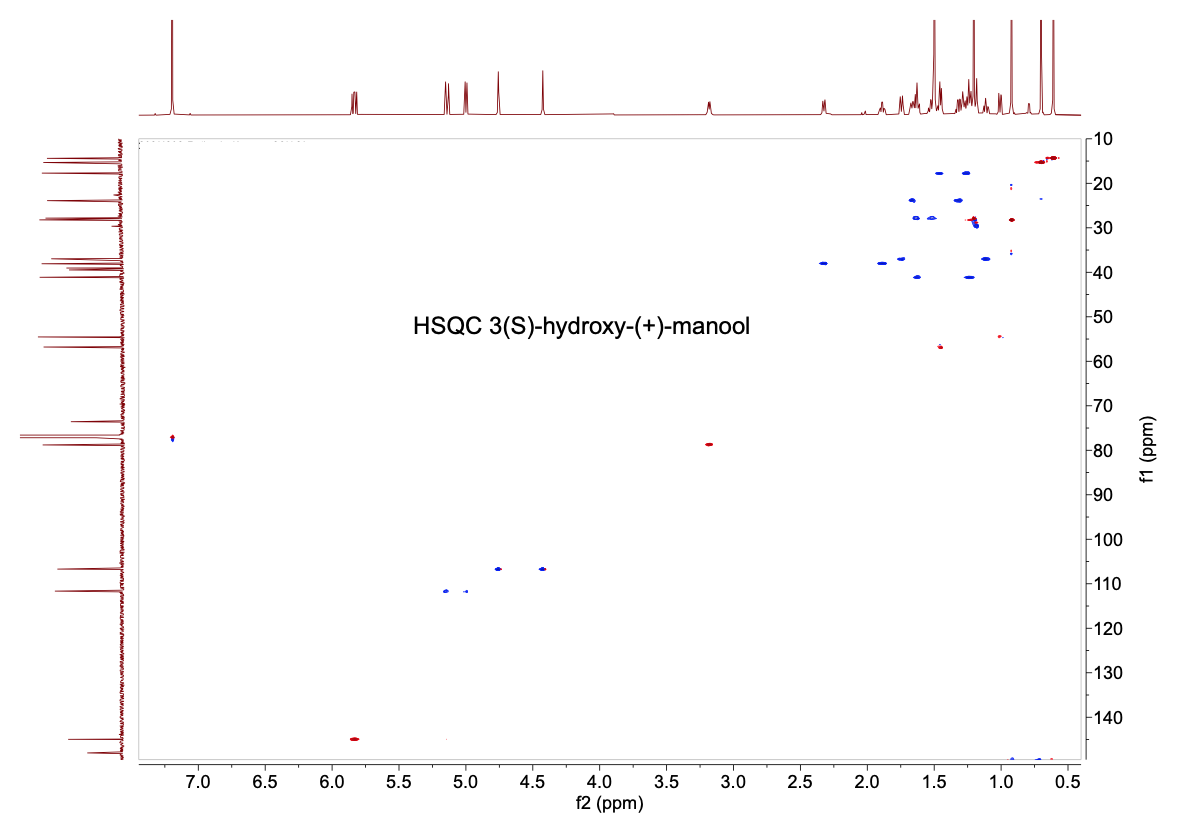
**Supplemental Fig. 7.** **Further analysis of CamCYP71D717 shows that it unexpectedly functions as a (+)-manool synthase when expressed in *N. benthamiana*.** We compared retention times and mass spectra of products when SsSS, a promiscuous class I TPS known to afford (+)-manool from (+)-CPP24, was substituted for CamCYP71D17. Based on recent work demonstrating CYP oxidation of kolavenol but not kolavenyl diphosophate25, we suggest that CamCYP71D717 is catalyzing an isomerization of (+)-copalol **(6)** which is produced by endogenous phosphatases in *N. benthamiana* when a (+)-CPP synthase is expressed.

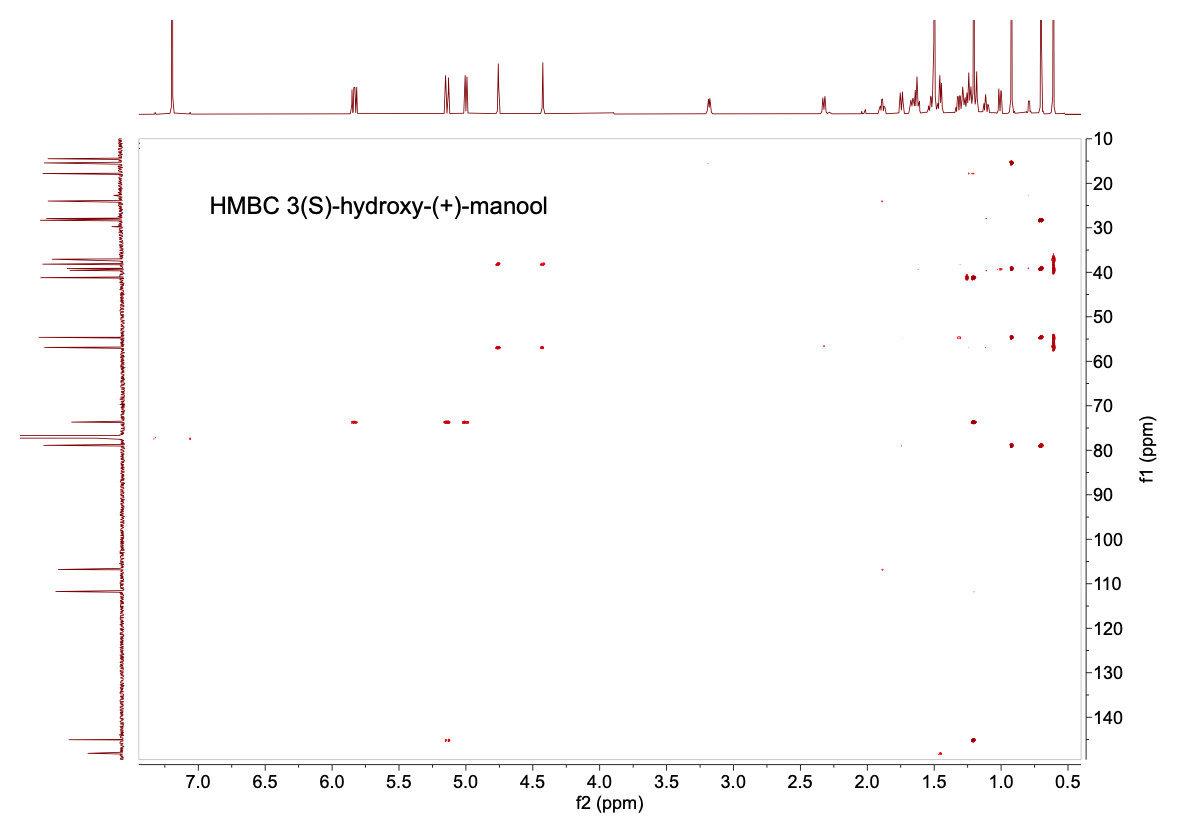


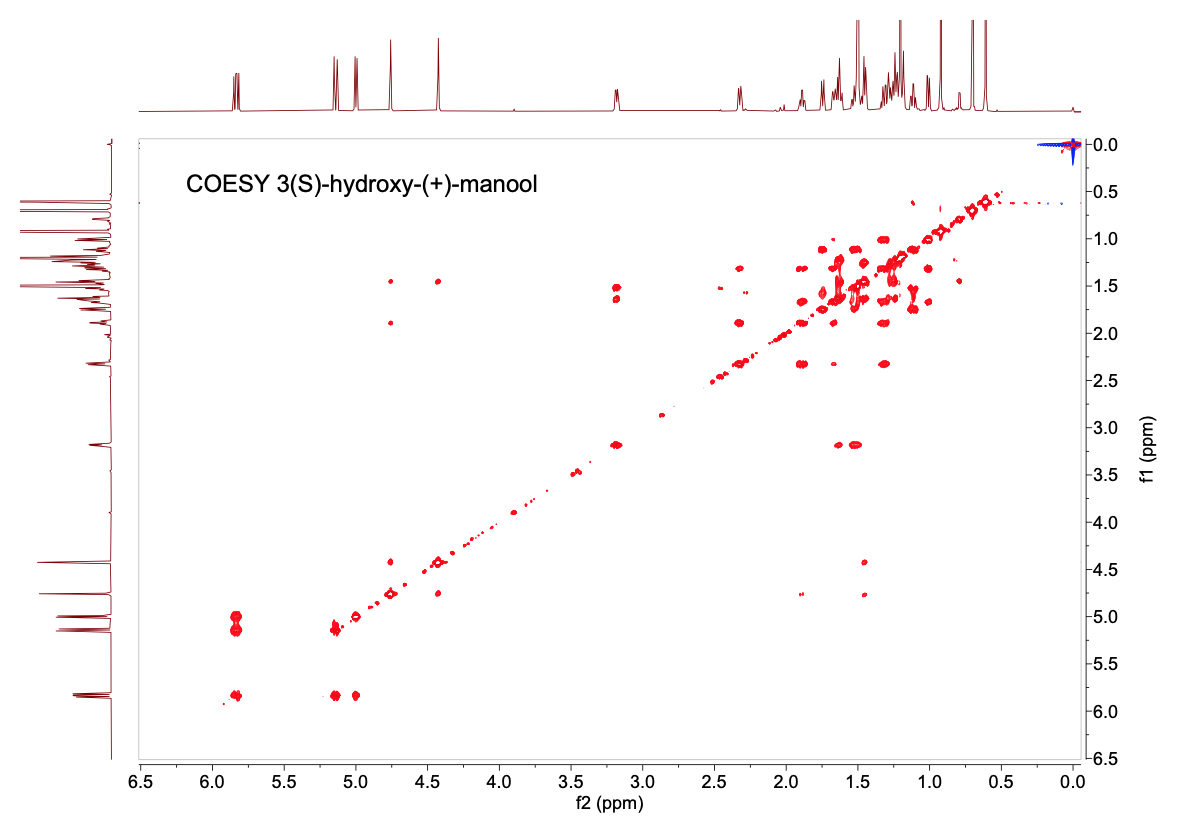
**Supplemental Fig 8. NMR analysis of 3(S)-hydroxy-(+)-manool (7).** Connectivity was deduced from 1 H, 13C, HSQC, HMBC, and COESY correlations. CDCl3 was used as the solvent, and CDCl3 peaks were referenced to 7.26 and 77.00 ppm for 1H and 13C spectra, respectively. Selected NOESY correlations are shown to support assignment of relative stereochemistry. Absolute stereochemistry was assigned based on the configuration of (+)-CPP, the precursor of **7**. Continued on the following pages.

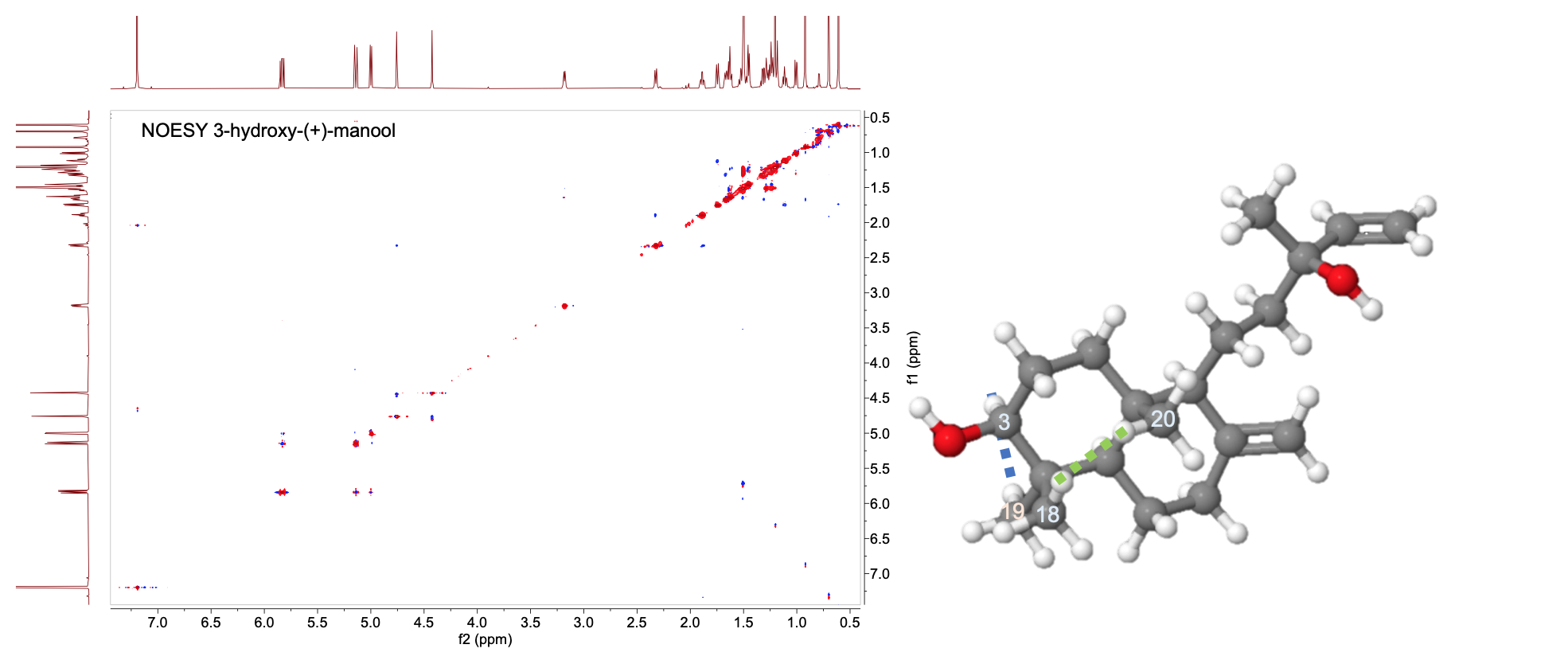


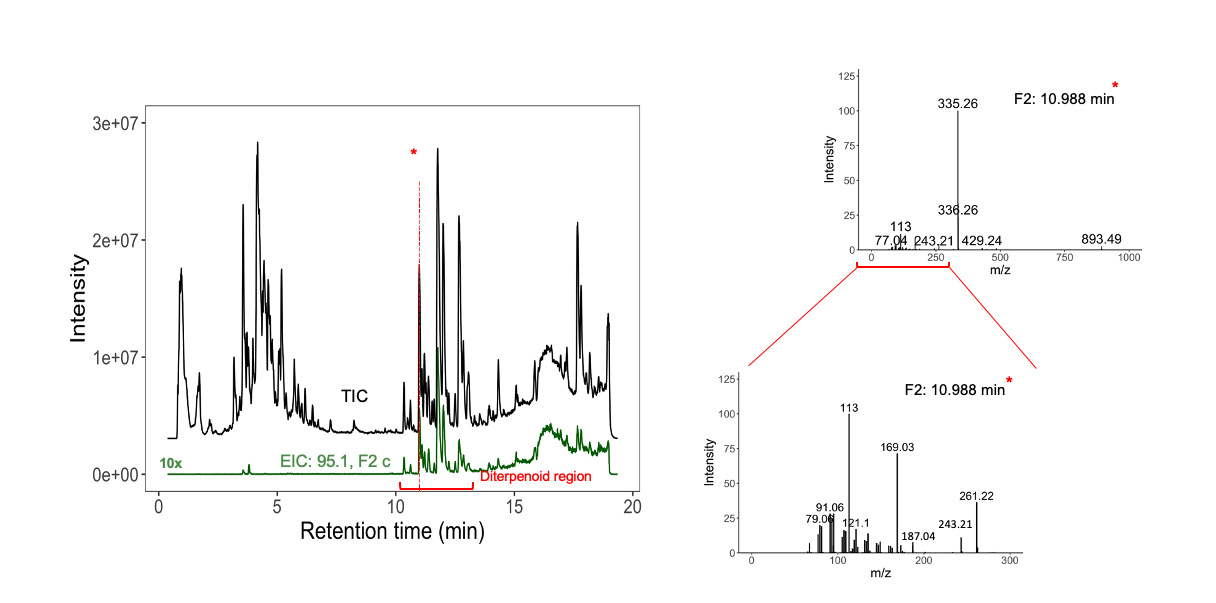




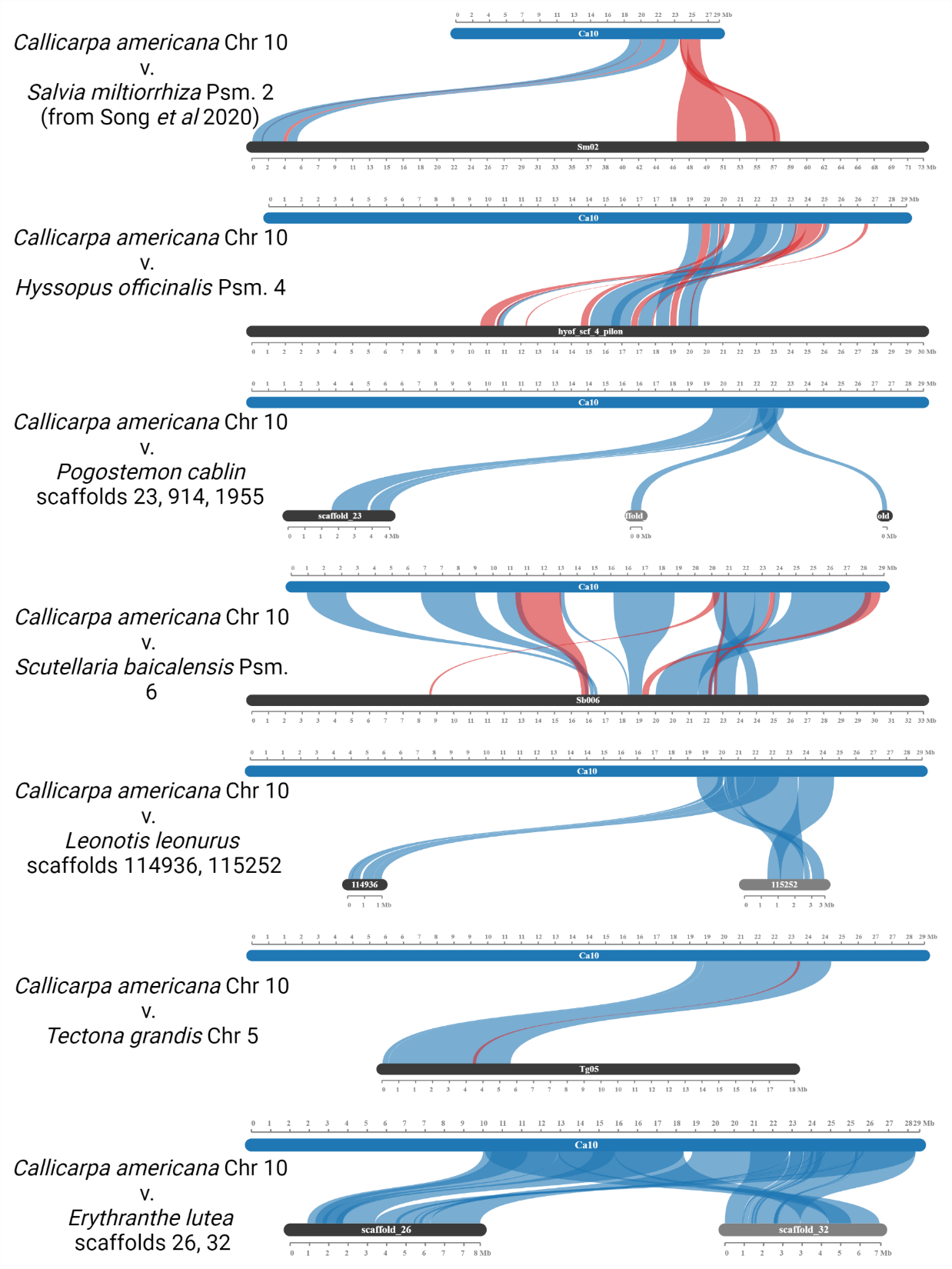








**Supplemental Fig. 9. LC-MS/MS analysis of a methanol root extract from *C. americana*.** The total ion chromatogram compared with an extracted ion chromatogram of m/z 95.1 (function 2 (F2), MS/MS), a characteristic fragment ion of diterpenes, show that most diterpenoids elute around 10-12 minutes with this method. Inspection of the mass spectrum of F2, the MSe spectrum, shows a characteristic diterpenoid fragmentation pattern.



**Supplemental Fig. 10. Synteny of chromosomes and scaffolds containing the miltiradiene BGC.** This cluster ranges from approximately 21.92-22.33 Mb on the *Callicarpa americana* chromosome 10, represented here as the top chromosome in each comparison. Blue lines represent collinear blocks, red lines represent inverted collinear blocks. Analysis was performed using MCScanX software and visualized using SynVisio.

SUPPLEMENTAL REFERENCES

1. Hamilton, J. P. *et al.* Generation of a chromosome-scale genome assembly of the insect-repellent terpenoid-producing Lamiaceae species, *Callicarpa americana*. *GigaScience* **9**, giaa093 (2020).

2. Lichman, B. R., Godden, G. T. & Buell, C. R. Gene and genome duplications in the evolution of chemodiversity: perspectives from studies of Lamiaceae. *Curr. Opin. Plant Biol.* **55**, 74–83 (2020).

3. Malli, R. P. N., Adal, A. M., Sarker, L. S., Liang, P. & Mahmoud, S. S. De novo sequencing of the *Lavandula angustifolia* genome reveals highly duplicated and optimized features for essential oil production. *Planta* **249**, 251–256 (2019).

4. Li, J. *et al.* The chromosome-based lavender genome provides new insights into Lamiaceae evolution and terpenoid biosynthesis. *Hortic. Res.* **8**, 1–14 (2021).

5. Vining, K. J. *et al.* Draft genome sequence of *Mentha longifolia* and development of resources for mint cultivar improvement. *Mol. Plant* **10**, 323–339 (2017).

6. Bornowski, N. *et al.* Genome sequencing of four culinary herbs reveals terpenoid genes underlying chemodiversity in the Nepetoideae. *DNA Res.* **27**, dsaa016 (2020).

7. Rastogi, S. *et al.* Unravelling the genome of Holy basil: an “incomparable” “elixir of life” of traditional Indian medicine. *BMC Genomics* **16**, 413 (2015).

8. Upadhyay, A. K. *et al.* Genome sequencing of herb Tulsi (*Ocimum tenuiflorum*) unravels key genes behind its strong medicinal properties. *BMC Plant Biol.* **15**, 212 (2015).

9. Zhang, Y. *et al.* Incipient diploidization of the medicinal plant *Perilla* within 10,000 years. *Nat. Commun.* **12**, 5508 (2021).

10. He, Y. *et al.* Building an octaploid genome and transcriptome of the medicinal plant *Pogostemon cablin* from Lamiales. *Sci. Data* **5**, 180274 (2018).

11. He, Y. *et al.* Survey of the genome of *Pogostemon cablin* provides insights into its evolutionary history and sesquiterpenoid biosynthesis. *Sci. Rep.* **6**, 26405 (2016).

12. Xu, H. *et al.* Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. *Mol. Plant* **9**, 949–952 (2016).

13. Zhang, G. *et al.* Hybrid de novo genome assembly of the Chinese herbal plant danshen (*Salvia miltiorrhiza* Bunge). *GigaScience* **4**, s13742-015-0104–3 (2015).

14. Song, Z. *et al.* A high-quality reference genome sequence of *Salvia miltiorrhiza* provides insights into tanshinone synthesis in its red rhizomes. *Plant Genome* **13**, e20041 (2020).

15. Ma, Y. *et al.* Expansion within the CYP71D subfamily drives the heterocyclization of tanshinones synthesis in *Salvia miltiorrhiza*. *Nat. Commun.* **12**, 685 (2021).

16. Dong, A.-X. *et al.* High-quality assembly of the reference genome for scarlet sage, *Salvia splendens*, an economically important ornamental plant. *GigaScience* **7**, giy068 (2018).

17. Jia, K.-H. *et al.* Chromosome-scale assembly and evolution of the tetraploid *Salvia splendens* (Lamiaceae) genome. *Hortic. Res.* **8**, 1–15 (2021).

18. Zhao, Q. *et al.* The reference genome sequence of *Scutellaria baicalensis* provides insights into the evolution of wogonin biosynthesis. *Mol. Plant* **12**, 935–950 (2019).

19. Xu, Z. *et al.* Comparative genome analysis of *Scutellaria baicalensis* and *Scutellaria barbata* reveals the evolution of active flavonoid biosynthesis. *Genomics Proteomics Bioinformatics* **18**, 230–240 (2020).

20. Yasodha, R. *et al.* Draft genome of a high value tropical timber tree, Teak (*Tectona grandis* L. f): Insights into SSR diversity, phylogeny and conservation. *DNA Res.* **25**, 409–419 (2018).

21. Zhao, D. *et al.* A chromosomal-scale genome assembly of *Tectona grandis* reveals the importance of tandem gene duplication and enables discovery of genes in natural product biosynthetic pathways. *GigaScience* **8**, (2019).

22. Mistry, N., Singh, R. & Ridley, J. A web-based stereochemistry tool to improve students’ ability to draw newman projections and chair conformations and assign R/S labels. *J. Chem. Educ.* **97**, 1157–1161 (2020).

23. Johnson, S. R. *et al.* A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae). *J. Biol. Chem.* **294**, 1349–1362 (2019).

24. Caniard, A. *et al.* 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 (2012).

25. Kwon, M. *et al.* Cytochrome P450-catalyzed biosynthesis of a dihydrofuran neoclerodane in magic mint ( *Salvia divinorum* ). *ACS Catal.* **12**, 777–782 (2022).