Nakamura et al.   Supplementary Information

Supplementary Table S1. The estimated content of triterpenoids in *A. marina* leaves
The estimated content of each triterpenoid in leaves was calculated based on peak areas from the mass chromatogram shown in Figure 1, by comparing the area of the authentic standard with known concentrations.

Compound (Peak No.)	**Content (mg/g-dry weight)**
β-amyrin (**1**)	0.38
α-amyrin (**2**)	0.56
lupeol (**3**)	2.94
erythrodiol (**4**)	0.03
uvaol (**5**)	0.13
betulin (**6**)	0.21
oleanolic acid (**7**)	0.12
betulinic acid (**8**)	0.97
ursolic acid (**9**)	4.06
maslinic acid (**10**)	0.06
corosolic acid (**11**)	0.35

Supplementary Table S2. Yeast strains generated in this study

Strain No.	**Genotype**
1	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*]
2	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*]
3	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A260*]
4	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C53*]
5	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C54*]
6	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716D59*]
7	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*]
8	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*]
9	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A260*]
10	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C53*]
11	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C54*]
12	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716D59*]
13	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*]
14	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*]
15	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A260*]
16	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C53*]
17	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716C54*]
18	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716D59*]
19	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*]; pYES-DEST52[*GAL1*/-]
20	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*]; pYES-DEST52[*GAL1*/*CYP716A260*]
21	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/*CYP716C53*]
22	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/*CYP716C54*]
23	INV*Sc*1; pYES3[*ADH*/*bAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/*CYP716D59*]
24	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/-]
25	INV*Sc*1; pYES3[*ADH*/*aAS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/*CYP716C53*]
26	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/-]
27	INV*Sc*1; pYES3[*ADH*/*LUS*]; pESC-LEU[*GAL10*/*CPR*, *GAL1*/*CYP716A259*] ; pYES-DEST52[*GAL1*/*CYP716C53*]

Supplementary Table S3. Primers used in this study

Primer No.	**Sequence (5' to 3')**
1	CACCATGGAGGTCTTCTATGCTACTTTCC
2	TTATGCTGCAGATTGGTGAGGATAGAGGC
3	CACCATGGAGGTCTTCTTCGTAACTTTAC
4	TTAAGCTTTGTGTGGGAGTAGGCGAAC
5	AAGCATTTTACGAGGCGAGCTATCC
6	GAACCTCGTTAGGGCTTCTGGCTTCAAG
7	GGTACTGAGATTATCCCCTTCAAGA
8	GCGTTGATGTGTTGCTGGGAAACTAG
9	GAAATACTCGCCGGACATCTTCAAGACC
10	CCCTCATCAGATTCCTGGCGAAGATGGA
11	CACCATGGCTACAGATTTCCTGTTGGCTG
12	TCAGTGCTGAGGGTGAAGCCGAATG
13	CACCATGGAGTTTTTCGCTGTTGCGTTTTC
14	TCAATGATGATACAGGCGGATTGGC
15	CACCATGGAGTTTCTGGGCGTTGCCATCT
16	TTAGTACTGATAAAAGAGGATGGGC

The underlined sequences were added to facilitate unidirectional cloning of the product into pENTR/D-TOPO (Thermo Fisher Scientific).

Supplementary Table S4. A list of plant P450s used to construct a phylogenetic tree shown in Figure 5

CYP name	**Species**	**Accession No.**
CYP51H10	*Avena strigosa*	DQ680852
CYP716A1	*Arabidopsis thaliana*	NM_123002
CYP716A110	*Aquilegia coerulea*	KU878864
CYP716A111	*Aquilegia coerulea*	KY047600
CYP716A113v1	*Aquilegia coerulea*	KU878866
CYP716A12	*Medicago truncatula*	DQ335781
CYP716A140	*Platycodon grandiflorus*	KU878853
CYP716A141	*Platycodon grandiflorus*	KU878855
CYP716A14v2	*Artemisia annua*	KF309251
CYP716A15	*Vitis vinifera*	AB619802
CYP716A154	*Catharanthus roseus*	JN565975
CYP716A17	*Vitis vinifera*	AB619803
CYP716A175	*Malus x domestica*	XM_008392874
CYP716A179	*Glycyrrhiza uralensis*	LC157867
CYP716A2	*Arabidopsis thaliana*	LC106013
CYP716A244	*Eleutherococcus senticosus*	KX354739
CYP716A252	*Ocimum basilicum*	JQ958967
CYP716A253	*Ocimum basilicum*	JQ958968
CYP716A52v2	*Panax ginseng*	JX036032
CYP716A75	*Maesa lanceolata*	KF318733
CYP716A78	*Chenopodium quinoa*	KX343075
CYP716A79	*Chenopodium quinoa*	KX343076
CYP716A80	*Barbarea vulgaris*	KP795926
CYP716A81	*Barbarea vulgaris*	KP795925
CYP716A83	*Centella asiatica*	KU878849
CYP716A86	*Centella asiatica*	KU878848
CYP716C11	*Centella asiatica*	KU878852
CYP716E41	*Centella asiatica*	KU878851
CYP716S1	*Panax ginseng*	JX036031
CYP716S5	*Platycodon grandiflorus*	KU878856
CYP716U1	*Panax ginseng*	JN604536

CYP716Y1	*Bupleurum falcatum*	KC963423
CYP72A67	*Medicago truncatula*	DQ335780

Supplementary Table S5. Contigs related to metabolic pathways

Contig	**KO assignment**	**EC number**	**Definition**
Flavonoid biosynthesis			
Contig00016645	K00475	EC:1.14.11.9	Naringenin 3-dioxygenase
Contig00001041	K00487	EC:1.14.13.11	Trans-cinnamate 4-monooxygenase
Contig00006136	K00487	EC:1.14.13.11	Trans-cinnamate 4-monooxygenase
Contig00002512	K00588	EC:2.1.1.104	Caffeoyl-CoA O-methyltransferase
Contig00017344	K00588	EC:2.1.1.104	Caffeoyl-CoA O-methyltransferase
Contig00032590	K00588	EC:2.1.1.104	Caffeoyl-CoA O-methyltransferase
Contig00000069	K00660	EC:2.3.1.74	Chalcone synthase
Contig00000070	K00660	EC:2.3.1.74	Chalcone synthase
Contig00002910	K01859	EC:5.5.1.6	Chalcone isomerase
Contig00003032	K01859	EC:5.5.1.6	Chalcone isomerase
Contig00031432	K05277	EC:1.14.11.19	Leucoanthocyanidin dioxygenase
Contig00003503	K05278	EC:1.14.11.23	Flavonol synthase
Contig00000141	K05280	EC:1.14.13.21	Flavonoid 3'-monooxygenase
Contig00000142	K05280	EC:1.14.13.21	Flavonoid 3'-monooxygenase
Contig00001063	K09754	EC:1.14.13.36	Coumaroylquinate(coumaroylshikimate) 3'-monooxygenase
Contig00009712	K13065	EC:2.3.1.133	Shikimate O-hydroxycinnamoyltransferase
Flavone and flavonol biosynthesis			
Contig00000141	K05280	EC:1.14.13.21	Flavonoid 3'-monooxygenase
Contig00000142	K05280	EC:1.14.13.21	Flavonoid 3'-monooxygenase
Contig00003088	K10757	EC:2.4.1.91	Flavonol 3-O-glucosyltransferase
Anthocyanin biosynthesis			
Contig00009065	K17193	EC:2.4.2.51	Anthocyanidin 3-O-glucoside 2'''-O-xylosyltransferase
Terpenoid backbone biosynthesis			
Contig00002991	K00021	EC:1.1.1.34	Hydroxymethylglutaryl-CoA reductase (NADPH)

Contig00027198	K00021	EC:1.1.1.34	Hydroxymethylglutaryl-CoA reductase (NADPH)
Contig00004953	K00099	EC:1.1.1.267	1-deoxy-D-xylulose-5-phosphate reductoisomerase
Contig00010665	K00587	EC:2.1.1.100	Protein-S-isoprenylcysteine O-methyltransferase
Contig00001560	K00626	EC:2.3.1.9	Acetyl-CoA C-acetyltransferase
Contig00000300	K00787	EC:2.5.1.1	Farnesyl diphosphate synthase
Contig00013905	K00869	EC:2.7.1.36	MvaK1; mevalonate kinase
Contig00007619	K00919	EC:2.7.1.148	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
Contig00012361	K00919	EC:2.7.1.148	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
Contig00016808	K00938	EC:2.7.4.2	Phosphomevalonate kinase
Contig00014536	K00991	EC:2.7.7.60	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
Contig00016181	K01597	EC:4.1.1.33	Diphosphomevalonate decarboxylase
Contig00000487	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase
Contig00000533	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase
Contig00007119	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase
Contig00007289	K01662	EC:2.2.1.7	1-deoxy-D-xylulose-5-phosphate synthase
Contig00001549	K01770	EC:4.6.1.12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
Contig00011922	K01770	EC:4.6.1.12	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
Contig00008632	K01823	EC:5.3.3.2	Isopentenyl-diphosphate Delta-isomerase
Contig00021964	K01823	EC:5.3.3.2	Isopentenyl-diphosphate Delta-isomerase
Contig00004156	K03526	EC:1.17.7.1	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase
Contig00007194	K03526	EC:1.17.7.1	(E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase
Contig00001843	K03527	EC:1.17.7.4	4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase
Contig00000067	K05356	EC:2.5.1.84	All-trans-nonaprenyl-diphosphate synthase
Contig00013354	K05906	EC:1.8.3.5	Prenylcysteine oxidase / farnesylcysteine lyase
Contig00004669	K05954	EC:2.5.1.58	Protein farnesyltransferase subunit beta
Contig00010059	K05955	EC:2.5.1.58	Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha
Contig00025026	K06013	EC:3.4.24.84	STE24 endopeptidase
Contig00018397	K08658	EC:3.4.22.-	Prenyl protein peptidase
Contig00001034	K10960	EC:1.3.1.83	Geranylgeranyl diphosphate/geranylgeranyl-bacteriochlorophyllide a reductase

Contig00002983	K11778	EC:2.5.1.87	Ditrans,polycis-polyprenyl diphosphate synthase
Contig00012970	K11778	EC:2.5.1.87	Ditrans,polycis-polyprenyl diphosphate synthase
Contig00015907	K11778	EC:2.5.1.87	Ditrans,polycis-polyprenyl diphosphate synthase
Contig00001824	K13789	EC:2.5.1.1	Geranylgeranyl diphosphate synthase, type II
Contig00014790	K13789	EC:2.5.1.1	Geranylgeranyl diphosphate synthase, type II
Contig00020329	K13789	EC:2.5.1.1	Geranylgeranyl diphosphate synthase, type II
Contig00003105	K14066	EC:2.5.1.1	Geranyl diphosphate synthase
Contig00019678	K15889	EC:3.1.1.-	Prenylcysteine alpha-carboxyl methylesterase
Contig00014662	K15891	EC:1.1.1.354	NAD+-dependent farnesol dehydrogenase
Contig00010533	K15892	EC:2.7.1.216	Farnesol kinase
Monoterpenoid biosynthesis			
Contig00012347	K15095	EC:1.1.1.208	(+)-neomenthol dehydrogenase
Sesquiterpenoid and triterpenoid biosynthesis			
Contig00002510	K00511	EC:1.14.14.17	Squalene monooxygenase
Contig00019347	K00511	EC:1.14.14.17	Squalene monooxygenase
Contig00000979	K00801	EC:2.5.1.21	Farnesyl-diphosphate farnesyltransferase
Contig00000012	K14184	EC:4.2.3.104	Alpha-humulene/beta-caryophyllene synthase
Contig00004015	K15815	EC:5.4.99.39	Lupeol synthase 2
Contig00014662	K15891	EC:1.1.1.354	NAD+-dependent farnesol dehydrogenase
Diterpenoid biosynthesis			
Contig00029901	K04121	EC:4.2.3.19	Ent-kaurene synthase
Contig00027530	K04122	EC:1.14.13.78	Ent-kaurene oxidase
Contig00035467	K04123	EC:1.14.13.79	Ent-kaurenoic acid hydroxylase
Contig00022284	K04124	EC:1.14.11.15	Gibberellin 3-beta-dioxygenase
Contig00027341	K04125	EC:1.14.11.13	Gibberellin 2-oxidase
Contig00028311	K04125	EC:1.14.11.13	Gibberellin 2-oxidase
Contig00011784	K05282	EC:1.14.11.12	Gibberellin 20-oxidase
Contig00013148	K05282	EC:1.14.11.12	Gibberellin 20-oxidase

KO, KEGG orthology

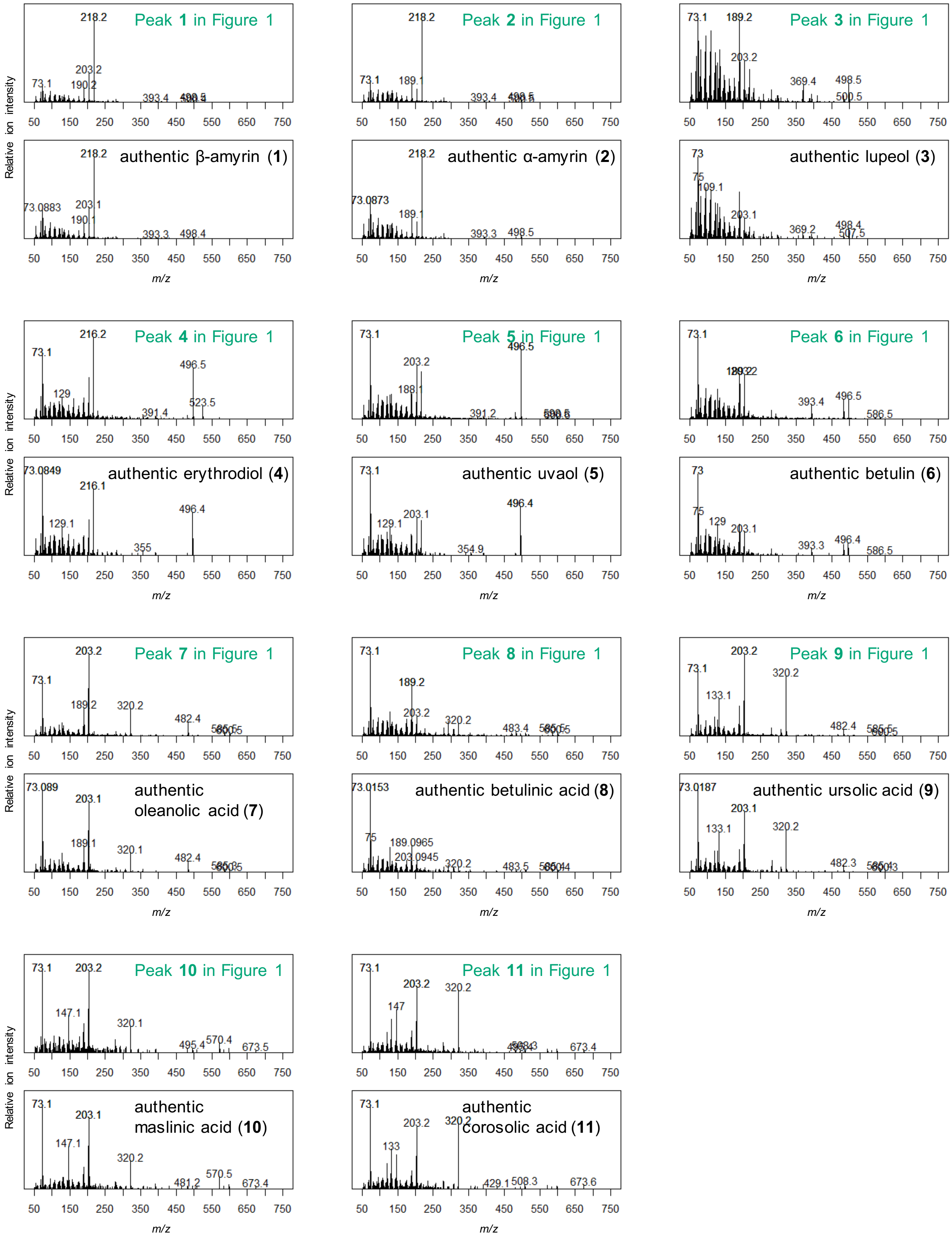
Supplementary Table S6. Contigs related to the mevalonate pathway and triterpenoid biosynthesis

Contig	KO assignment	EC number	Definition	RPKM (leaf)
Contig00002991	K00021	EC:1.1.1.34	Hydroxymethylglutaryl-CoA reductase (NADPH)	29.38
Contig00027198	K00021	EC:1.1.1.34	Hydroxymethylglutaryl-CoA reductase (NADPH)	6.17
Contig00002510	K00511	EC:1.14.14.17	Squalene monooxygenase	40.74
Contig00019347	K00511	EC:1.14.14.17	Squalene monooxygenase	4.74
Contig00001560	K00626	EC:2.3.1.9	Acetyl-CoA C-acetyltransferase	71.92
Contig00000300	K00787	EC:2.5.1.10	Farnesyl diphosphate synthase	15.98
Contig00000979	K00801	EC:2.5.1.21	Farnesyl-diphosphate farnesyltransferase	49.73
Contig00013905	K00869	EC:2.7.1.36	Mevalonate kinase	12.24
Contig00016808	K00938	EC:2.7.4.2	Phosphomevalonate kinase	14.97
Contig00016181	K01597	EC:4.1.1.33	Diphosphomevalonate decarboxylase	31.29
Contig00000487	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase	10.75
Contig00000533	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase	18.95
Contig00007119	K01641	EC:2.3.3.10	Hydroxymethylglutaryl-CoA synthase	9.72
Contig00001824	K13789	EC:2.5.1.10	Geranylgeranyl diphosphate synthase, type II	129.10
Contig00014790	K13789	EC:2.5.1.10	Geranylgeranyl diphosphate synthase, type II	16.97
Contig00020329	K13789	EC:2.5.1.10	Geranylgeranyl diphosphate synthase, type II	33.17
Contig00004015	K15815	EC:5.4.99.39	Lupeol synthase 2	14.77
Contig00004015	K15815	EC:5.4.99.40	Lupeol synthase 2	14.77
Contig00004015	K15815	EC:5.4.99.41	Lupeol synthase 2	14.77

KO, KEGG orthology

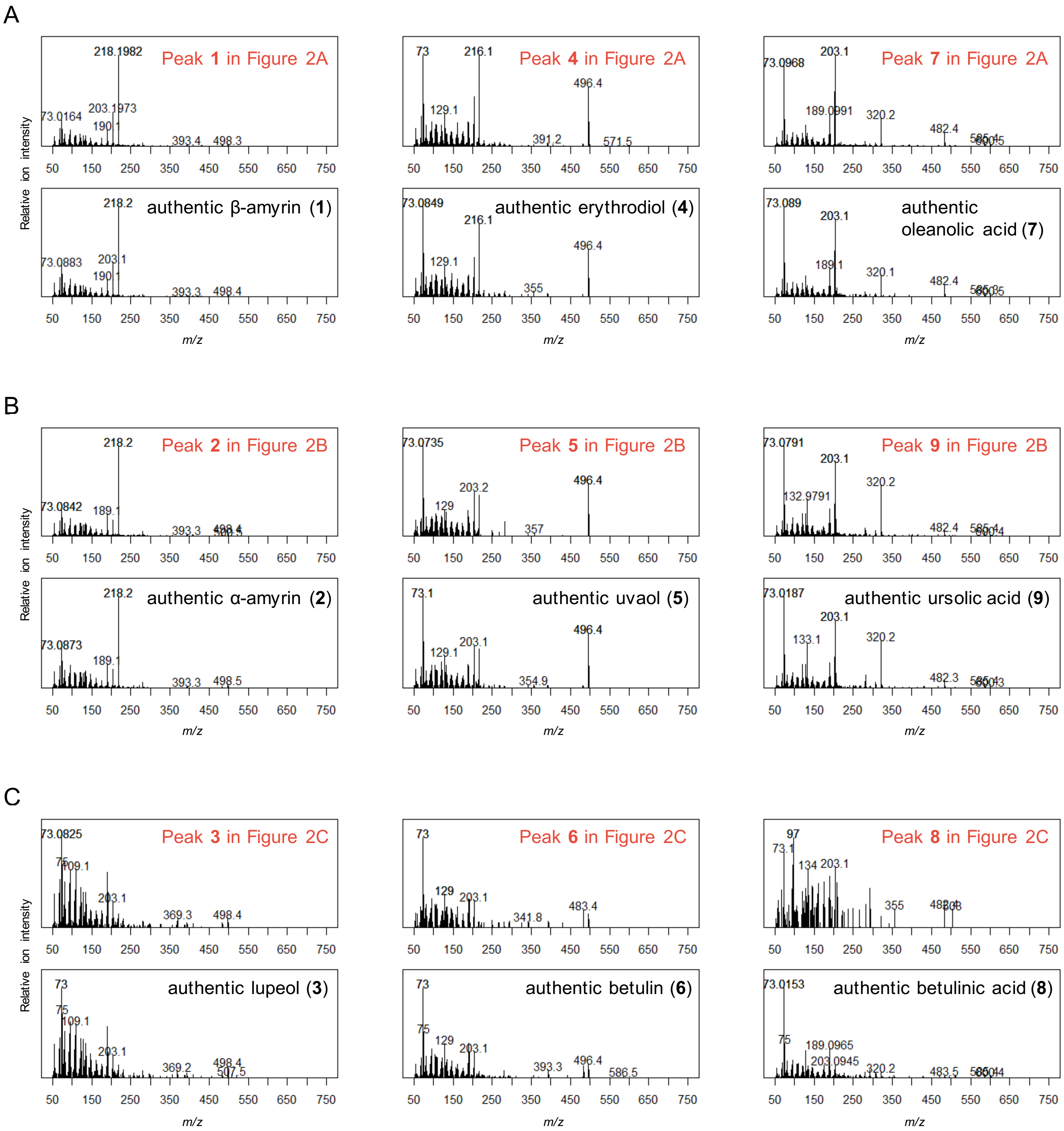
Supplementary Table S7. Results of a tblastn search using CYP716A12 as a query sequence

Query	**Subject**	**Amino acid identities (%)**	**E value**	**Score (bits)**	**RPKM (leaf)**
CYP716A12	Contig00013756 / CYP716A259	76	0.0	773	14.36
CYP716A12	Contig00009575 / CYP716A260	73	0.0	731	24.58
CYP716A12	Contig00032757 / CYP716D59	53	1.00E-141	412	2.32
CYP716A12	Contig00031299 / CYP716C53	45	1.00E-116	349	3.90



Supplementary Figure S1. Mass spectra of the triterpenoids detected in extracts from *A. marina* leaves

Mass spectra of peaks **1-11** from the GC profile (indicated in green) shown in Figure 1 are shown.

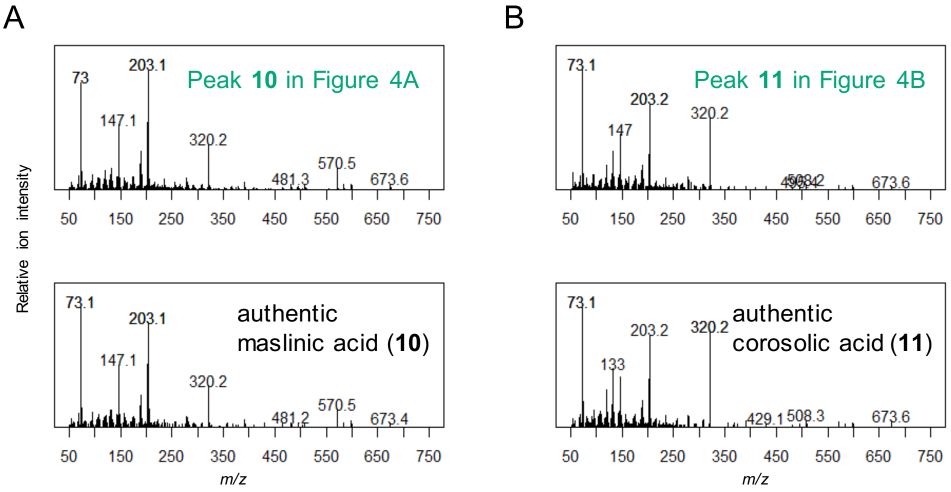


Supplementary Figure S2. Mass spectra of the products from yeast co-expressing OSC and CYP716A259

(A) Mass spectra of peaks **1**, **4**, and **7** from the GC profile shown in Figure 2A and the corresponding authentic standards are shown.

(B) Mass spectra of peaks **2**, **5**, and **9** from the GC profile shown in Figure 2B and the corresponding authentic standards are shown.

(C) Mass spectra of peaks **3**, **6**, and **8** from the GC profile shown in Figure 2C and the corresponding authentic standards are shown.



Supplementary Figure S3. Mass spectra of the products from yeast co-expressing OSC, CPR, CYP716A259 and CYP716C53

(A) Mass spectra of peak **10** from the GC profile shown in Figure 4A and the corresponding authentic standard are shown.

(B) Mass spectra of peak **11** from the GC profile shown in Figure 4B and the corresponding authentic standard are shown.