Plant Communications, Volume *4*

Supplemental information

Comparative genomics reveals the diversification of triterpenoid biosynthesis and origin of ocotillol-type triterpenes in *Panax*

Zijiang Yang, Xiaobo Li, Ling Yang, Sufang Peng, Wanling Song, Yuan Lin, Guisheng Xiang, Ying Li, Shuang Ye, Chunhua Ma, Jianhua Miao, Guanghui Zhang, Wei Chen, Shengchao Yang, and Yang Dong

Zijiang Yang, Xiaobo Li, Ling Yang, Wanling Song, Yuan Lin, Guisheng Xiang, Ying Li, Shuang Ye, Chunhua Ma, Jianhua Miao, Guanghui Zhang, Wei Chen, Shengchao Yang, Yang Dong.

National & Local Joint Engineering Research Center on Germplasm Innovation & Utilization of Chinese Medicinal Materials in Southwest China, Yunnan Agricultural University, Kunming, China

The Key Laboratory of Medicinal Plant Biology of Yunnan Province, Yunnan Agricultural University, Kunming, China

College of Food Science and Technology, Yunnan Agricultural University, Kunming, China

Guangxi Key Laboratory of Medicinal Resources Protection and Genetic Improvement, Guangxi Botanical Garden of Medicinal Plants, Nanning, China

Yunnan Plateau Characteristic Agriculture Industry Research Institute, Kunming, China

Comparative genomics reveals the diversification of triterpenoid biosynthesis and origin of ocotillol-type triterpenes in *Panax*

Supporting Information Methods S1. Characterization of compounds produced from enzymatic reactions

Gas chromatography-mass spectrometry (GC-MS) analysis

The purified yeast extract was derivatized by resuspending in 100 µl of trimethylsilyl cyanide (TMSCN) with a 1:1 ratio followed by incubation of 30 min at 65 °C. GC analysis was performed by Agilent 7890A with a HP-5MS quartz capillary column (30 m×0.25 mm i.d., 0.25 µm film thickness). The temperature was set as 250 °C for injector port, source, and transfer line. The column temperature was programmed as follows: 80 °C for 2 min; increase to 290 °C at a rate of 20 °C min^{-1}; hold at 290 °C for 30 min. The flow rate of carrier gas helium was 1.2 ml min^{-1}. Samples were injected in splitless mode with either a 1-µl or a 3-µl volume. MS analysis was performed using Agilent 6540 Accurate-Mass Q-TOF system.

Triterpenoid standards preparation

δ-amyrin, β-amyrin, α-amyrin, cycloartenol, ψ-taraxasterol, taraxasterol, and dammarendiol-II were purchased from Chengdu DeSiTe Biological Technology Co. Ltd, China; 3-epicabraleadiol was purchased from BioBioPha Co. Ltd, China. Standards were first dissolved in hexane, followed by derivatization using TMSCN.

Nuclear Magnetic Resonance (NMR) analysis

The purified yeast extract was subjected to column chromatography (CC) on silica gel (200-300 mesh, Qingdao Marine Chemical Factory, China) eluting with petroleum ether and then with petroleum ether/ethyl acetate stepwise-gradient system (from 13:1 to 5:1, v/v) to obtain four fractions (denoted as Fr.1–Fr.4). Fr.4 was purified by semi-preparative high-performance liquid chromatography (HPLC) on Agilent 1290 Infinity II system with off-line monitoring by thin-layer chromatography (TLC). The column used for HPLC was a reversed-phase column (Agilent ZORBAX StableBond SB-C18, 9.4 × 250 mm, 5 μm). The setting for mobile phase was 100% acetonitrile at a flow rate of 3 ml min^{-1}. TLC analysis was carried out on silica gel plates (GF254F, 10 - 40 μm, Qingdao Marine Chemical Factory) by spraying with 5% H2SO4 in EtOH (v/v) followed by heating to 120 °C for 5 min. The above process yielded compound **8** (3 mg, containing trace amount of compound **9**) and compound **9** (12 mg). The purified compound 8 and 9 was analyzed by ^{1}H-NMR and ^{13}C-NMR spectroscopy at 600 and 150 MHz in $CDCl_3$ solution using Bruker AV-600 MHz spectrometer.

Identification of compounds

Through GC analysis, the product profile for nine OSCs were identified (Table S24). The naming of compounds was in consistent with Figure 4C. Based on the GC retention times and mass spectral fragmentation patterns from existing literatures (Shan *et al.*, 2005; Salmon *et al.*, 2016; Kim *et al.*, 2018), the compounds were identified as follows: compound **1**: δ-amyrin; compound **2**: β-amyrin; compound **3**: α-amyrin; compound **5**: ψ-taraxasterol, compound **6**: taraxasterol; compound **7**: dammarendiol-II; compound **8**: 3-epicabraleadiol; compound **9**: ocotillol (Figures S19-S22).

Since authentic standard for ocotillol is not available. The NMR analysis was further performed to characterize the compound **8** and compound **9**. By comparison of NMR and mass spectroscopic data with previous study (Shan *et al.*, 2005), C-24S or C-24R epimers of the epoxydammaranes can be distinguished by the ^{1}H-NMR chemical shifts of H-24 and ^{13}C-NMR chemical shifts of C-24, C-25, C-21, C-23 and C-22 positions. Chemical shifts and coupling constants at H-24 vary remarkably for molecules with locally diastereomeric configurations at C-20 and C-24 (Figures S24-S26, Table S23).

References for Methods S1

Shan, H., Segura, M.J., Wilson, W.K., Lodeiro, S., Matsuda, S.P. (2005). Enzymatic cyclization of dioxidosqualene to heterocyclic triterpenes. J. Am. Chem. Soc. **127**: 18008–18009.

Salmon, M., Thimmappa, R.B., Minto, R.E., Melton, R.E., Hughes, R.K., O'Maille, P.E., Hemmings, A.M., Osbourn, A. (2016). A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases. Proc. Natl. Acad. Sci. U.S.A. **113**: E4407–E4414.

Kim, O.T., Um, Y., Jin, M.L., Kim, J.U., Hegebarth, D., Busta, L., Racovita, R.C., Jetter, R. (2018). A Novel Multifunctional C-23 Oxidase, CYP714E19, is Involved in Asiaticoside Biosynthesis. Plant Cell Physiol. **59**: 1200–1213.

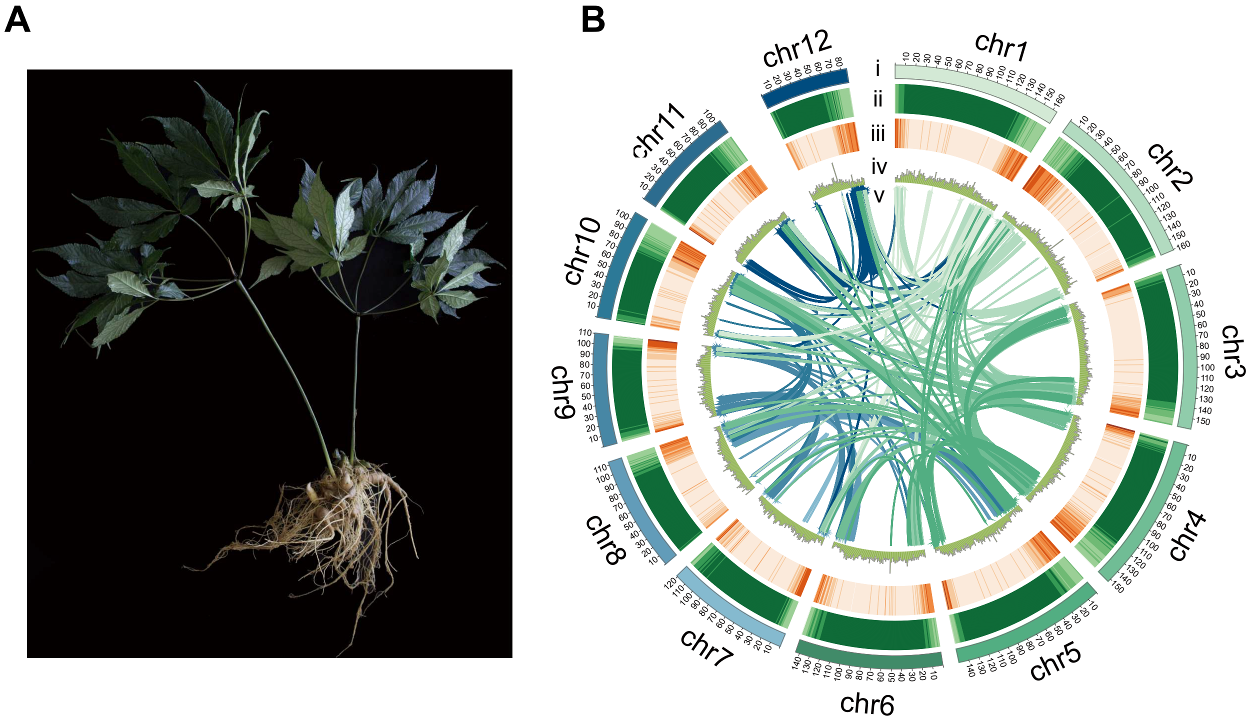
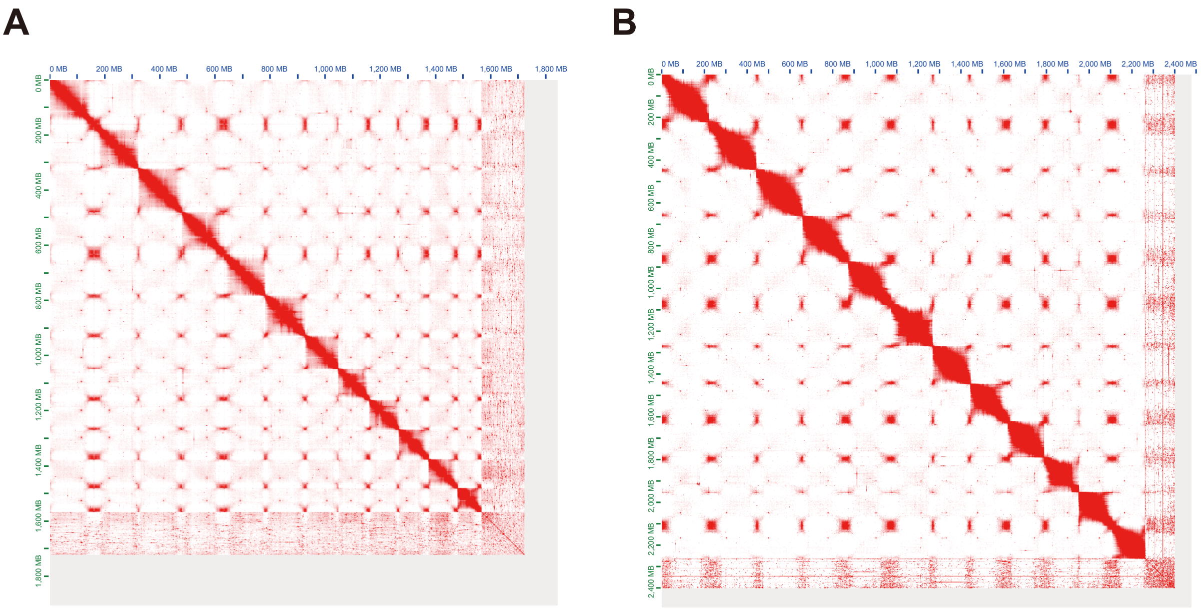


Figure S1 Morphology and genome of *P. vietnamensis* var. *fuscidiscus*. (A) Morphology of *P. vietnamensis* var. *fuscidiscus*. (B) Overview of *P. vietnamensis* var. *fuscidiscus* assembly. (I) chromosomes; (II) transposable elements density heatmap (1 Mb sliding window); (III) gene density heatmap (1 Mb sliding window); (IV) GC content (1 Mb sliding window); (V) collinear regions within *P. vietnamensis* var. *fuscidiscus* genome.



Panax vietnamensis var. *fuscidiscus*

Panax notoginseng

Figure S2 Hi-C contact heatmaps. (A) *P. vietnamensis* var. *fuscidiscus* assembly. (B) Updated *P. notoginseng* assembly.

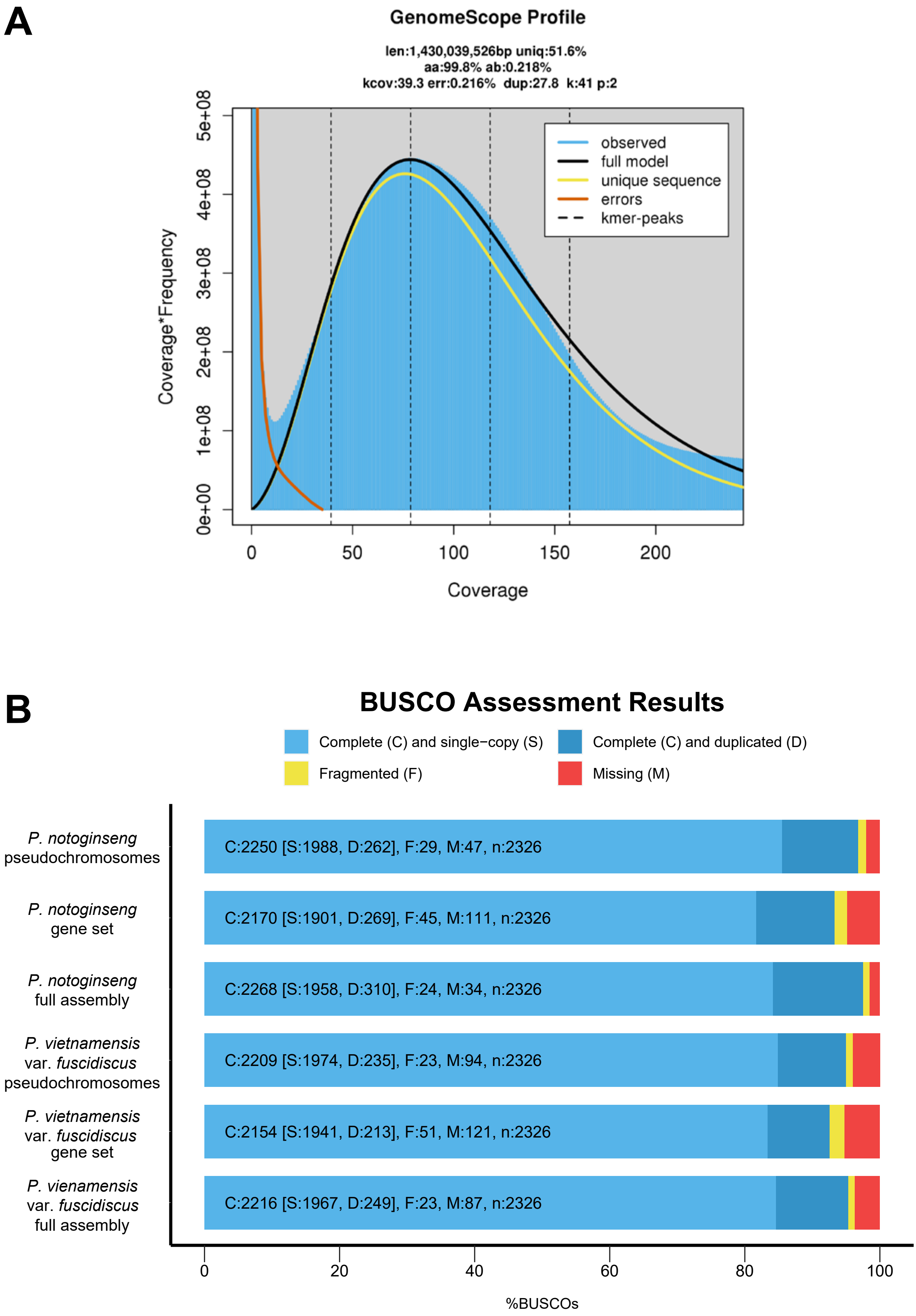


Figure S3 Genome survey and evaluation. (A) Survey of *P. vietnamensis* var. *fuscidiscus* genome by Genomescope. Ploidy and kmer length were set as 2 and 41, respectively. (B) BUSCO analysis of *P. vietnamensis* var. *fuscidiscus* and *P. notoginseng* assemblies and gene sets using the lineage dataset eudicots_odb10 (2,326 BUSCOs).

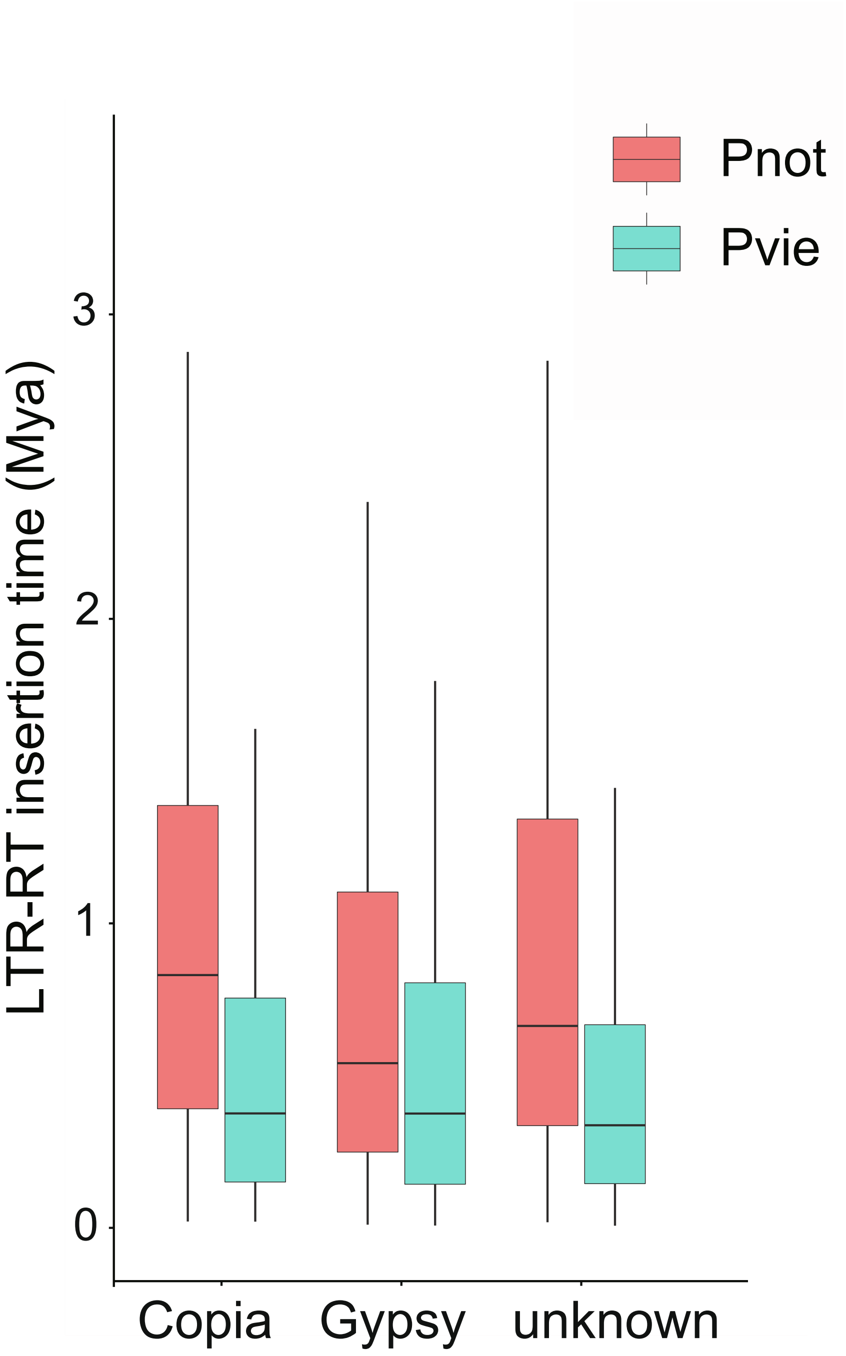


Figure S4 Estimated insertion time for Copia, Gypsy, and unknown type of LTRs in *P. vietnamensis* var. *fuscidiscus* and *P. notoginseng* genomes. Pvie: *P. vietnamensis* var. *fuscidiscus*, Pnot: *P. notoginseng*.

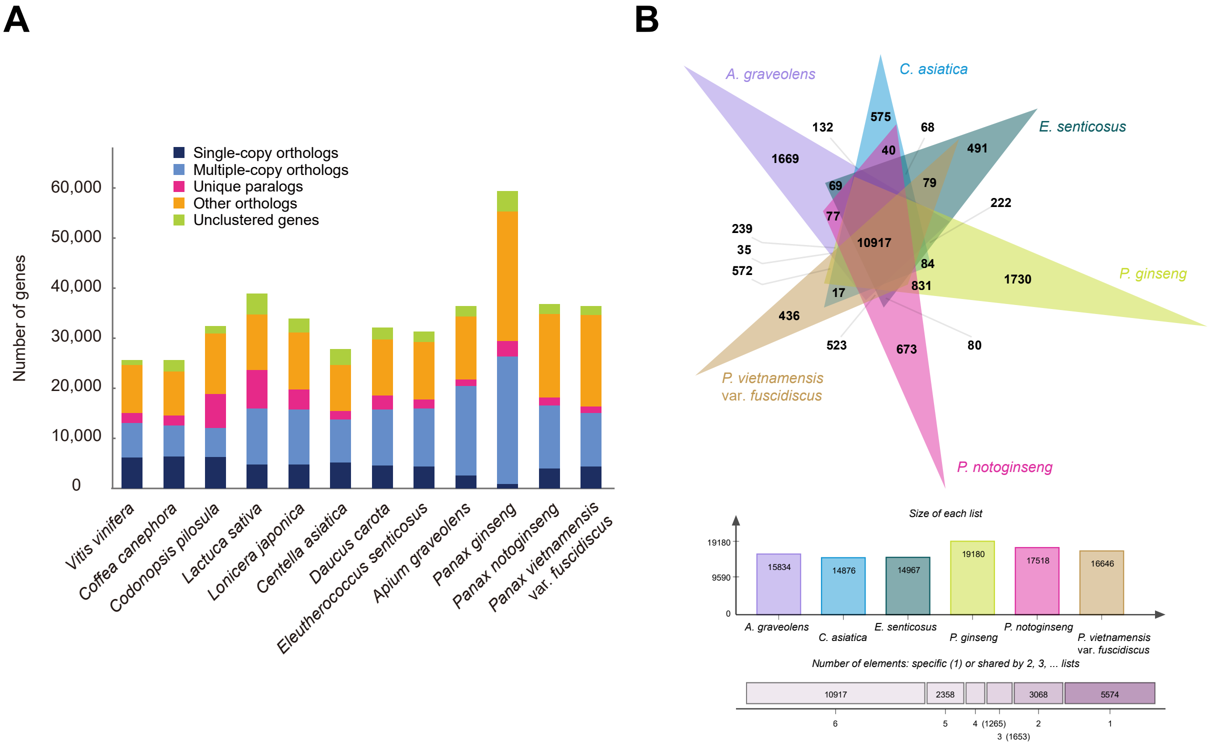


Figure S5 Orthology analyses. (A) Gene family analysis result for 12 eudicots. (B) Comparison of orthogroups from six Apiales species. Numbers in the upper venn diagram indicate the number of orthogroups. Numbers in the middle bar plot indicate the genes in the orthogroups for six species. Numbers in the bottom chart indicate number of shared or species-specific orthogroups.

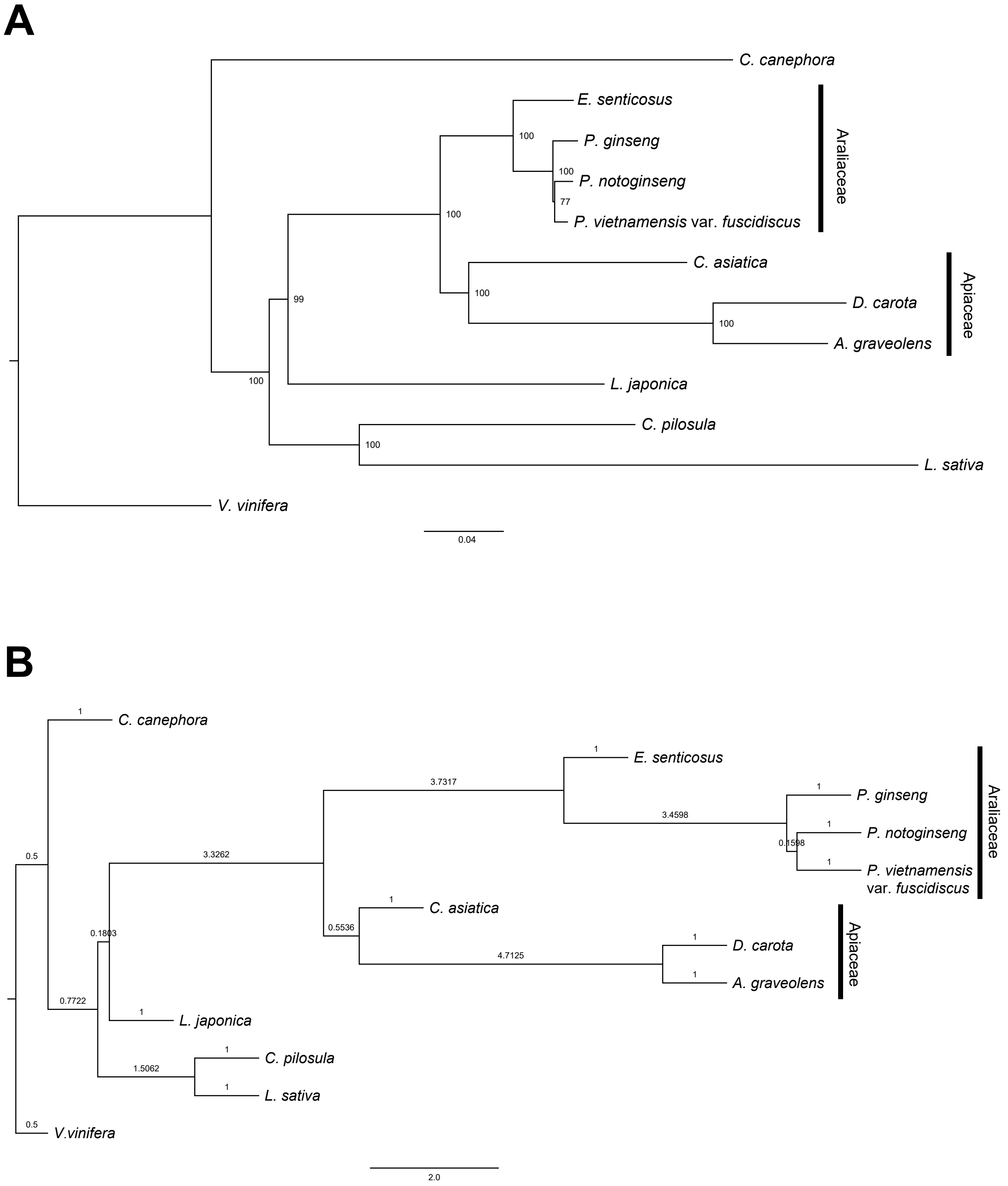


Figure S6 Phylogenetic analyses. (A) Species tree for 12 species inferred by concatenation-based method. Numbers indicate bootstrap values with 1,000 replicates. (B) Species tree for 12 species inferred by coalescence-based method. Branch lengths are shown in coalescent units. The numbers of each node represents the local posterior probabilities. Since the ASTRAL tree leaves the branch length of terminal branches empty, the length of terminal branches were all set as one.

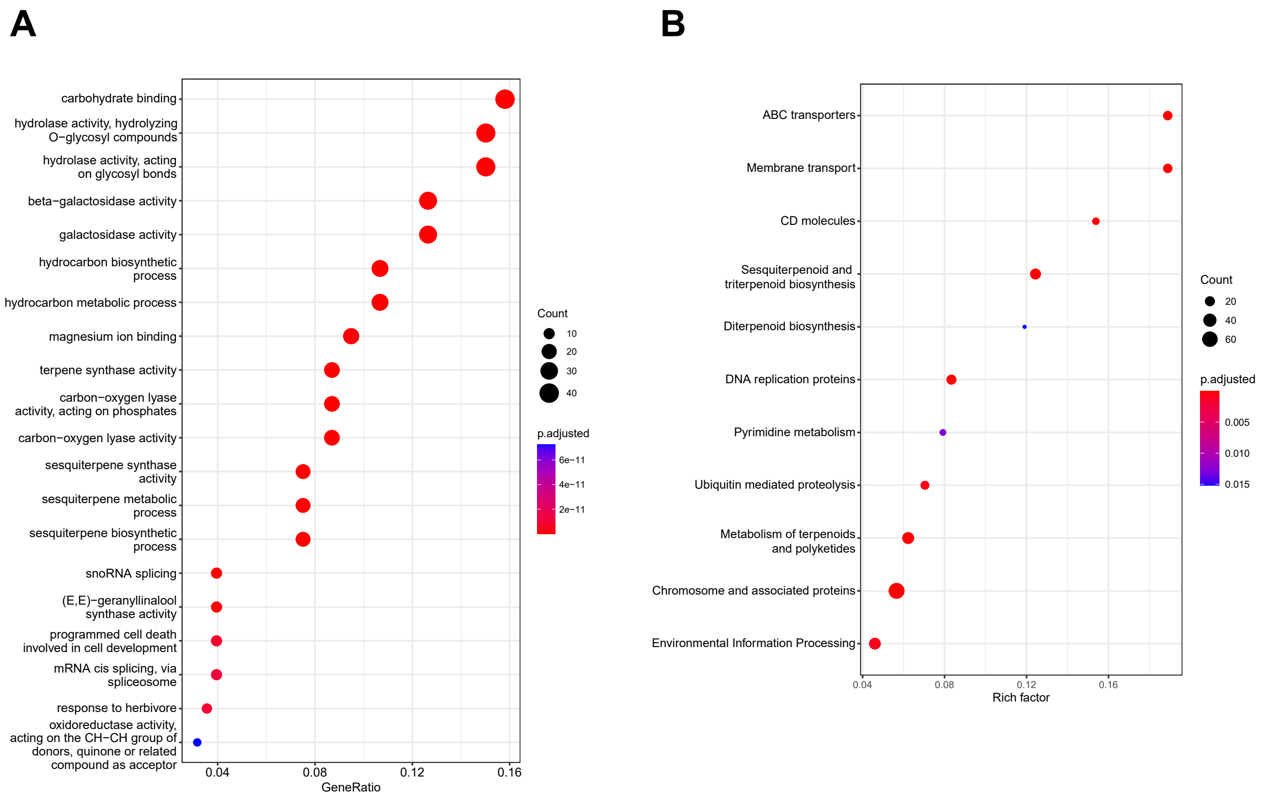


Figure S7 Functional enrichment analyses. (A) GO enrichment analysis of expanded gene families in *P. vietnamensis* var. *fuscidiscus*. (B) KEGG enrichment analysis of expanded gene families in *P. vietnamensis* var. *fuscidiscus*.

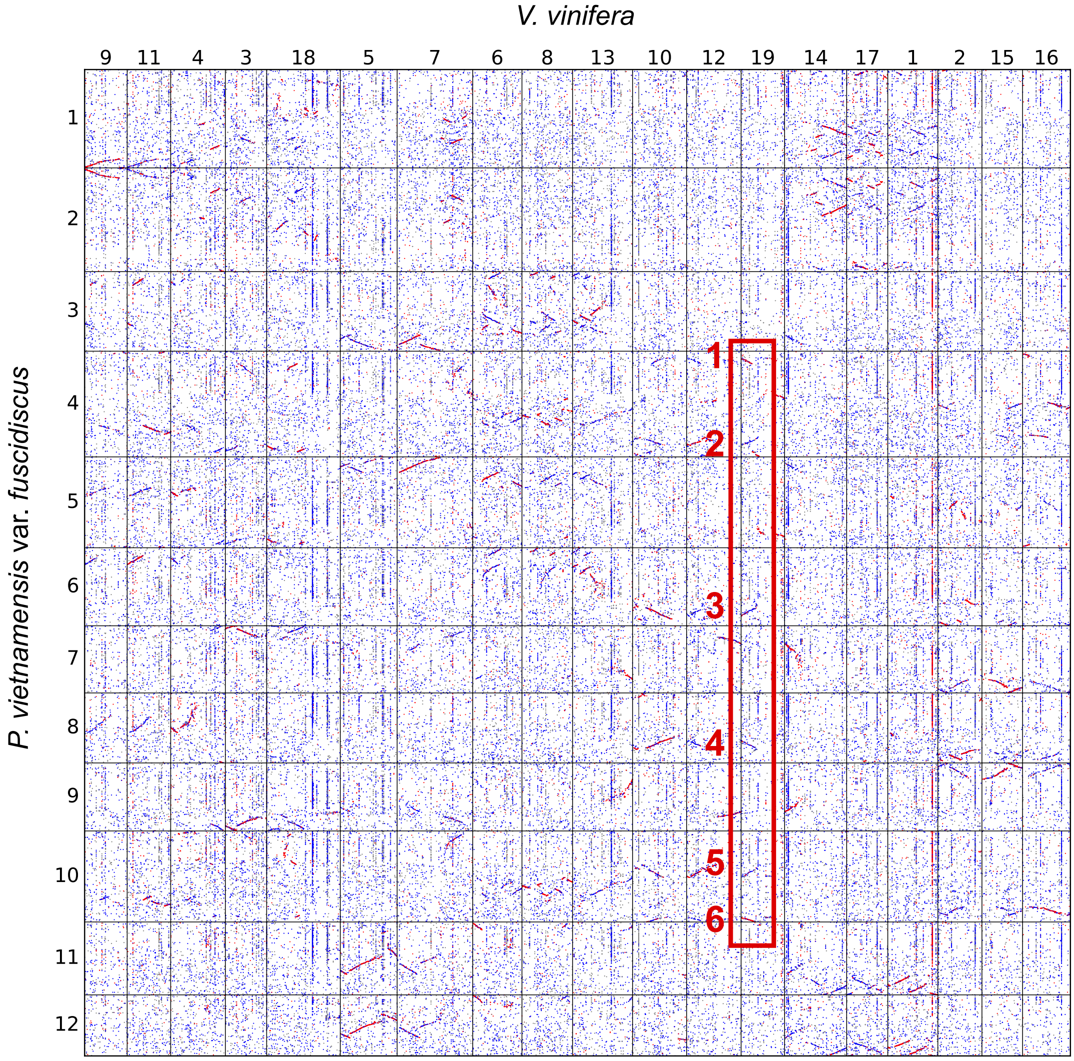


Figure S8 Synteny dot plot between *P. vietnamensis* var. *fuscidiscus* and *V. vinifera*. The red box highlighted regions between *V. vinifera* and *P. vietnamensis* var. *fuscidiscus* with a ratio of 1:6.

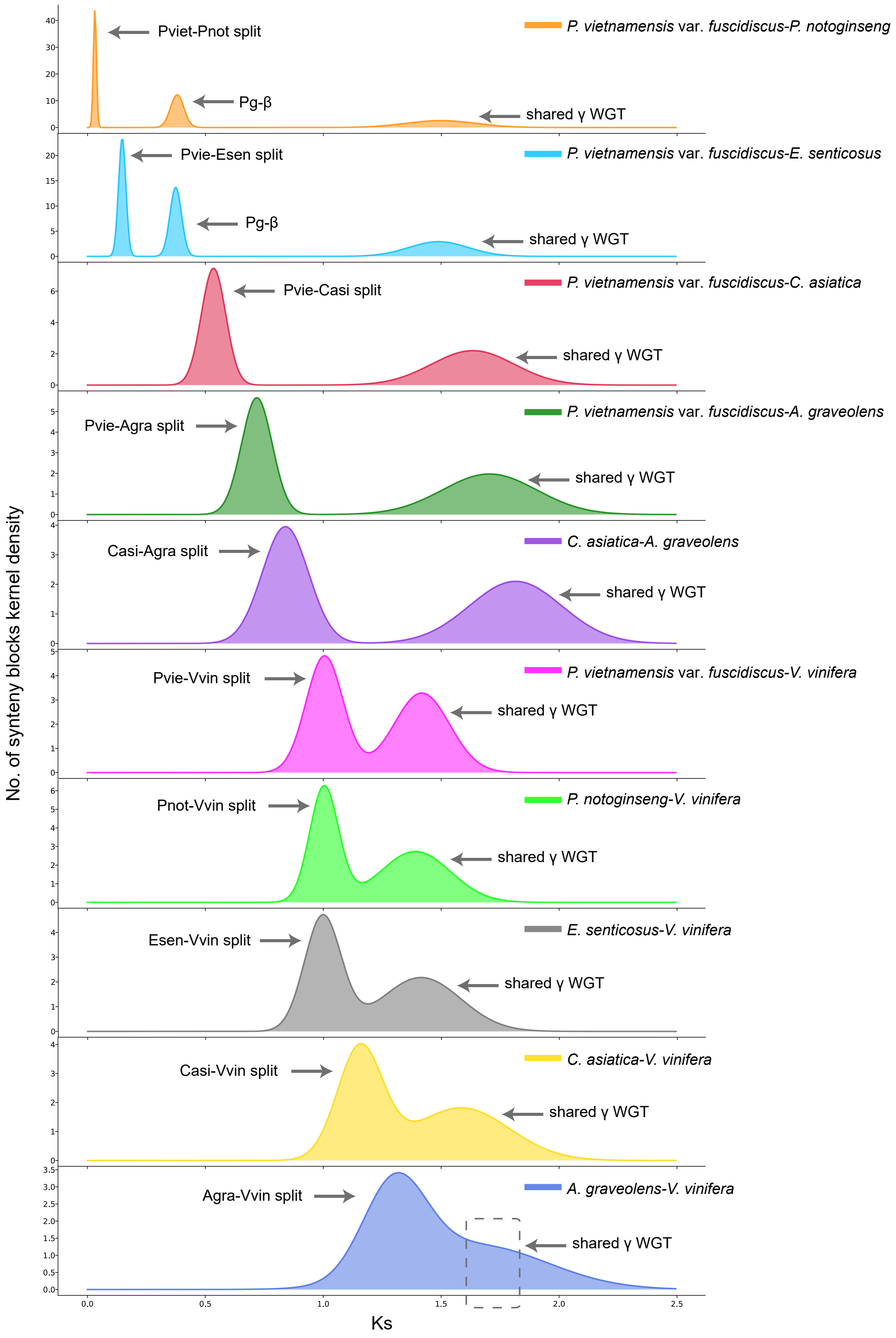


Figure S9 *Ks* distribution of interspecific collinear blocks. *Ks* peaks of speciation and shared polyploidizations between the studies species were labeled. Pvie: *P. vietnamensis* var. *fuscidiscus*, Pnot: *P. notoginseng*, Esen: *E. senticosus*, Casi: *C. asiatica*, Agra: *A. graveolens*, Vvin: *V. vinifera*.

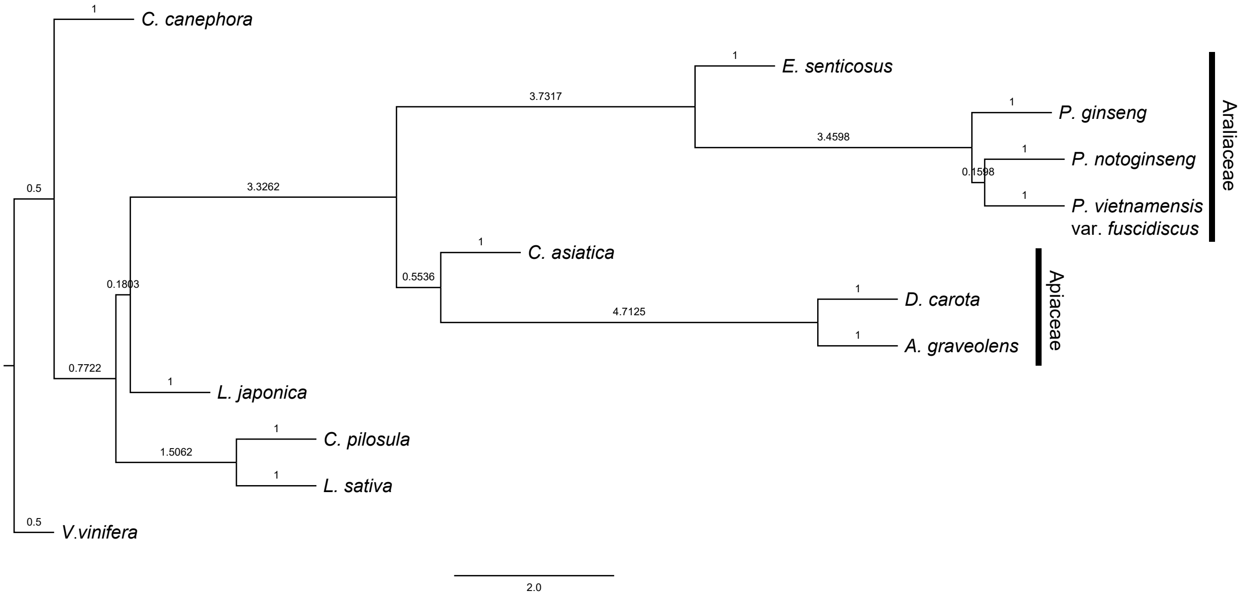


Figure S10 Synteny analyses. (A) Synteny dot plot between *P. notoginseng* and *E. senticosus*. *P. notoginseng* genome version PBJ-2021 (Yang *et al.*, 2021a). For *P. notoginseng*, number 13-18 represent Scaffold 13-18. (B) Synteny dot plot between *P. notoginseng* (updated by this study) and *E. senticosus*.

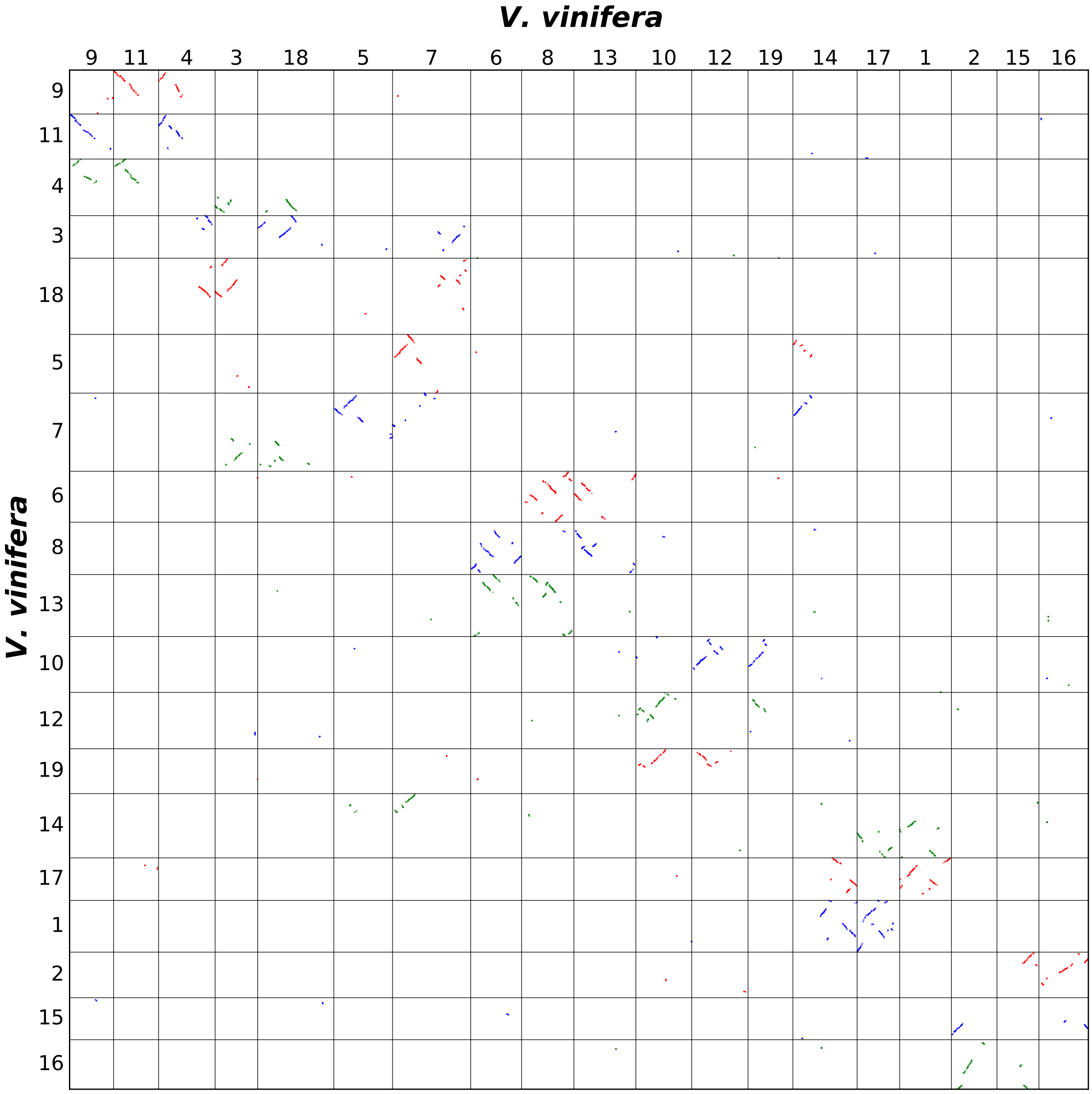


Figure S11 Collinear gene extraction of *V. vinifera*. The three candidate subsets resulted from polyploidizations were highlighted using red (S1), blue (S2), green (S3).

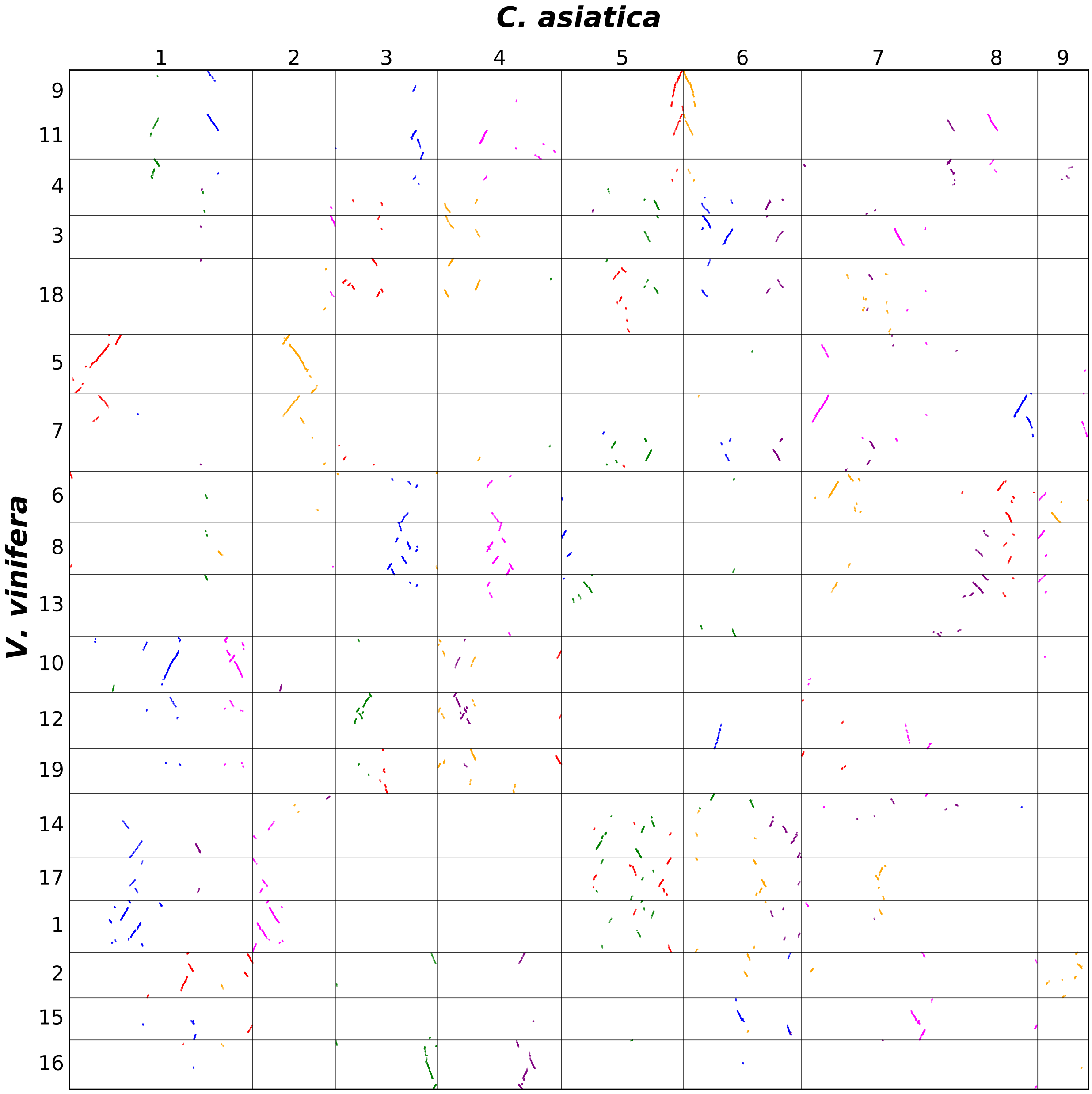


Figure S12 Collinear gene extraction between *V. vinifera* and *C. asiatica*. The six candidate subsets resulted from polyploidizations were highlighted using red (S1), blue (S2), green (S3), orange (S4), fuchsia (S5), purple (S6).

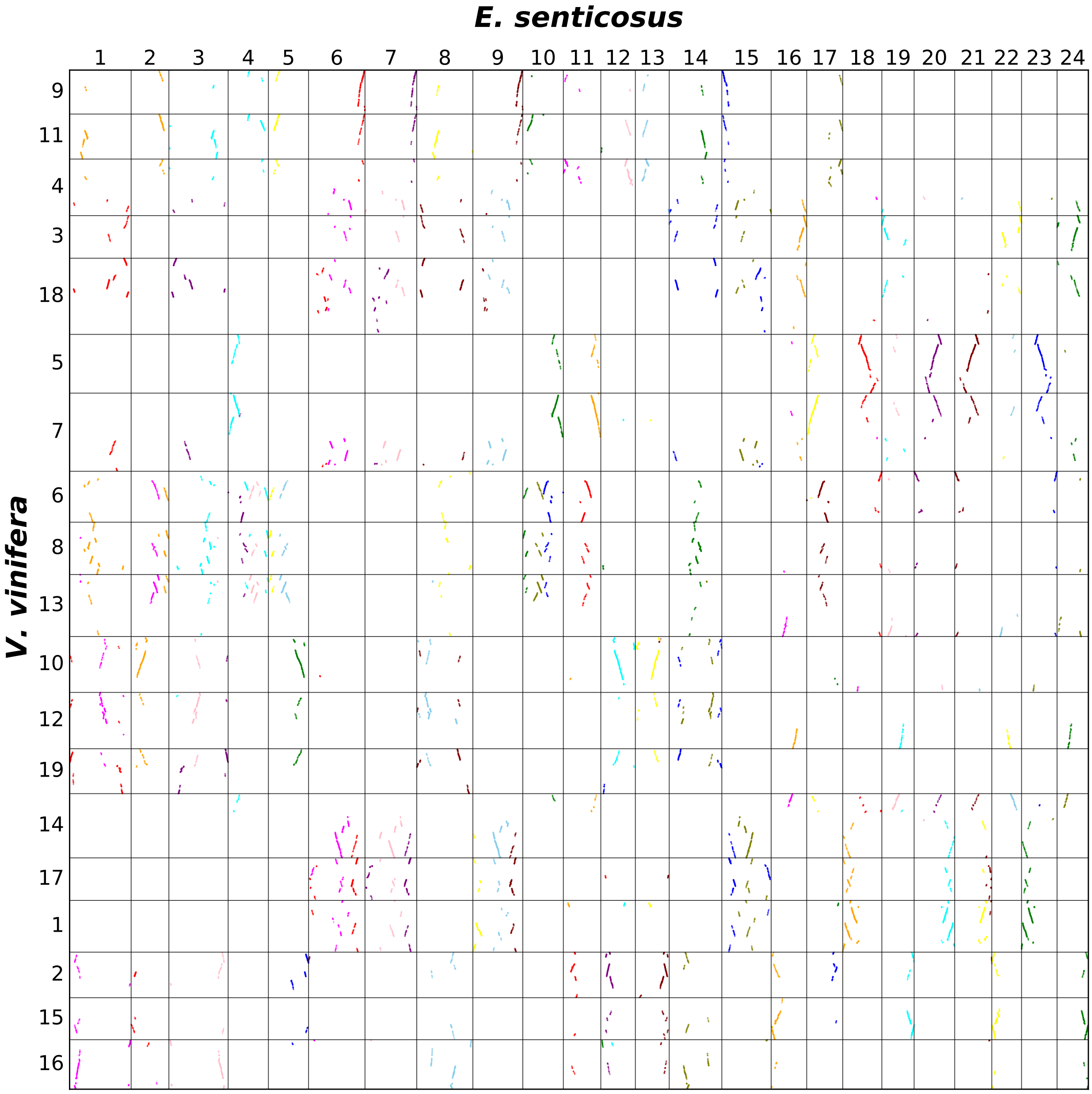


Figure S13 Collinear gene extraction between *V. vinifera* and *E. senticosus*. The 12 candidate subsets resulted from polyploidizations were highlighted using red (S1), blue (S2), green (S3), olive (S4), orange (S5), fuchsia (S6), purple (S7), cyan (S8), pink (S9), maroon (S10), yellow (S11), skyblue (S12).

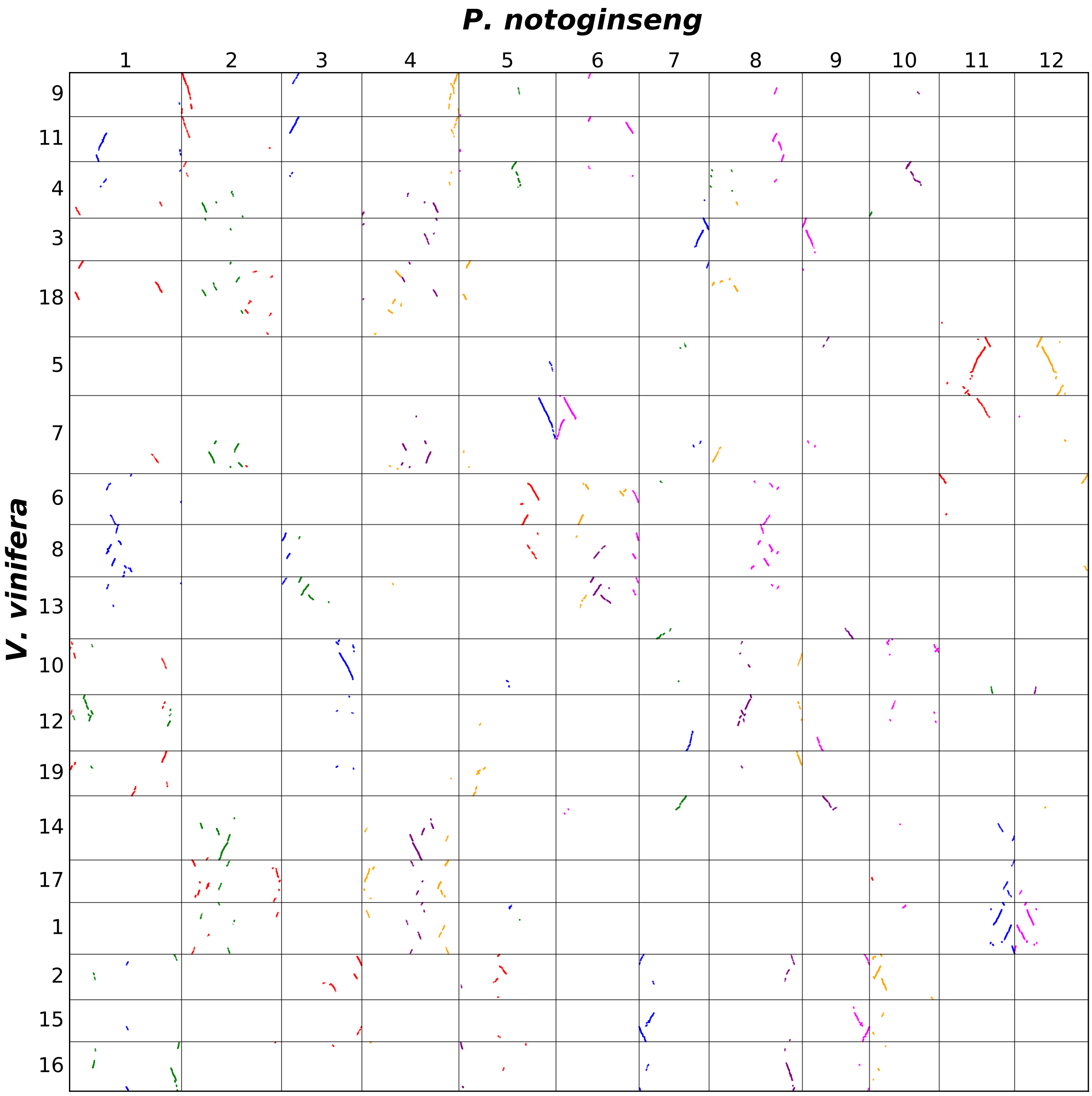


Figure S14 Collinear gene extraction between *V. vinifera* and *P. notoginseng*. The six candidate subsets resulted from polyploidizations were highlighted using red (S1), blue (S2), green (S3), orange (S4), fuchsia (S5), purple (S6).

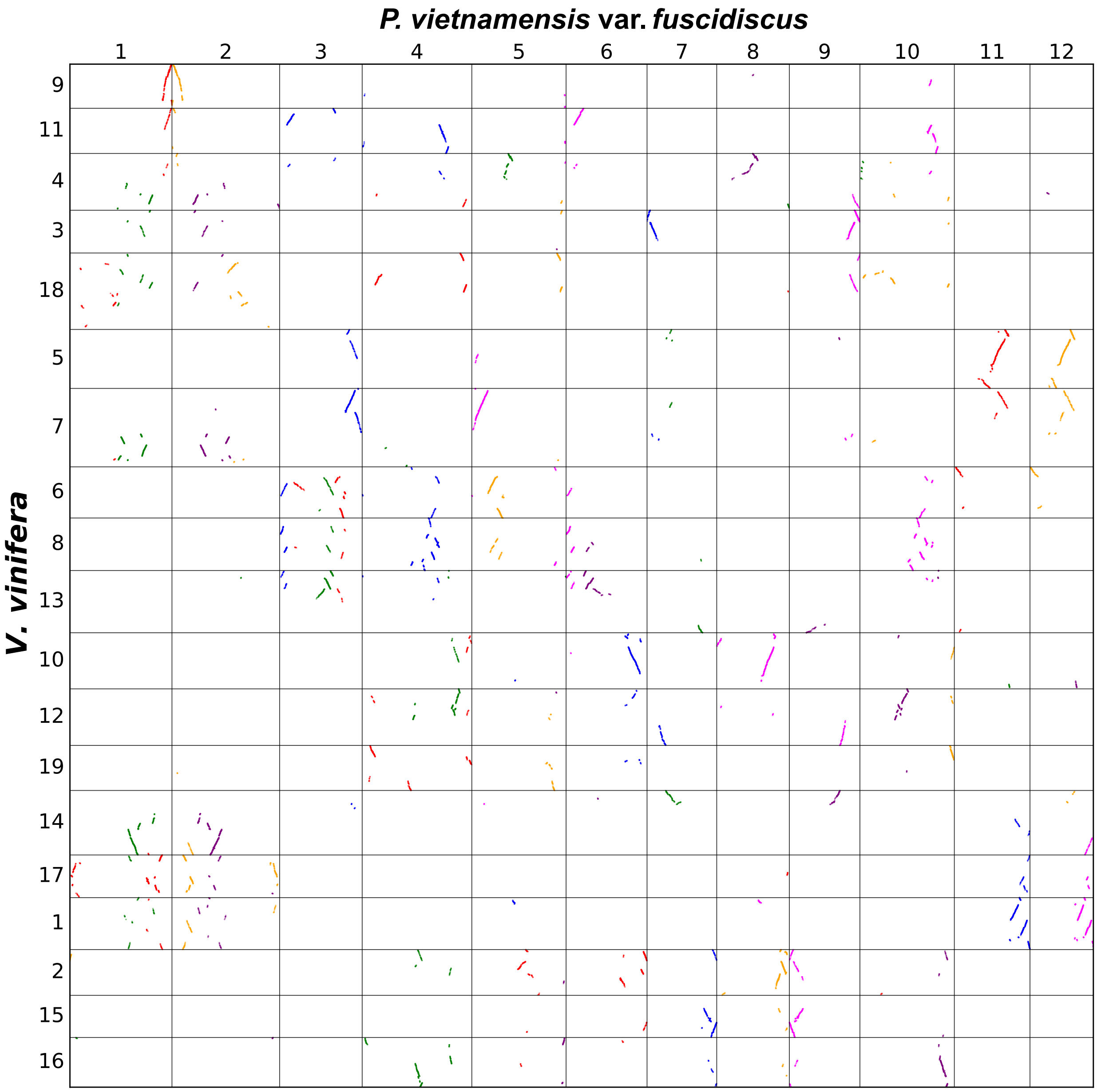


Figure S15 Collinear gene extraction between *V. vinifera* and *P. vietnamensis* var. *fuscidiscus*. The six candidate subsets resulted from polyploidizations were highlighted using red (S1), blue (S2), green (S3), orange (S4), fuchsia (S5), purple (S6).

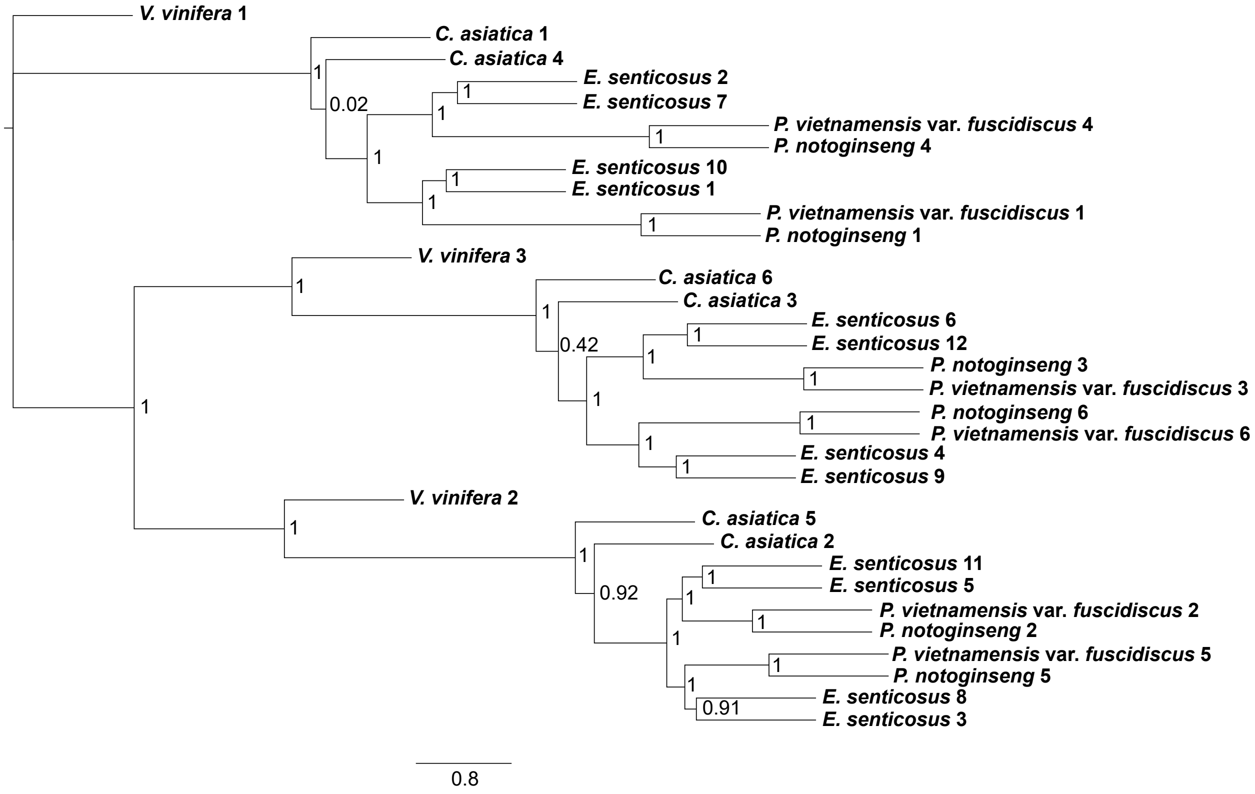


Figure S16 Synteny-based coalescent species tree including *V. vinifera* and four Apiales species. Branch lengths are shown in coalescent units. The numbers of each node represents the local posterior probabilities. Since the ASTRAL tree leaves the branch length of terminal branches empty, the length of terminal branches were all set as one.

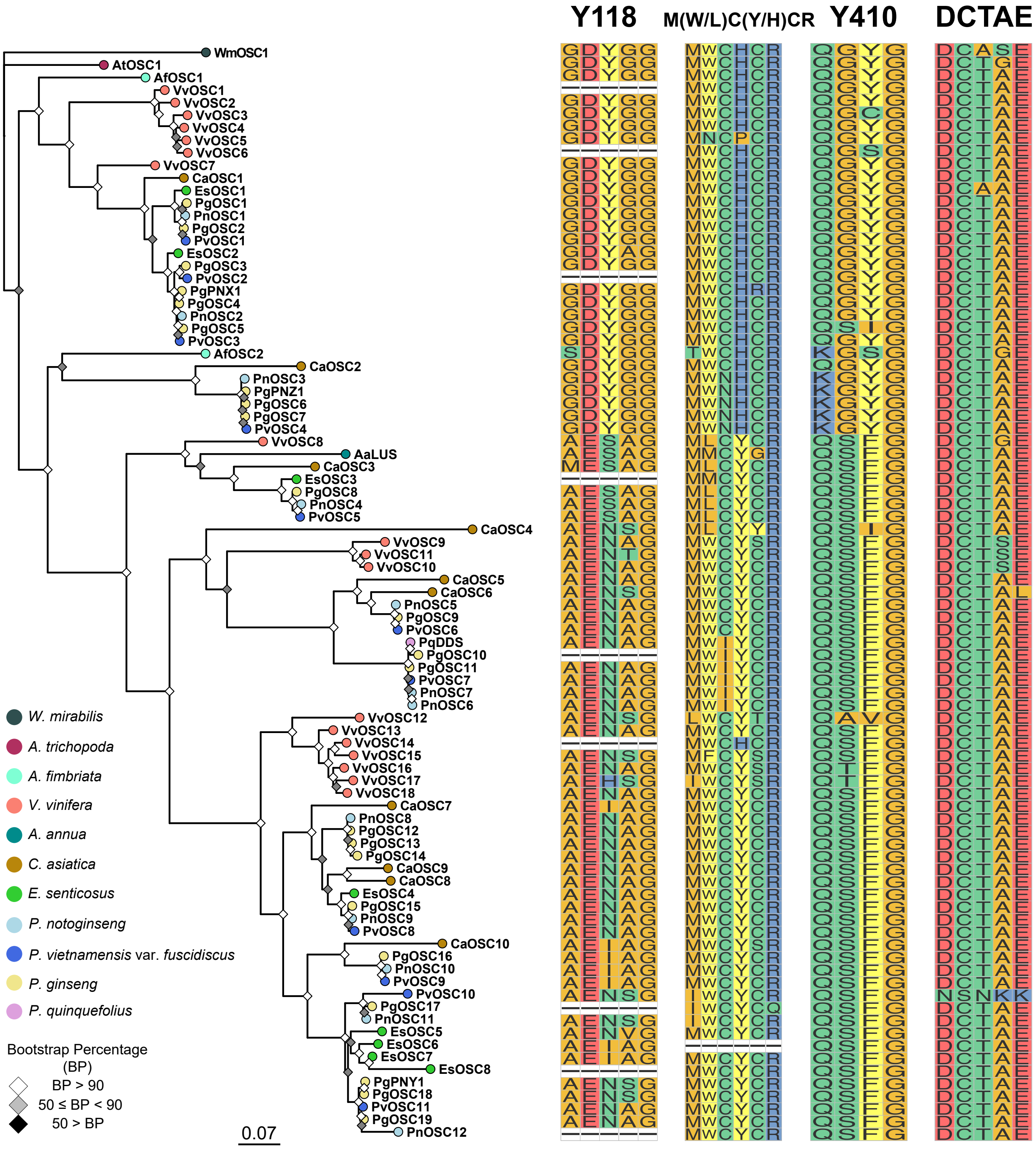


Figure S17 Maximum likelihood phylogenetic tree of OSCs with motifs aligned. Four deterministic motifs were visualized including Y118, M(W/L)C(Y/H)CR, Y410, and DCTAE.

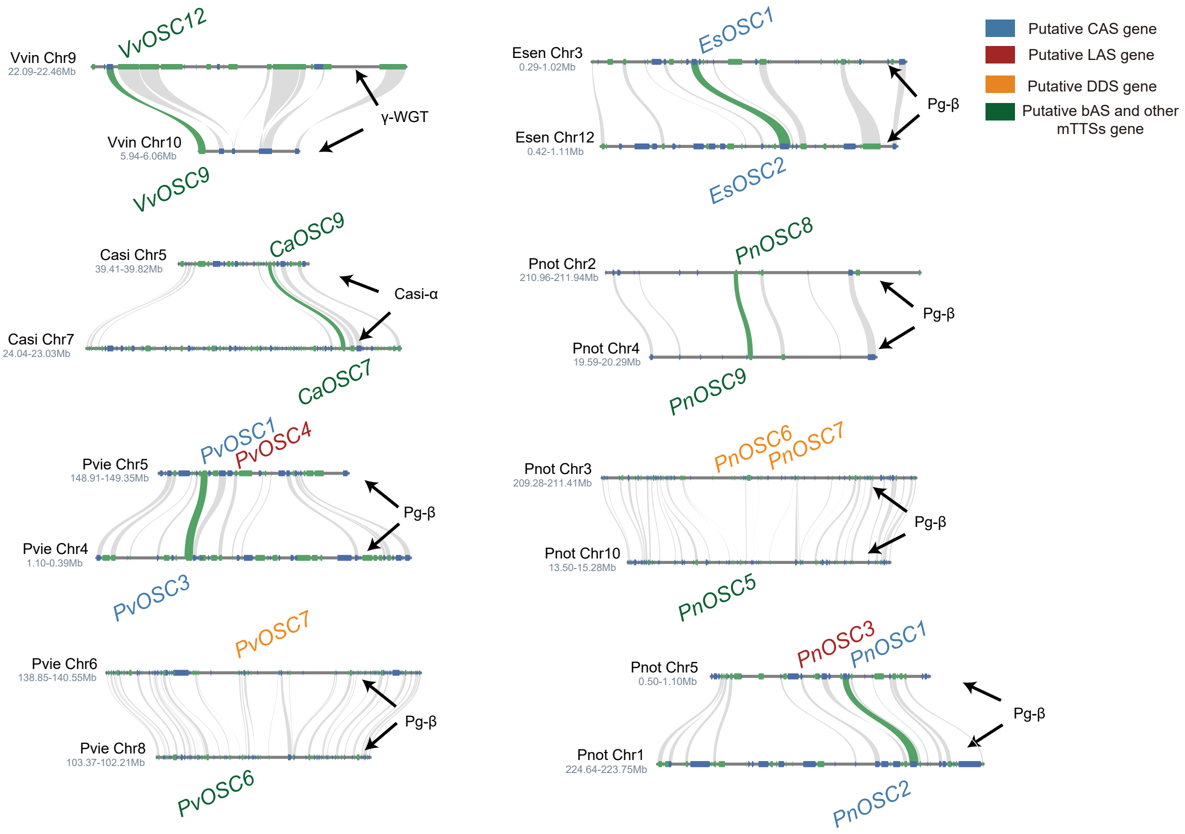


Figure S18 Intraspecific micro-synteny relations of OSC genes. Direct collinear relations for OSC genes were highlighted in green (Vvin: *V. vinifera*, Casi: *C. asiatica*, Pvie: *P. vietnamensis* var. *fuscidiscus*, Esen: *E. senticosus*, Pnot: *P. notoginseng*).

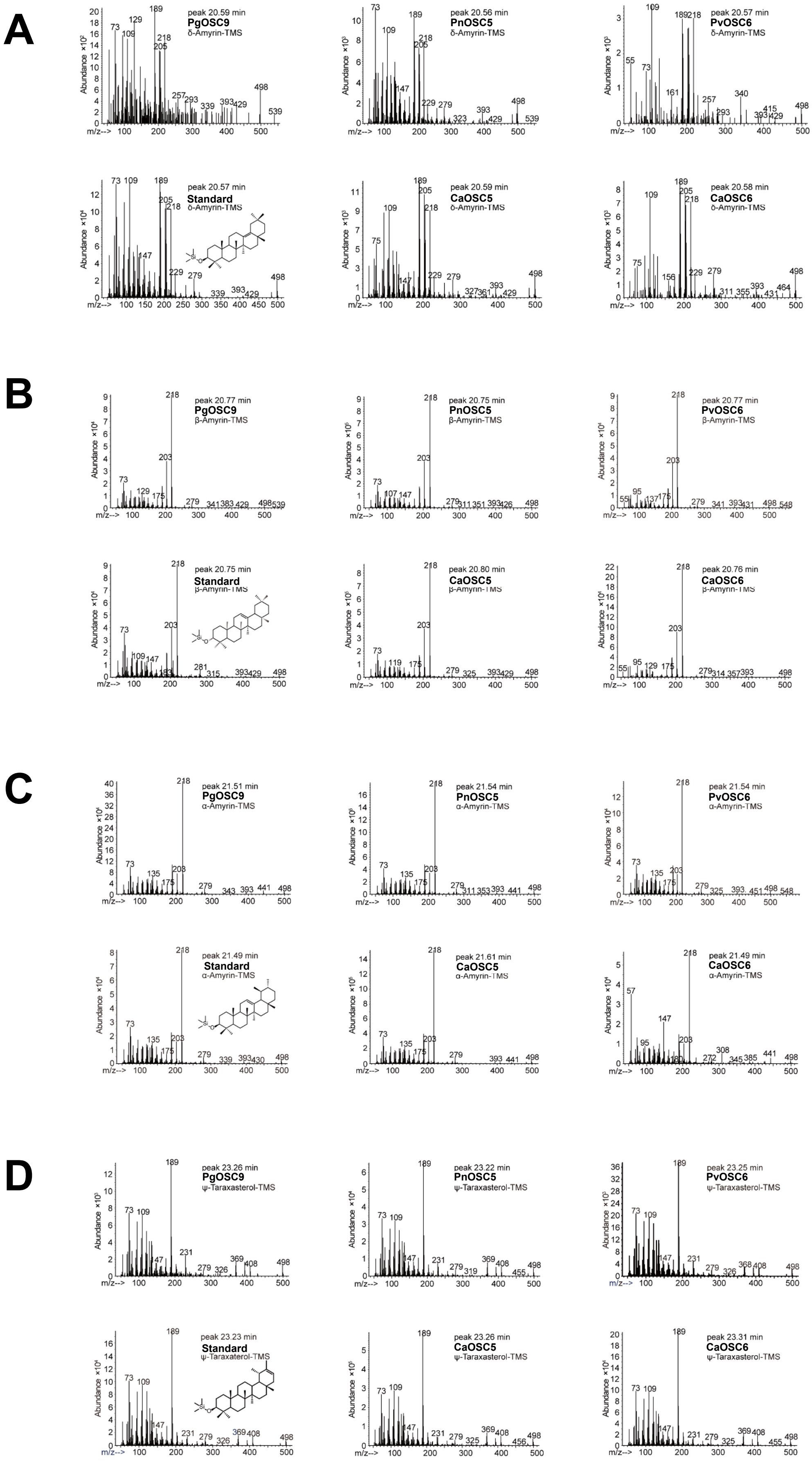


Figure S19 Mass spectra identification for compound 1 (δ-amyrin) (A), compound 2 (β-amyrin) (B), compound 3 (α-amyrin) (C), and compound 5 (ψ-taraxasterol) (D).

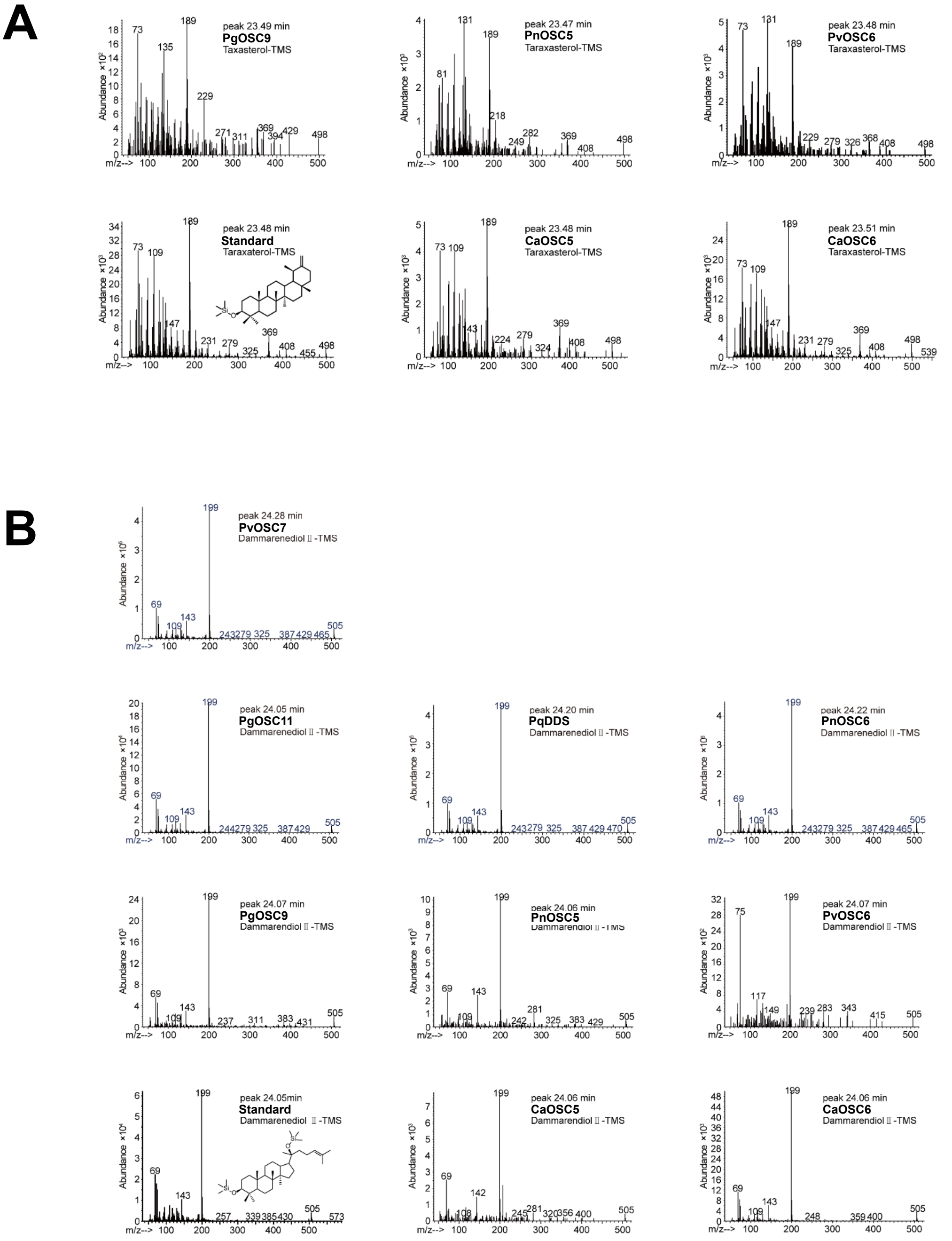


Figure S20 Mass spectra identification for compound 6 (taraxasterol) (A) and compound 7 (dammarendiol-II) (B).

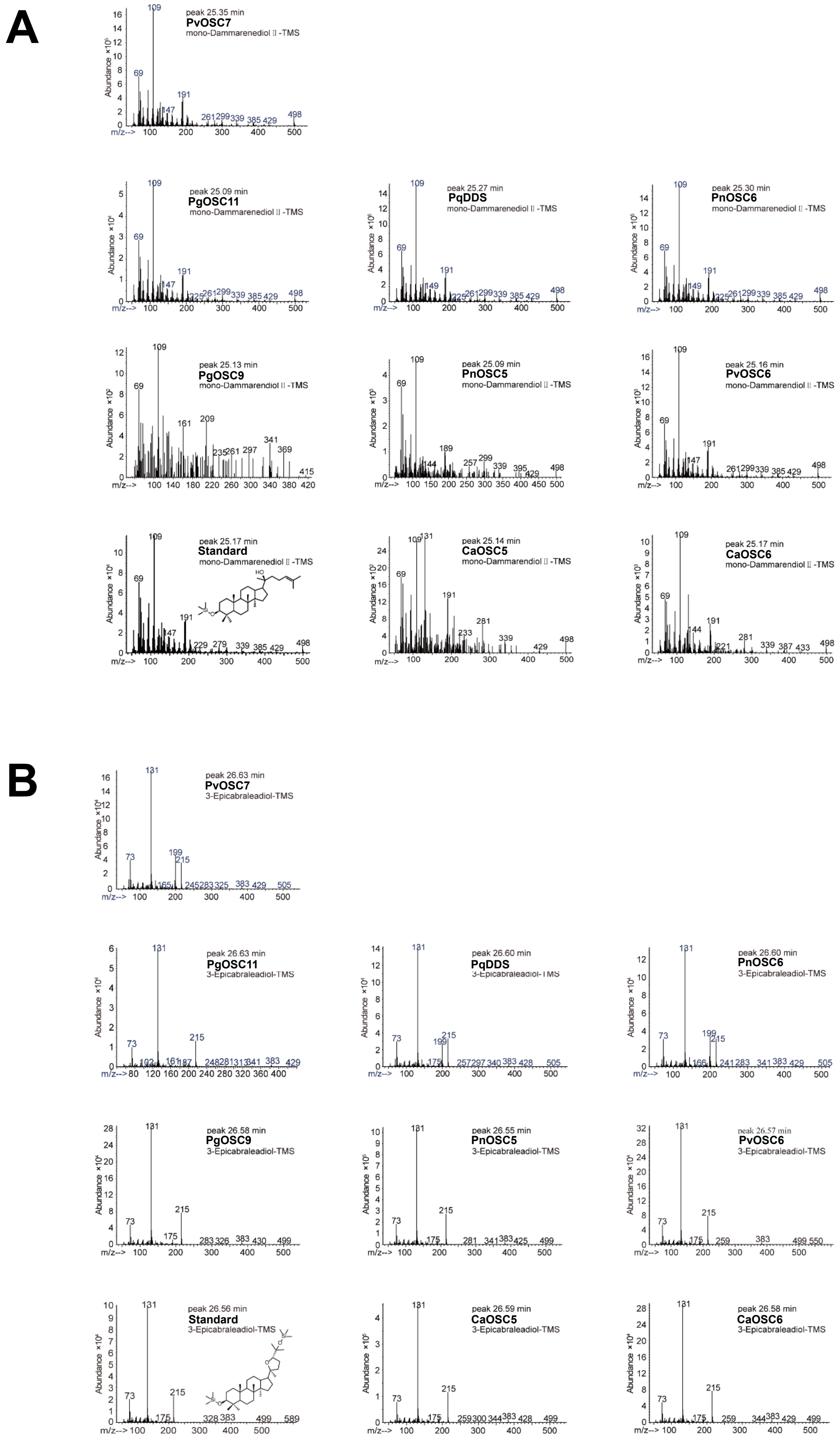


Figure S21 Mass spectra identification for compound # (A) and compound 8 (3-epicabraleadiol) (B). Compound # was identified as the mono-TMS derivative of dammarendiol-II.

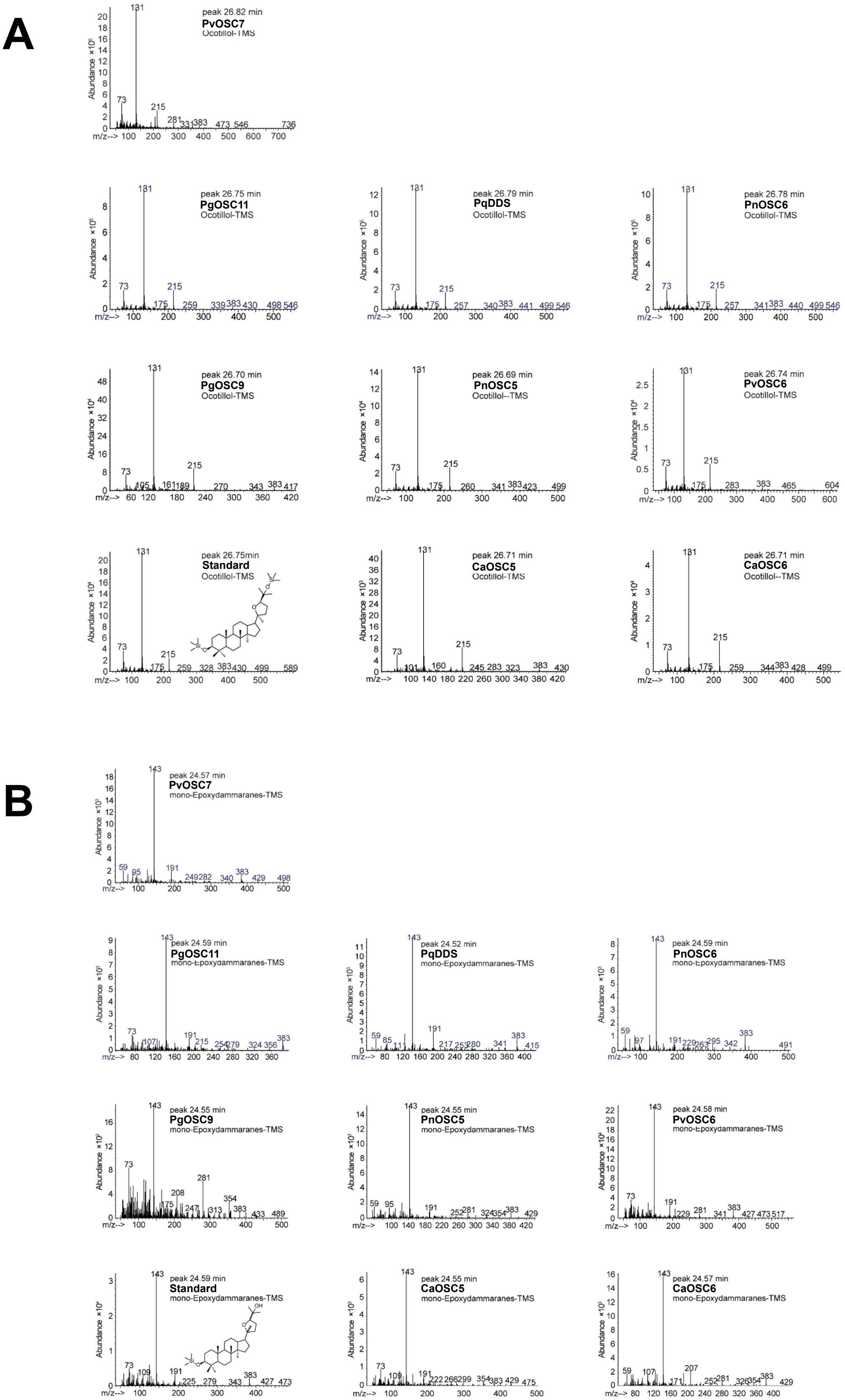


Figure S22 Mass spectra identification for compound 9 (ocotillol) (A) and compound * (B). Compound * was identified as the mono-TMS derivatives of epoxydammaranes (compound 8 and compound 9).

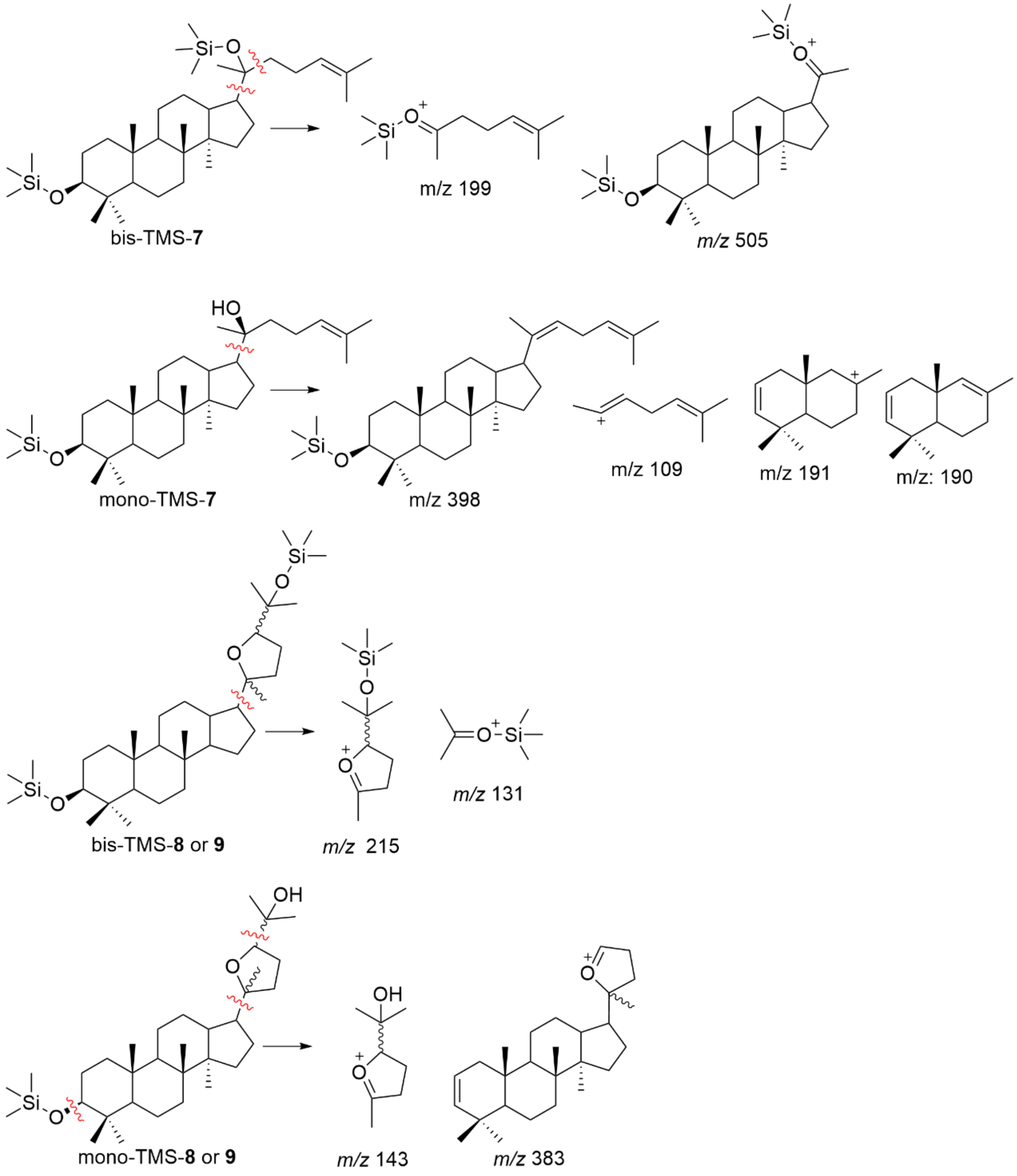


Figure S23 Proposed fragmentation patterns for compound 7, 8, and 9.

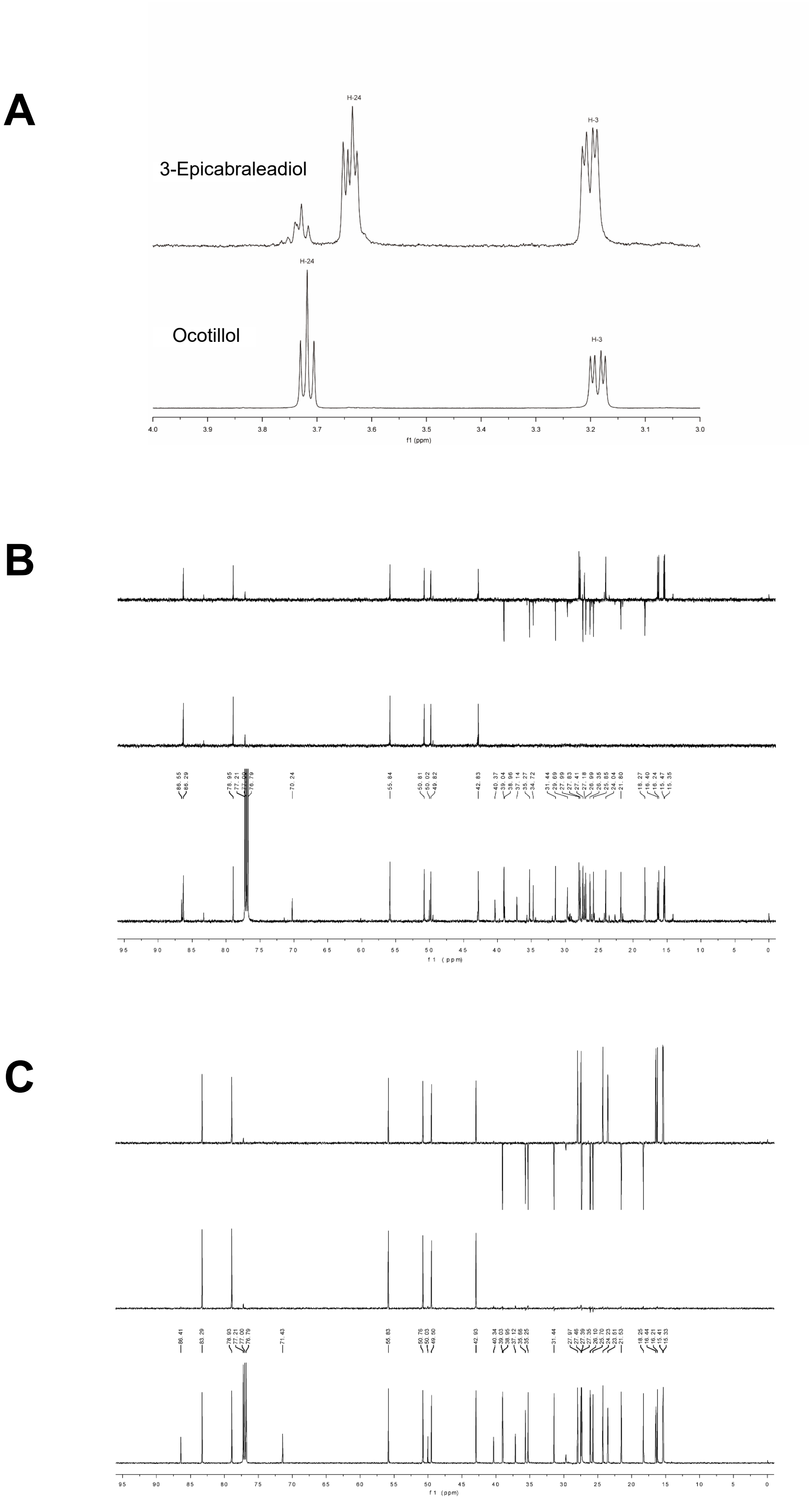


Figure S24 (A) ^{1}H NMR spectra of compound 8 (3-epicabraleadiol) and 9 (ocotillol). (B) ^{13}C NMR and DEPT spectra of compound 8 measured at 150 MHz in $CDCl_3$. (C) ^{13}C NMR and DEPT spectra of compound 9 measured at 150 MHz in $CDCl_3$.

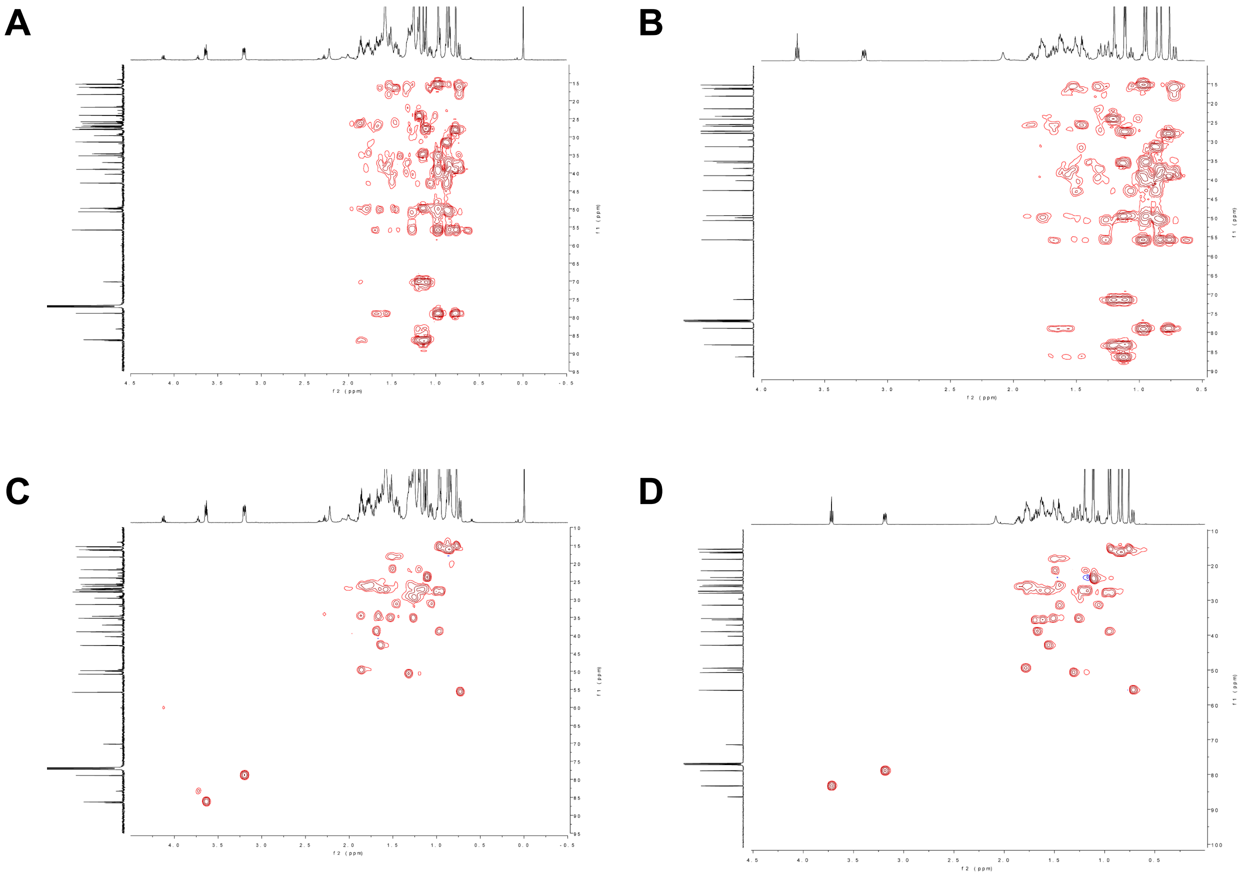


Figure S25 (A) HMBC spectrum of compound 8 measured at 600 MHz in $CDCl_3$. (B) HMBC spectrum of compound 9 measured at 600 MHz in $CDCl_3$. (C) HSQC spectrum of compound 8 measured at 600 MHz in $CDCl_3$. (D) HSQC spectrum of compound 9 measured at 600 MHz in $CDCl_3$.

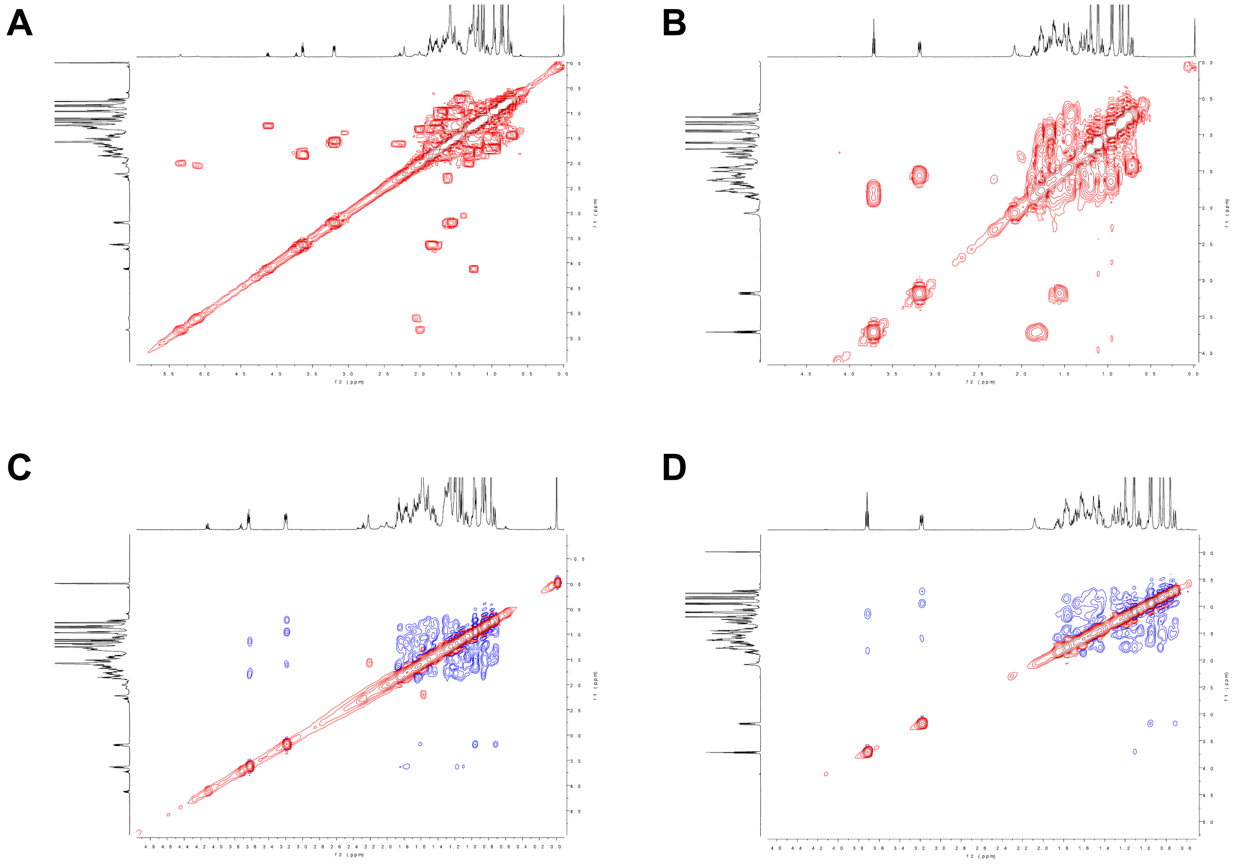


Figure S26 (A) ^{1}H-^{1}H COSY spectrum of compound 8 measured at 600 MHz in $CDCl_3$. (B) ^{1}H-^{1}H COSY spectrum of compound 9 measured at 600 MHz in $CDCl_3$. (C) ^{1}H-^{1}H ROESY spectrum of compound 8 measured at 600 MHz in $CDCl_3$. (D) ^{1}H-^{1}H ROESY spectrum of compound 9 measured at 600 MHz in $CDCl_3$.

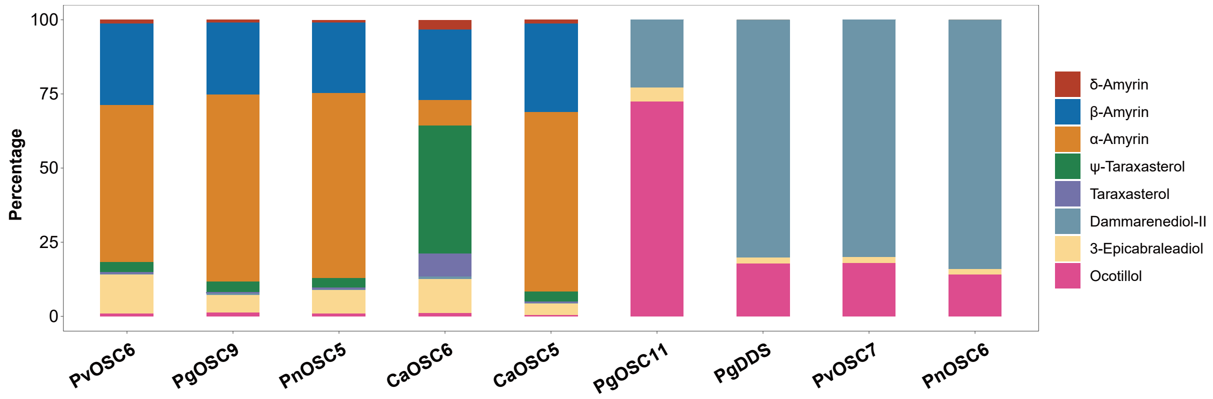


Figure S27 Relative composition of identified products for the nine OSCs. The relative abundance of each compound is calculated based on the area of the corresponding peak.

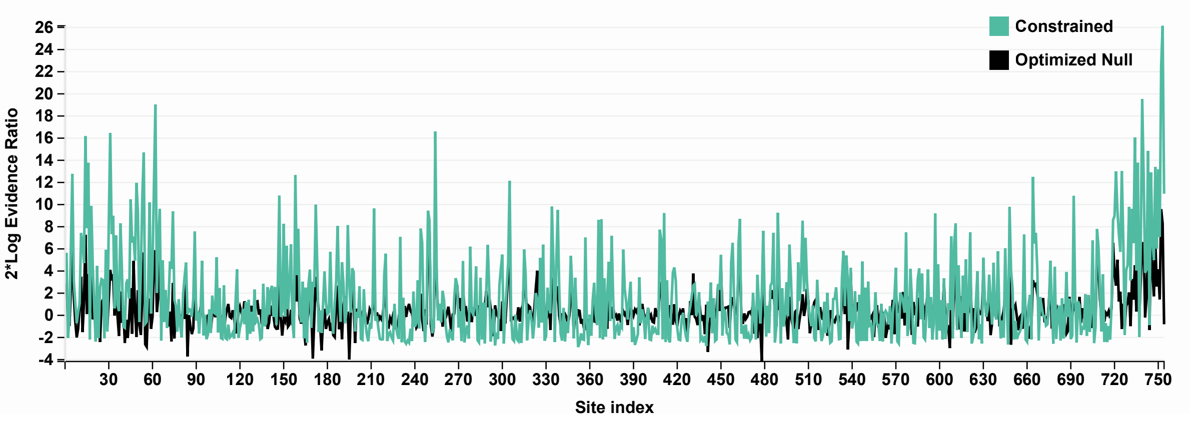


Figure S28 Evidence ratio for BUSTED model in OSCs phylogeny. BUSTED with synyonymous rate variation found evidence (LRT, $P = 0.0000 \leq 0.05$) of gene-wide episodic diversifying selection. The Evidence ratio (y-axis) gives the likelihood ratio (on a log-scale) that the alternative model (selection along test branches) was a better fit to the data as compared to the null model.

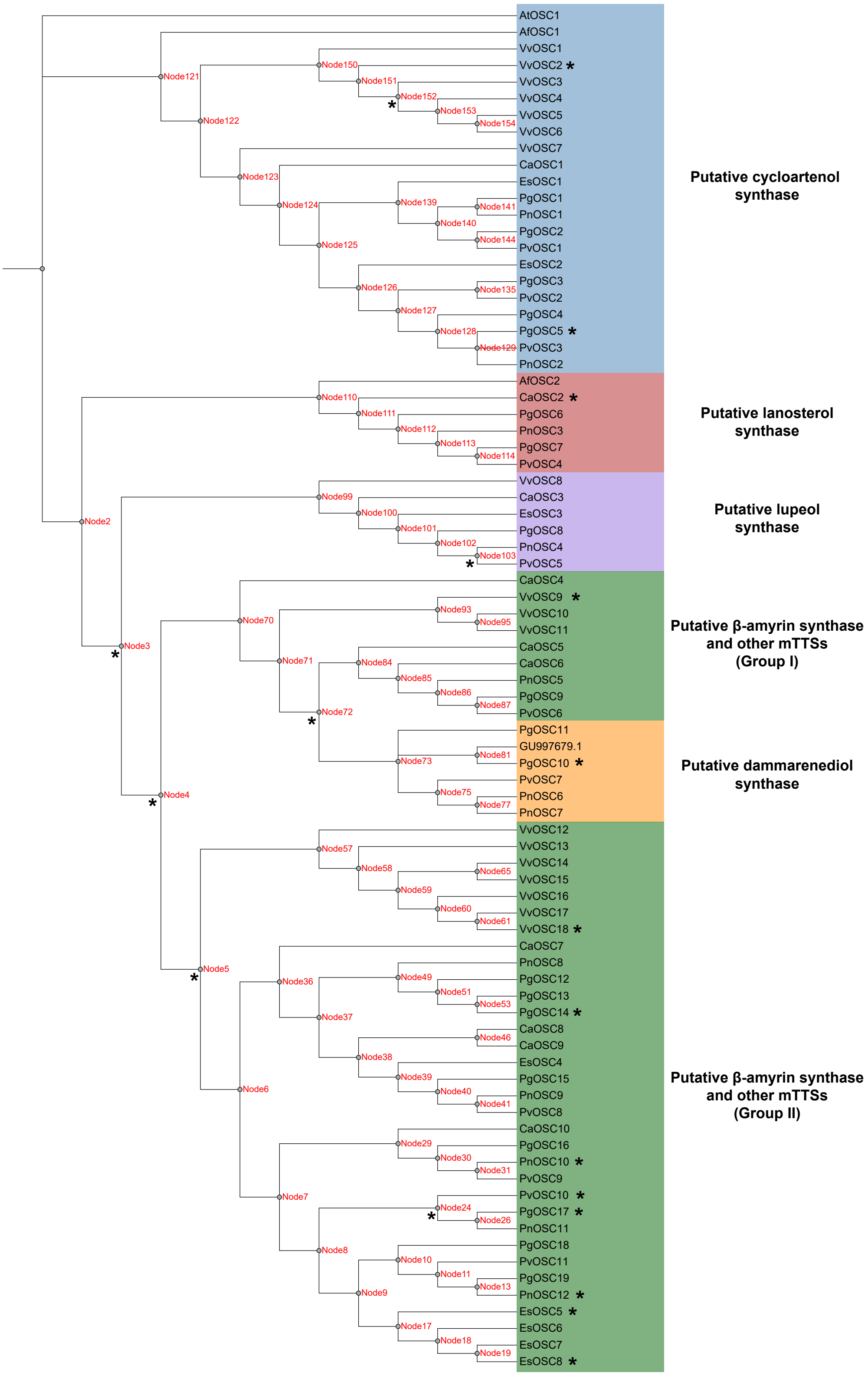
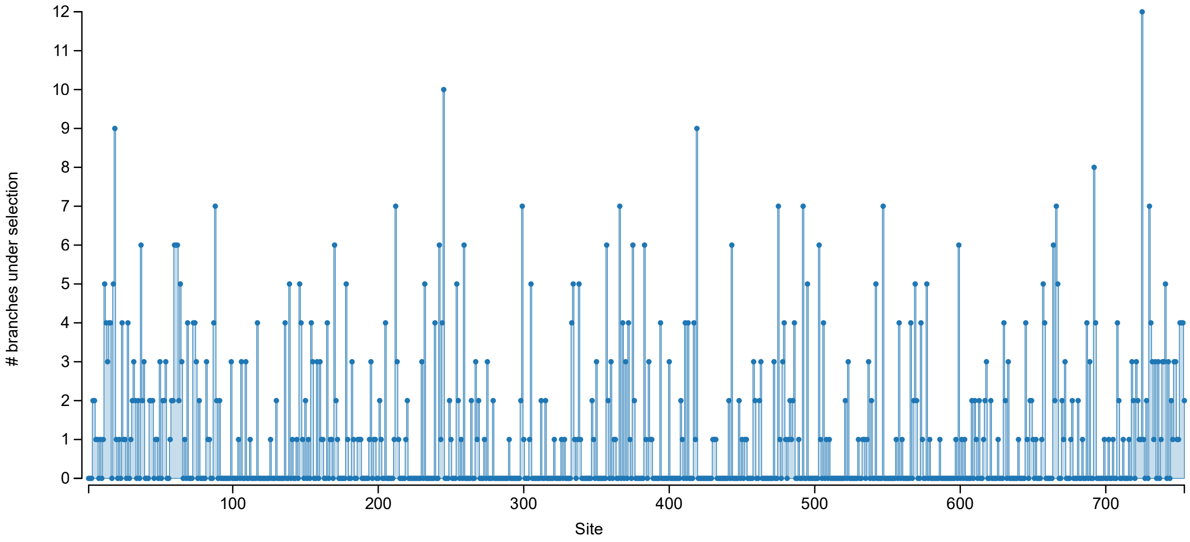


Figure S29 Maximum likelihood phylogenetic tree for OSCs showing aBSREL result for branch specific selection. The branches and internal nodes showing evidence of episodic diversifying selection were labeled with asterisks.

A



B

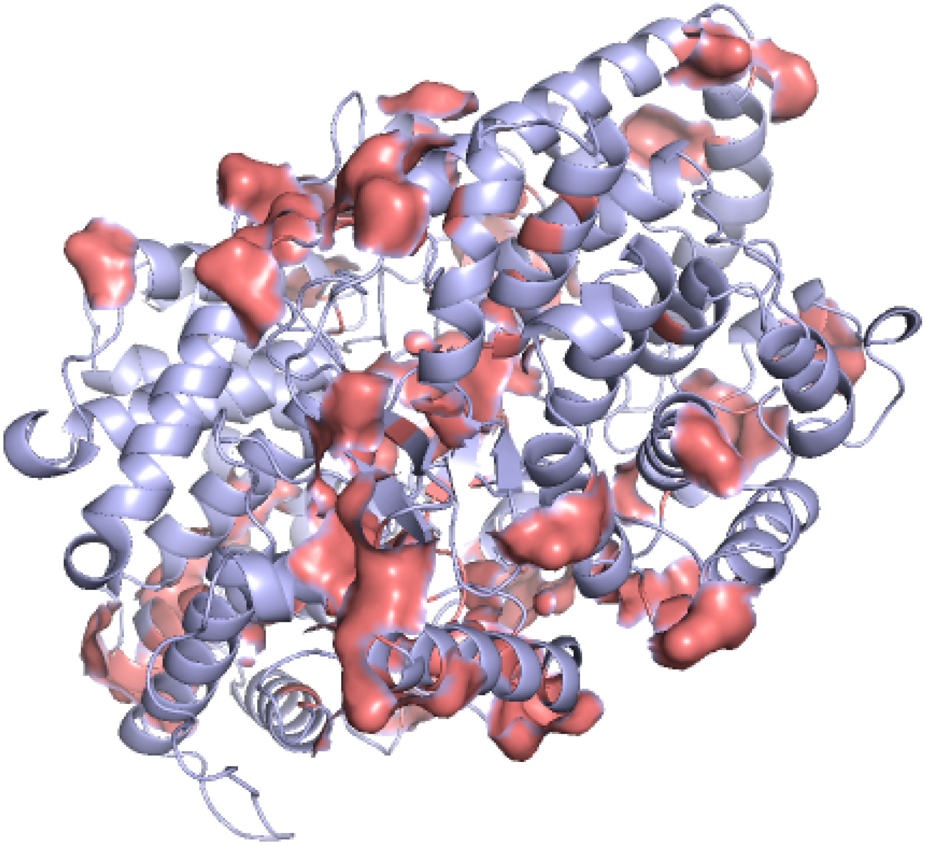


Figure S30 (A) Sites under episodic positive selection detected by MEME. Y-axis showing how many branches may have been under selection under this site (very approximate and rough). (B) 3D structure of P. ginseng CAS (AF-O82139-F1). Sites that have experienced positive selection detected by MEME were highlighted in red.

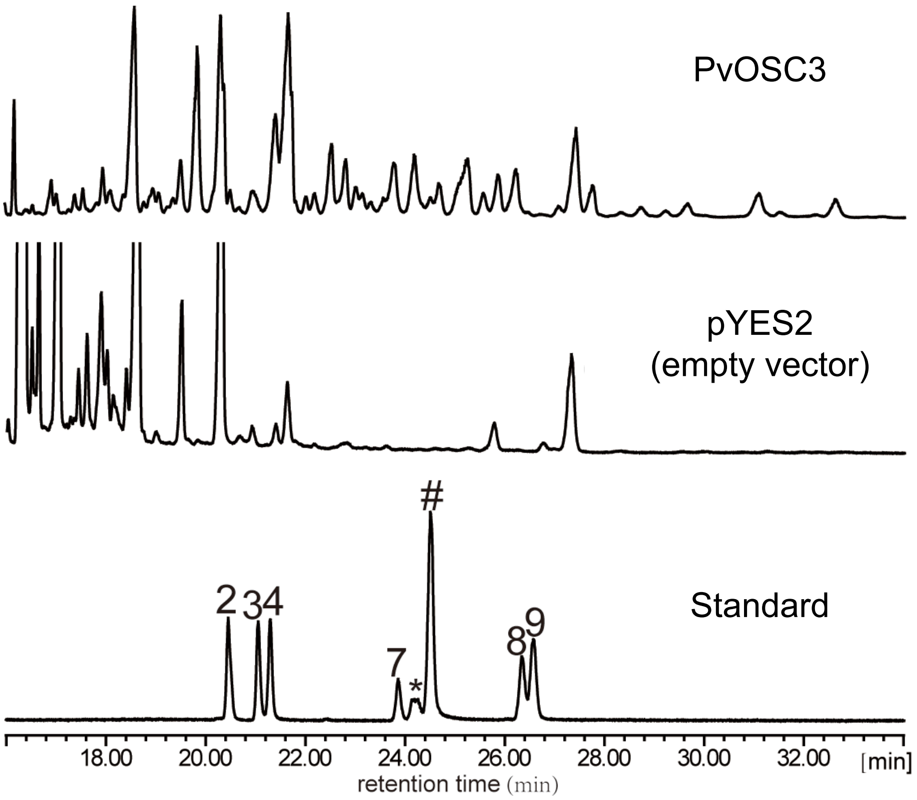


Figure S31 Functional characterization of one *P. vietnamensis* var. *fuscidiscus* CAS using heterelogous expression. * and in total ion chromatograms (TICs) represent epoxydammaranes mono-trimethylsilyl ether and dammarenediol-II mono-trimethylsilyl ether, respectively. 2: β-amyrin, 3: α-amyrin, 4: cycloartenol, 7: dammarenediol II, 8: 20S,24S-3-epicabraleadiol, 9: 20S,24R-ocotillol.

Table S1. Statistics on *P. vietnamensis* var. *fuscidiscus* genome assembly.

Item	**Contig-level assembly**		**Hic chromosome-level assembly**	
	length (bp)	**Number**	**length (bp)**	**Number**
N90	142,377	4,074	85,413,289	12
N80	206,622	3,082	107,381,201	10
N70	265,737	2,344	110,022,002	9
N60	334,899	1,770	120,995,942	7
N50	410,271	1,304	144,079,729	6
N40	493,050	919	149,510,626	5
N30	597,689	600	153,102,694	4
N20	730,980	339	159,980,755	3
N10	997,954	134	160,487,695	2
Max length	2,342,647	-	164,156,069	-
Total length	1,723,337,714	-	1,727,411,714	-
Total number	-	5,866	-	6,305
Average length	293,787	-	273,974	-
Number of sequences >=500bp	-	5,866	-	6,305
Number of sequences >=1000bp	-	5,866	-	6,304
Number of sequences >=2000bp	-	5,866	-	5,676
Number of sequences >=5000bp	-	5,866	-	4,664

Table S2. Genome size estimation of *P. vietnamensis* var. *fuscidiscus* using flow cytometry.

Sample ID	***Solanum lycopersicum* fluorescence intensity**	***P. vietnamensis* var. *fuscidiscus* fluorescence intensity**	**Ratio**	***S. lycopersicum* genome size (Gb)**	***P. vietnamensis* var. *fuscidiscus* genome size (Gb)**	**Average Size (Gb)**	**Standard deviation**
1	108.72	193.78	1.78	0.90	1.60	1.61	0.01
2	109.86	198.5	1.81		1.63		
3	110.09	196.11	1.78		1.60		

Table S3. Genome survey of *P. vietnamensis* var. *fuscidiscus* using Genomescope.

Property	**min**	**max**
Homozygous (aa)	99.14%	100%
Heterozygous (ab)	0%	0.86%
Genome Haploid Length	1,188,704,553 bp	1,430,039,526 bp
Genome Repeat Length	575,700,139 bp	692,580,802 bp
Genome Unique Length	613,004,415 bp	737,458,723 bp
Model Fit	53.47%	96.99%
Read Error Rate	0.22%	0.22%

Table S4. Mapping statistics of Illumina reads to *P. vietnamensis* var. *fuscidiscus* assembly.

	Total reads	**Supplementary reads**	**Mapped reads**	**Reads paired in sequencing**	**Properly paired reads**	**Singletons (only one read mapped)**	**With mate mapped to a different sequence**	**Percentage of mapped reads**	**Percentage of properly paired reads**	**Genome coverage rate**
P. vietnamensis* var. *fuscidiscus	1,546,294,209	16,501,123	1,520,692,682	1,529,793,086	1,445,135,176	3,160,273	50,057,132	98.34%	94.47%	97.40%

Table S5. Comparison of gene space of *P. vietnamensis* var. *fuscidiscus* with other species.

Species	**Gene number**	**Average mRNA length (bp)**	**Total exon number**	**Average exon length (bp)**	**Average cds length per gene (bp)**	**Average exon number**	**Total intron number**	**Average intron length (bp)**	**Average intron length per gene (bp)**
E. senticosus	36,372	5575.58	215,069	241.84	1429.99	5.91	180,779	834.08	4145.60
P. vietnamensis* var. *fuscidiscus	36,454	6166.47	189,971	293.88	1531.5	5.21	158,204	976.21	4236.58
P. notoginseng	36,747	6298.53	199,700	231.69	1259.06	5.43	169,123	979.01	4506.02
P. ginseng	59,352	4394.38	297,411	223.53	1120.12	5.01	241,351	760.71	3093.37
C. asiatica	27,785	3624.05	139,555	245.10	1231.04	5.02	111,770	594.88	2393.00

Table S6. Functional annotation of the predicted genes in *P. vietnamensis* var. *fuscidiscus*.

Type		Number	Percentage (%)
Total		36,454	100
	eggNOG	33,570	92.09
Annotated	**GO**	17,491	47.98
	KEGG	16,023	43.95
	PFAM domains	31,444	86.26
Unannotated		2,884	7.91

Table S7. BUSCO analysis results of genome assemblies and annotated gene sets for *P. vietnamensis* var. *fuscidiscus* and *P. notoginseng*.

Species	**Items**	**Complete BUSCOs**	**Complete and single-copy BUSCOs**	**Complete and duplicated BUSCOs**	**Fragmented BUSCOs**	**Missing BUSCOs**	**Total BUSCO groups searched**
P. vietnamensis* var. *fuscidiscus	Genome (full assembly)	95.3%	84.6%	10.7%	1.0%	3.7%	2,326
	Genome (pseudochromosomes)	95.0%	84.9%	10.1%	1.0%	4.0%	2,326
	Gene set	92.6%	83.4%	9.2%	2.2%	5.2%	2,326
P. notoginseng	Genome (full assembly)	97.5%	84.2%	13.3%	1.0%	1.5%	2,326
	Genome (pseudochromosomes)	96.8%	85.5%	11.3%	1.2%	2.0%	2,327
	Gene set	93.3%	81.7%	11.6%	1.9%	4.8%	2,326

Table S8. Statistics on *P. notoginseng* genome assembly (updated).

Item	length (bp)	Number
N90	152,644,078	12
N80	161,973,221	10
N70	169,268,257	9
N60	176,594,900	7
N50	196,656,904	6
N40	200,662,011	5
N30	213,721,665	4
N20	216,551,805	3
N10	221,362,758	2
Max length	225,175,661	-
Total length	2,402,896,139	-
Total number	-	5,223
Average length	460,060	-
Number of sequences >=500bp	-	5,223
Number of sequences >=1000bp	-	5,223
Number of sequences >=2000bp	-	4,755
Number of sequences >=5000bp	-	4,097

Table S9. Functional annotation of the predicted genes in *P. notoginseng*.

Type		Number	Percentage (%)
Total		36,747	100
Annotated	**eggNOG**	34,098	92.79
	GO	18,151	49.39
	KEGG	17,214	46.84
	PFAM domains	30,993	84.34
Unannotated		2,649	7.21

Table S10. Repeat annotation of *P. vietnamensis* var. *fuscidiscus*.

Type		Length (bp)	Percentage of genome (%)
Tandem repeats	**Simple repeats**	101,728,874	5.90
	Satellite repeats	64,151	0.00
Interspersed repeats	**LTR**	1,360,330,196	78.94
	LTR (Copia)	97,780,472	5.67
	LTR (Gypsy)	939,613,786	54.52
	SINE	39,882	0.00
	LINE	9,521,199	0.55
	DNA transposons	50,052,051	2.90
	Unclassified	100,930,458	5.86
Total (non-redundant)		1,495,684,042	86.79
LTR identity		95.26	
raw LTR assembly index (LAI)		15.18	
LAI		11.63	

Table S11. Repeat annotation of *P. notoginseng*.

Type		Length (bp)	Percentage of genome (%)
Tandem repeats	**Simple repeats**	135,431,446	5.64
	Satellite repeats	1,939	0.00
Interspersed repeats	**LTR**	1,938,176,020	80.66
	LTR (Copia)	109,347,216	4.55
	LTR (Gypsy)	1,333,339,204	55.49
	SINE	8,342	0.00
	LINE	9,181,808	0.38
	DNA transposons	75,167,439	3.13
	Unclassified	134,389,290	5.59
Total (non-redundant)		2,118,941,998	88.18
LTR identity		94.33	
raw LTR assembly index (LAI)		11.87	
LAI		10.95	

Table S12. Classification of transposable elements in *P. vietnamensis* var. *fuscidiscus*.

Class	Order	Superfamily	Clade	Number
Class I retrotransposons	**LTR**	**Copia**	**Ale**	3,358
			Alesia	111
			Angela	7,560
			Bianca	2,232
			Bryco	17
			Lyco	17
			Gymco-III	5
			Gymco-I	4
			Gymco-II	22
			Ikeros	1,537
			Ivana	950
			Gymco-IV	24
			Osser	16
			SIRE	12,000
			TAR	2,070
			Tork	1,487
			mixture/unknown	98,537
		Gypsy	**non-chromo-outgroup**	16
			Phygy	3
			Selgy	3
			Athila	12,422
			TatI	8
			TatII	126
			TatIII	59
			Ogre	19,942
			Retand	7,292
			Chlamyvir	82

				Tcn1	18
				chromo-outgroup	64
				CRM	4,253
				Galadriel	167
				Tekay	233,307
				Reina	3,261
				chromo-unclass	8
				mixture/unknown	639,713
		Retrovirus	**unknown**	**unknown**	2,145
		pararetrovirus	**unknown**	**unknown**	3,263
		DIRS	**unknown**	**unknown**	316
		LINE	**unknown**	**unknown**	6,028
Class II DNA transposons	**Subclass 1**	**TIR**	**EnSpm_CACTA**	**unknown**	3,752
			hAT	**unknown**	3,322
			Merlin	**unknown**	1,762
			MuDR_Mutator	**unknown**	2,708
			PIF_Harbinger	**unknown**	482
			Sola1	**unknown**	1
			Tc1_Mariner	**unknown**	280
	Subclass 2	**Helitron**	**unknown**	**unknown**	667
		Maverick	**unknown**	**unknown**	4,085
mixture		**mixture**	**mixture**	**unknown**	377

Table S13. Classification of transposable elements in *P. notoginseng*.

Class	Order	Superfamily	Clade	Number
Class I retrotransposons	**LTR**	**Copia**	**Ale**	4,060
			Alesia	155
			Angela	15,715
			Bianca	2,933
			Bryco	13
			Lyco	13
			Gymco-III	6
			Gymco-I	15
			Gymco-II	22
			Ikeros	1,820
			Ivana	1,389
			Gymco-IV	16
			Osser	14
			SIRE	10,494
			TAR	2,481
			Tork	1,708
			mixture/unknown	114,324
		Gypsy	**non-chromo-outgroup**	29
			Phygy	4
			Selgy	3
			Athila	18,508
			TatI	7
			TatII	236
			TatIII	75
			Ogre	34,678
			Retand	15,239
			Chlamyvir	99

				Tcn1	16
				chromo-outgroup	57
				CRM	4,791
				Galadriel	232
				Tekay	332,594
				Reina	4,502
				chromo-unclass	22
				mixture/unknown	999,717
		Retrovirus	**unknown**	**unknown**	584
		pararetrovirus	**unknown**	**unknown**	4,309
		DIRS	**unknown**	**unknown**	1,226
		LINE	**unknown**	**unknown**	3,373
Class II DNA transposons	**Subclass 1**	**TIR**	**EnSpm_CACTA**	**unknown**	5,060
			hAT	**unknown**	3,539
			Merlin	**unknown**	1,693
			MuDR_Mutator	**unknown**	5,574
			PIF_Harbinger	**unknown**	1,059
			Sola1	**unknown**	18
			Tc1_Mariner	**unknown**	1,090
	Subclass 2	**Helitron**	**unknown**	**unknown**	1,769
		Maverick	**unknown**	**unknown**	3,616
mixture		**mixture**	**mixture**	**unknown**	560

Table S14. Summary of gene family analysis of *P. vietnamensis* var. *fuscidiscus* with other species.

Type	Number
Number of species	12
Number of genes	416,694
Number of genes in orthogroups	387,841
Number of unassigned genes	28,853
Percentage of genes in orthogroups	93.1
Percentage of unassigned genes	6.9
Number of orthogroups	30,074
Number of species-specific orthogroups	7,487
Number of genes in species-specific orthogroups	36,109
Percentage of genes in species-specific orthogroups	8.7
Mean orthogroup size	12.9
Median orthogroup size	10
G50 (assigned genes)	20
G50 (all genes)	19
O50 (assigned genes)	5,736
O50 (all genes)	6,486
Number of orthogroups with all species present	8,542
Number of single-copy orthogroups	168

Table S15. Results of GO enrichment analysis of expanded gene families in *P. vietnamensis* var. *fuscidiscus* ($P < 0.05$).

ID	Description	GeneRatio	BgRatio	p.adjust	qvalue	Count
GO:0030246	carbohydrate binding	40/253	406/17485	3.37325E-20	2.24067E-20	40
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	38/253	410/17485	3.15485E-18	2.0956E-18	38
GO:0016798	hydrolase activity, acting on glycosyl bonds	38/253	446/17485	4.66915E-17	3.10146E-17	38
GO:0004565	beta-galactosidase activity	32/253	106/17485	4.15053E-31	2.75698E-31	32
GO:0015925	galactosidase activity	32/253	114/17485	3.86914E-30	2.57007E-30	32
GO:0120251	hydrocarbon biosynthetic process	27/253	68/17485	4.60183E-30	3.05675E-30	27
GO:0120252	hydrocarbon metabolic process	27/253	77/17485	1.17591E-28	7.81096E-29	27
GO:0000287	magnesium ion binding	24/253	224/17485	9.14711E-13	6.07594E-13	24
GO:0010333	terpene synthase activity	22/253	32/17485	4.15053E-31	2.75698E-31	22
GO:0016838	carbon-oxygen lyase activity, acting on phosphates	22/253	37/17485	1.80523E-29	1.19912E-29	22
GO:0016835	carbon-oxygen lyase activity	22/253	167/17485	1.69374E-13	1.12506E-13	22
GO:0010334	sesquiterpene synthase activity	19/253	24/17485	2.4489E-29	1.62668E-29	19
GO:0051761	sesquiterpene metabolic process	19/253	24/17485	2.4489E-29	1.62668E-29	19
GO:0051762	sesquiterpene biosynthetic process	19/253	24/17485	2.4489E-29	1.62668E-29	19
GO:0034247	snoRNA splicing	10/253	10/17485	1.95372E-17	1.29776E-17	10

GO:0080013	(E,E)-geranyllinalool synthase activity	10/253	10/17485	1.95372E-17	1.29776E-17	10
GO:0010623	programmed cell death involved in cell development	10/253	22/17485	7.74677E-12	5.14577E-12	10
GO:0045292	mRNA cis splicing, via spliceosome	10/253	23/17485	1.23559E-11	8.20739E-12	10
GO:0080027	response to herbivore	9/253	16/17485	9.99025E-12	6.63599E-12	9
GO:0016635	oxidoreductase activity, acting on the CH-CH group of donors, quinone or related compound as acceptor	8/253	13/17485	7.23975E-11	4.80898E-11	8

Table S16. Results of KEGG enrichment analysis of expanded gene families in *P. vietnamensis* var. *fuscidiscus* ($P < 0.05$).

Term Name	MainClass	GeneRatio	BgRatio	enrichFactor	corrected p-value(BH method)
B 09131 Membrane transport	A09130 Environmental Information Processing	17/331	90/15326	8.745955	1.993E-10
02010 ABC transporters	A09130 Environmental Information Processing	17/331	90/15326	8.745955	1.993E-10
04090 CD molecules	A09180 Brite Hierarchies	10/331	65/15326	7.1234023	1.427E-05
00909 Sesquiterpenoid and triterpenoid biosynthesis	A09100 Metabolism	25/331	201/15326	5.7589695	6.421E-11
00904 Diterpenoid biosynthesis	A09100 Metabolism	5/331	42/15326	5.5121565	0.015191
03032 DNA replication proteins	A09180 Brite Hierarchies	20/331	240/15326	3.8585096	3.761E-06
00240 Pyrimidine metabolism	A09100 Metabolism	8/331	101/15326	3.6674942	0.0127851
04120 Ubiquitin mediated proteolysis	A09120 Genetic Information Processing	15/331	213/15326	3.2607123	0.0006457
B 09109 Metabolism of terpenoids and polyketides	A09100 Metabolism	31/331	498/15326	2.8822602	2.124E-06
03036 Chromosome and associated proteins	A09180 Brite Hierarchies	64/331	1131/15326	2.620102	5.7E-11
A09130 Environmental Information Processing	A09130 Environmental Information Processing	30/331	652/15326	2.1304654	0.0007545

Table S17. Summary of Gaussian kernel analysis on Ks distribution of intraspecific and interspecific colinear gene blocks.

Intraspecific/interspecific colinear gene blocks	**Peak of Ks distribution**	**Deviation**	**R square of linear regression**
A. graveolens (gamma)	1.793	0.1824	0.8052
A. graveolens (Apiaceae-β)	1.0479	0.1343	0.9315
A. graveolens (Apiaceae-α)	0.5749	0.086	0.9434
C. asiatica (gamma)	1.7442	0.2825	0.9207
C. asiatica （Casi-α)	0.7481	0.1003	0.9513
E. senticosus (gamma)	1.4778	0.1737	0.9262
E. senticosus (Pg-β)	0.3818	0.0338	0.9718
E. senticosus (Esen-α)	0.137	0.021	0.9968
P. notoginseng (gamma)	1.5018	0.2433	0.9272
P. notoginseng (Pg-β)	0.3837	0.0362	0.9312
P. vietnamensis var. *fuscidiscus* (gamma)	1.5116	0.2141	0.9418
P. vietnamensis var. *fuscidiscus* (Pg-β)	0.3773	0.0502	0.9604
V. vinifera (gamma)	1.2852	0.155	0.9624
A. graveolens-V. vinifera (gamma)	1.6519	0.4224	0.9612
A. graveolens-V. vinifera (speciation)	1.3022	0.1878	0.9686
C. asiatica-V. vinifera (gamma)	1.5838	0.2906	0.9431
C. asiatica-V. vinifera (speciation)	1.1554	0.1357	0.9766
E. senticosus-V. vinifera (gamma)	1.4135	0.2386	0.9719
E. senticosus-V. vinifera (speciation)	0.9966	0.1098	0.983
P. notoginseng-V. vinifera (gamma)	1.3906	0.2103	0.9516
P. notoginseng-V. vinifera (speciation)	1.0036	0.0868	0.9433
P. vietnamensis var. *fuscidiscus-V. vinifera* (gamma)	1.4178	0.1673	0.9193
P. vietnamensis var. *fuscidiscus-V. vinifera* (speciation)	1.0055	0.1114	0.9448
A. graveolens-C. asiatica (gamma)	1.8155	0.2772	0.9523

A. graveolens-C. asiatica (speciation)	0.8394	0.1392	0.9548
P. vietnamensis var. *fuscidiscus-A. graveolens* (gamma)	1.7051	0.2832	0.9618
P. vietnamensis var. *fuscidicu -A. graveolens* (speciation)	0.7184	0.0922	0.9781
P. vietnamensis var. *fuscidiscus-C. asiatica* (gamma)	1.6343	0.248	0.9469
P. vietnamensis var. *fuscidiscus-C. asiatica* (speciation)	0.5348	0.0734	0.9618
P. vietnamensis var. *fuscidiscus-E. senticosus* (gamma)	1.4891	0.1774	0.9418
P. vietnamensis var. *fuscidicu -E. senticosus* (Pg-β)	0.3746	0.0351	0.9637
P. vietnamensis var. *fuscidiscus-E. senticosus* (speciation)	0.1476	0.0222	0.9772
P. vietnamensis var. *fuscidiscus-P. notoginseng* (gamma)	1.4978	0.2106	0.9145
P. vietnamensis var. *fuscidiscus-P. notoginseng* (Pg-β)	0.3801	0.0404	0.9577
P. vietnamensis var. *fuscidiscus-P. notoginseng* (speciation)	0.0311	0.0094	0.9839

Table S18. Summary of Collinear genomics subsets for five species.

Species	***V. vinifera***	***C. asiatica***	***E. senticosus***	***P. notoginseng***	***P. vietnamensis* var. *fuscidiscus***
Whole genome duplications	γ triplication	γ triplication	γ triplication	γ triplication	γ triplication
	-	Casi-α/Apiaceae-common WGD	Pg-β	Pg-β	Pg-β
	-	-	Esen-α	-	-
Number of collinear copies	3	6	12	6	6

Table S19. General statistics on oxidosqualene cyclase genes from eight species identified using HMMER.

Species	Sequence id	Name	CDS length (bp)	Function/Putative function
A. trichopoda	ERN12565	AtOSC1	2,286	Cycloartenol synthase
A. fimbriata	KAG9449364.1	AfOSC1	2,283	Cycloartenol synthase
	KAG9449355.1	AfOSC2	2,265	Lanosterol synthase
V. vinifera	GSVIVT01029468001	VvOSC1	1,683	Cycloartenol synthase
	GSVIVT01029527001	VvOSC2	2,274	Cycloartenol synthase
	GSVIVT01029514001	VvOSC3	2,274	Cycloartenol synthase
	GSVIVT01029525001	VvOSC4	2,409	Cycloartenol synthase
	GSVIVT01029524001	VvOSC5	2,409	Cycloartenol synthase
	GSVIVT01015994001	VvOSC6	1,683	Cycloartenol synthase
	GSVIVT01032285001	VvOSC7	2,283	Cycloartenol synthase
	GSVIVT01032217001	VvOSC8	2,406	Lupeol synthase
	GSVIVT01021473001	VvOSC9	2,262	β-amyrin synthase and other mTTSs (Group I)
	GSVIVT01021474001	VvOSC10	2,280	β-amyrin synthase and other mTTSs (Group I)
	GSVIVT01021494001	VvOSC11	2,280	β-amyrin synthase and other mTTSs (Group I)
	GSVIVT01021495001		24,726	β-amyrin synthase and other mTTSs (Group I)
	GSVIVT01029510001	VvOSC12	2,445	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029509001	VvOSC13	2,541	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029491001	VvOSC14	1,539	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029488001	VvOSC15	2,310	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029489001	VvOSC16	2,619	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029508001	VvOSC17	2,733	β-amyrin synthase and other mTTSs (Group II)
	GSVIVT01029474001	VvOSC18	2,658	β-amyrin synthase and other mTTSs (Group II)
C. asiatica	evm.model.Scaffold_7.3187	CaOSC1	2,274	Cycloartenol synthase
	evm.model.Scaffold_3.13	CaOSC2	2,391	Lanosterol synthase

	evm.model.Scaffold_1.5050	CaOSC3	4,389	Lupeol synthase
	evm.model.Scaffold_1.163	CaOSC4	2,283	β-amyrin synthase and other mTTSs (Group I)
	evm.model.Scaffold_1.3735	CaOSC5	2,283	β-amyrin synthase and other mTTSs (Group I)
	evm.model.Scaffold_1.3299	CaOSC6	2,298	β-amyrin synthase and other mTTSs (Group I)
	evm.model.Scaffold_7.2291	CaOSC7	2,286	β-amyrin synthase and other mTTSs (Group II)
	evm.model.Scaffold_5.2605	CaOSC8	2,298	β-amyrin synthase and other mTTSs (Group II)
	evm.model.Scaffold_5.2606	CaOSC9	2,295	β-amyrin synthase and other mTTSs (Group II)
	evm.model.Scaffold_7.2982	CaOSC10	2,178	β-amyrin synthase and other mTTSs (Group II)
E. senticosus	Ese03G002792.t1	EsOSC1	2,277	Cycloartenol synthase
	Ese12G002640.t1	EsOSC2	2,274	Cycloartenol synthase
	Ese18G000732.t1	EsOSC3	2,487	Lupeol synthase
	Ese07G000079.t1	EsOSC4	2,286	β-amyrin synthase and other mTTSs (Group II)
	Ese17G001441.t1	EsOSC5	2,160	β-amyrin synthase and other mTTSs (Group II)
	Ese24G002150.t1	EsOSC6	2,253	β-amyrin synthase and other mTTSs (Group II)
	Ese11G000382.t1	EsOSC7	2,295	β-amyrin synthase and other mTTSs (Group II)
	Ese11G000379.t1	EsOSC8	1,752	β-amyrin synthase and other mTTSs (Group II)
P. notoginseng	Pno05G000040.t1	PnOSC1	2,274	Cycloartenol synthase
	Pno01G006888.t1	PnOSC2	2,277	Cycloartenol synthase
	Pno05G000039.t1	PnOSC3	2,331	Lanosterol synthase
	Pno05G003783.t1	PnOSC4	2,178	Lupeol synthase
	Pno10G001026.t1	PnOSC5	2,280	β-amyrin synthase and other mTTSs (Group I)
	Pno03G005730.t1	PnOSC6	2,310	Dammarenediol synthase
	Pno03G005732.t1	PnOSC7	2,310	Dammarenediol synthase
	Pno02G006287.t1	PnOSC8	2,286	β-amyrin synthase and other mTTSs (Group II)
	Pno04G000589.t1	PnOSC9	2,274	β-amyrin synthase and other mTTSs (Group II)
	Pno09G004398.t1	PnOSC10	2,316	β-amyrin synthase and other mTTSs (Group II)
	Pno12G004330.t1	PnOSC11	2,292	β-amyrin synthase and other mTTSs (Group II)
	Pno02G002482.t1	PnOSC12	1,977	β-amyrin synthase and other mTTSs (Group II)
P. vietnamensis	Pvi05G002838.t1	PvOSC1	2,274	Cycloartenol synthase

	Pvi79G000004.t1	PvOSC2	1,977	Cycloartenol synthase
	Pvi04G017143.t1	PvOSC3	2,277	Cycloartenol synthase
	Pvi05G002839.t1	PvOSC4	2,334	Lanosterol synthase
	Pvi05G007905.t1	PvOSC5	2,187	Lupeol synthase
	Pvi08G000242.t1	PvOSC6	2,280	β-amyrin synthase and other mTTSs (Group I)
	Pvi06G002447.t1	PvOSC7	2,310	Dammarenediol synthase
	Pvi02G002534.t1	PvOSC8	2,286	β-amyrin synthase and other mTTSs (Group II)
	Pvi07G000589.t1	PvOSC9	2,289	β-amyrin synthase and other mTTSs (Group II)
	Pvi12G004397.t1	PvOSC10	2,361	β-amyrin synthase and other mTTSs (Group II)
	Pvi01G001914.t1	PvOSC11	2,337	β-amyrin synthase and other mTTSs (Group II)
P. ginseng	Pg_S0266.37	PgOSC1	2,277	Cycloartenol synthase
	Pg_S0762.36	PgOSC2	2,277	Cycloartenol synthase
	Pg_S0910.3	PgOSC3	2,490	Cycloartenol synthase
	Pg_S2798.13	PgOSC4	2,277	Cycloartenol synthase
	Pg_S0701.10	PgOSC5	2,205	Cycloartenol synthase
	Pg_S0266.35	PgOSC6	2,214	Lanosterol synthase
	Pg_S0762.35	PgOSC7	2,331	Lanosterol synthase
	Pg_S0577.13	PgOSC8	2,280	Lupeol synthase
	Pg_S4166.7	PgOSC9	2,289	β-amyrin synthase and other mTTSs (Group I)
	Pg_S3517.9	PgOSC10	1,524	Dammarenediol synthase
	Pg_S3318.3	PgOSC11	2,310	Dammarenediol synthase
	Pg_S4815.4	PgOSC12	2,286	β-amyrin synthase and other mTTSs (Group II)
	Pg_S0034.9	PgOSC13	2,286	β-amyrin synthase and other mTTSs (Group II)
	Pg_S0034.2	PgOSC14	2,295	β-amyrin synthase and other mTTSs (Group II)
	Pg_S2801.2	PgOSC15	2,286	β-amyrin synthase and other mTTSs (Group II)
	Pg_S2939.4	PgOSC16	2,289	β-amyrin synthase and other mTTSs (Group II)
	Pg_S0361.30	PgOSC17	2,043	β-amyrin synthase and other mTTSs (Group II)
	Pg_S2492.7	PgOSC18	2,292	β-amyrin synthase and other mTTSs (Group II)
	Pg_S0888.6	PgOSC19	2,292	β-amyrin synthase and other mTTSs (Group II)

	AB009029	PgPNX1	2,259	Cycloartenol synthase
	AB009030	PgPNY1	2,589	β-amyrin synthase
	AB009031	PgPNZ1	2,684	Lanosterol synthase
Welwitschia mirabilis	W.mirabilis.02578	WmOSC1	2,280	Cycloartenol synthase
Artemisia annua	KM670094	AaLUS	2,274	Lupeol synthase
Panax quinquefolius	GU997679	PqDDS	2,310	Dammarenediol synthase

Table S20. Product profile for OSCs based on GC analysis. √ and × represents presence and absence of compounds.

Protein ID	1	2	3	5	6	7	8	9	*	#	Unidentified
PvOSC6	√	√	√	√	√	√	√	√	√	√	√
PgOSC9	√	√	√	√	√	√	√	√	√	√	√
PnOSC5	√	√	√	√	√	√	√	√	√	√	√
CaOSC6	√	√	√	√	√	√	√	√	√	√	√
CaOSC5	√	√	√	√	√	√	√	√	√	√	√
PgOSC11	×	×	×	×	×	√	√	√	√	√	×
PqDDS	×	×	×	×	×	√	√	√	√	√	×
PvOSC7	×	×	×	×	×	√	√	√	√	√	×
PnOSC6	×	×	×	×	×	√	√	√	√	√	×

Table S21. ^{13}C and ^{1}H NMR assignments for compound 8 and 9.

No.	**Compound 8** (3-epicabraleadiol)		**Compound 9** (ocotillol)	
	δH (J in Hz)	δC	δH (J in Hz)	δC
C1(CH_2)	1.69(1H, m), 0.97(1H, m)	39.04	1.67(1H, m), 0.95(1H, m)	39.03
C2(CH_2)	1.65(1H, m), 1.57(1H, m)	27.41	1.64(1H, m), 1.57(1H, m)	27.39
C3(CH)	3.20 (1H, dd, *J* = 11.5, 4.7)	78.95	3.19 (1H, dd, *J* = 11.5, 4.8)	78.93
C4(qC)		38.96		38.95
C5(CH)	0.73 (1H, dd, J = 11.9, 2.2)	55.84	0.72 (dd, J = 12.0, 2.0)	55.83
C6(CH_2)	1.52 (1H, m), 1.421H, (m)	18.27	1.50(1H, m), 1.44(1H, m)	18.25
C7(CH_2)	1.53(1H, m), 1.27(1H, m)	35.27	1.51(1H, m), 1.26(1H, m)	35.25
C8(qC)		40.37		40.34
C9(CH)	1.32 (1H, m)	50.81	1.31(1H, m)	50.76
C10(qC)		37.14		37.12
C11(CH_2)	1.87(1H, m)，1.50(1H, m)	21.8	1.49(1H, m),1.46(1H, m)	21.53
C12(CH_2)	1.75(1H, m), 1.33(1H, m)	25.85	177(1H, m),1.47(1H, m)	25.7
C13(CH)	1.64 (1H, m)	42.83	1.56 (1H, m)	42.93
C14(qC)		50.02		50.03
C15(CH_2)	1.46 (1H, m), 1.06 (1H, m)	26.99	1.62(1H, m), 1.44(1H, m)	26.1
C16(CH_2)	1.76 (1H, m), 1.63 (1H, m)	31.44	1.84 (1H, m), 1.06 (1H, m)	31.44
C17(CH)	1.86 (1H, m)	49.82	1.79 (1H, m)	49.5
C18(CH_3)	0.97(3H, s)	15.47	0.94(3H, s)	15.41
C19(CH_3)	0.85(3H, s)	16.24	0.83(3H, s)	16.21
C20(qC)		86.55		86.41
C21(CH_3)	1.11(3H, s)	27.18	1.12(3H, s)	23.51
C22(CH_2)	1.87(1H, m), 1.67(1H, m)	34.72	1.62(1H, m), 1.55(1H, m)	35.66
C23(CH_2)	1.80(1H, m), 1.75(1H, m)	26.35	1.85(1H, m), 1.77(1H, m)	27.35
C24(CH)	3.64 (dd, *J* = 10.2, 5.2)	86.29	3.72 (t, J = 7.4)	83.29
C25(qC)		70.24		71.43
C26(CH_3)	1.19(3H, s)	27.83	1.20(3H, s)	27.46

C27(CH_3)	1.11(3H, s)	24.04	1.11(3H, s)	24.23
C28(CH_3)	0.97 (3H, s)	27.99	0.96 (3H, s)	27.97
C29(CH_3)	0.77(3H, s)	15.35	0.76(3H, s)	15.33
C30(CH_3)	0.87(3H, s)	16.4	0.86(3H, s)	16.44

Table S22. Relative composition of identified products for the nine OSCs. The relative abundance of each compound is calculated based on the area of the corresponding peak.

Protein ID	δ-Amyrin	β-Amyrin	α-Amyrin	ψ-Taraxasterol	Taraxasterol	Dammarenediol-II	3-Epicabraleadiol	Ocotillol	Total
PvOSC6	1.35%	27.41%	52.91%	3.35%	0.81%	0.19%	12.99%	0.99%	100.00%
PgOSC9	0.93%	24.29%	63.06%	3.57%	0.46%	0.52%	5.86%	1.31%	100.00%
PnOSC5	1.01%	23.71%	62.46%	3.12%	0.71%	0.16%	7.90%	0.93%	100.00%
CaOSC6	3.34%	23.78%	8.51%	43.33%	7.63%	0.88%	11.37%	1.16%	100.00%
CaOSC5	1.20%	29.97%	60.47%	3.48%	0.51%	0.10%	3.82%	0.45%	100.00%
PgOSC11	0%	0%	0%	0%	0%	22.82%	4.77%	72.41%	100.00%
PqDDS	0%	0%	0%	0%	0%	80.25%	2.04%	17.71%	100.00%
PvOSC7	0%	0%	0%	0%	0%	80.05%	2.05%	17.90%	100.00%
PnOSC6	0%	0%	0%	0%	0%	84.11%	1.85%	14.04%	100.00%

Table S23. Summary of *P. vietnamensis* var. *fuscidiscus* sequencing data.

Type	**Library**	**Insert size (bp)**	**Reads number**	**GC content (%)**	**Mean reads length (bp)**	**Reads length N50 (bp)**	**Base (Gb)**	**Coverage depth**
Illumina genomic sequencing	YSQ-1	150	207,536,126	39.98	-	-	31.13	132.64 X
	YSQ-3	150	421,194,574	40.01	-	-	63.18	
	YSQ-5	150	370,979,332	40.03	-	-	55.65	
	YSQ-7	150	530,083,054	39.14	-	-	79.51	
Pacbio genomic sequencing	YSQ-pacbio	-	11,577,317	-	10,117	17,312	117.13	67.71 X
Hic sequencing	YSQ-Hic	150	1,688,961,966	37	-	-	253.34	146.44 X
RNA sequencing	Leaf-1	150	49,842,528	-	150	150	7.46	12.32 X
	Stem-1	150	43,288,694	-	150	150	6.49	
	Root-1	150	49,166,928	-	150	150	7.36	
Total	-	-	-	-	-	-	621.25	370.94 X

Table S24. Source information of species used in phylogenetic analysis.

Species	**Family**	**Source**
Vitis vinifera	Vitaceae	10.1038/nature06148
Coffea canephora	Rubiaceae	10.1126/science.1255274
Codonopsis pilosula	Campanulaceae	medicinalplants.ynau.edu.cn/genome
Welwitschia mirabilis	Welwitschiaceae	10.1038/s41467-021-24528-4
Lactuca sativa	Asteraceae	10.1038/ncomms14953
Lonicera japonica	Lonicera japonica	10.1111/nph.16552
Centella asiatica	Apiaceae	10.1016/j.ygeno.2021.05.019
Daucus carota	Apiaceae	10.1038/ng.3574
Apium graveolens	Apiaceae	10.1111/pbi.13499
Eleutherococcus senticosus	Araliaceae	10.1111/1755-0998.13403
Panax ginseng	Araliaceae	10.1111/pbi.12926
Panax notoginseng	Araliaceae	This study
Panax vietnamensis var. *fuscidiscus*	Araliaceae	This study