

## Additional file 1

**Table S1. Clone and primer list.**

X = deoxyuracil  
 F = forward primer  
 R = reverse primer  
 UF = USER compatible forward primer  
 UR = USER compatible reverse primer  
 leaf = amplified from leaf from leaf derived cDNA

Clones only listed with UF and UR, were amplified from cDNA using these primers

Clones list with F, R, UF and UR, were amplified from cDNA using F and R primers

Type	Gene	Family	Clan	Primer type	Sequence
OSCs	OSC1			F	ATGTGGAGGCTAAAGATTGCAGAGG
				R	CTGCTAGACTTTGGAAGATAATGTTACTCTC
				UF	GGCTTAAXATGTGGAGGCTAAAGATTGCAGA
				UR	GGTTAAXGCTAGACTTTGGAAGATAATGTTACTCTC
	OSC2			F	ATGTGGAGGCTAAAGATTGCAGAGG
				R	TTAGGCATCAGAAGGTAGAGGAAGCAG
				UF	GGCTTAAXATGTGGAGGCTAAAGATTGCAGA
				UR	GGTTAAXTTAGGCATCAGAAGGTAGAGGAA
	OSC3			F	ATGTGGAGGCTTAAGGTGGGCGA
				R	GGACTGGGCATTGGTCTAGCATTAC
				UF	GGCTTAAXATGTGGAGGCTTAAGGTGGG
				UR	GGTTAAXTGGGCATTGGTCTAGCATTAC
	OSC4			F	ATGTGGAAGATAAAGATAGCTGAACGTGG
				R	TCATTGCAATTTAATTCCTTTGAAGGCAACT
				UF	GGCTTAAXATGTGGAAGATAAAGATAGCTGAACG
				UR	GGTTAAXTCATTGCAATTTAATTCCTTTGAAGGC
CYPs	CYP712K1	CYP712	71 clan	F	CCTTCTCGCAATCTCCAAT
				R	CCTTAGTTTTCTGCTTAACCG
				UF	GGCTTAAXATGGCCACCATCACTGACATCCAATACT
				UR	GGTTAAXTTAACCGCAAATGGATTGAAACGAACGACCCGG
	CYP712K2	CYP712	71 clan	F	GCCTAAACCACTCTAATAAGCCA
				R	AACCACAGCTTATCTCCAACCATC
				UF	GGCTTAAXATGGCCACTACCATCATTGACGTCCAAT
				UR	GGTTAAXTTAGGAAGAAAATGGATCAAATCGGACGATTGG

(Table S1 continued)

Type	Gene	Family	Clan	Primer type	Sequence
	CYP712K3	CYP712	71 clan	F	CGATAAGCCTAAACCACTCTAATAAGCCA
				R	GCAAGAAAAGGGATGGAATCGGAGA
				UF	GGCTTAAXATGGCCACCACTACCATCATTGAC
				UR	GGTTAAXTTAGCAAGAAAAGGGATGGAATCGGAGA
					ATGGATCTTGTTATGTTAATGCATAAATCAAATCAATTTTATGCCATCTCTGCTCTTTG
	CYP71AN56	CYP71	71 clan	F	GAGGAGATCAATGAGAATGC
				R	GAGGAGATCAATGAGAATGC
				UF	GGCTTAAXATGGATCTTGTTATGTTAATGCATAAAT
				UR	GGTTAAXTTAATGAGAATGCATTAGTGGGACAAGATTAAG
	CYP71AP40	CYP71	71 clan	F	TAATGCTGCCTTCCTTCAGT
				R	GGTGATTTATTGTCTTAGGGG
				UF	GGCTTAAXATGCTGCCTTCCTTCAGTTGCTGATCC
				UR	GGTTAAXTTAGGGGAACCATGGTTTGGCAACAACAATCAG
	CYP71B102	CYP71	71 clan	F	AGACGTAGCATAGTACCATTC
				R	GTGTATCTACGAGTAACTTCAGC
				UF	GGCTTAAXATGGCTCTGCATACCTTAGCAA
				UR	GGTTAAXTTAGCTTGTGTTATTTACTGGTAATGAC
	CYP71E23	CYP71	71 clan	F	TACACGGCTCACTTCTTGTT
				R	ACTTGATACTCCACCGTATG
				UF	GGCTTAAXATGGCCATCATTTCACAAACA
				UR	GGTTAAXTTAACAGAGCCCTATATCCTTCG
	CYP71BE86	CYP71	71 clan	F	CAGTAACCACTCACACCTATCTC
				R	CCAGTATGGAACCTTGTGCTTG
				UF	GGCTTAAXATGGATCTCCAACCTCCGTCCTTCCCGA
				UR	GGTTAAXTTAACCGGAGAGCGAGCGATAAAGGAGTTGGAAC
	CYP71BE85v1	CYP71	71 clan	F	CCTCATGGACCTCCTACAAT
				R	AAGGGAGATAAACTGCTAC
				UF	GGCTTAAXATGGACCTCCTACAATTTCCATCGGTCT
				UR	GGTTAAXTTAGTTAAAGGCAGGTGAAGGACAATAAGGAGT
	CYP71BE84	CYP71	71 clan	F	CCATCATGGATCATCACCTTACT
				R	GTCGTAATATAGGGGAACATCAC
				UF	GGCTTAAXATGGATCATCACCTTACTTCTTTCCAC
				UR	GGTTAAXTTACTTGTTTTCTCAACAACAGAGGACTTATA
	CYP71BE83	CYP71	71 clan	F	TCATCATGGATCTTCAGTTTTC
				R	CACTCTAGTTTCATTCACTCG
				UF	GGCTTAAXATGGATCTTCAGTTTTCTTCTTTCCGG
				UR	GGTTAAXTTACTCGACAATTTGGGATGATAAGGAATTGG
leaf	CYP71BE85v2	CYP71	71 clan	UF	GGCTTAAXATGGACCTCCTACAATTTCCATCGGTCT
				UR	GGTTAAXTTAGTTAAAGGCAGGTGAAGGACAATAAGGAGT
	CYP76T30	CYP76	71 clan	F	AGGAAGTGCTAGGAACAGAC
				R	CACACATGCATTGCCATTTAG
				UF	GGCTTAAXATGGACTACTTTACGCTATTAGCGCTGA
				UR	GGTTAAXTTATACTTGGATAGGGACAGCCAGGAGAGGAAG

(Table S1 continued)

Type	Gene	Family	Clan	Primer type	Sequence
	CYP76Y20	CYP76	71 clan	F	CCTTCCATGCTCTCATCCATAC
				R	ATTGACACACCATTGCT
				UF	GGCTTAAXATGGAACCTCACATCTTCCTTTCTGCC
				UR	GGTTAAXTTAATTGTTATTGGCTTGGATGGGGATAGCACA
	CYP711A62	CYP711	711 clan	F	GCGAATGTTCTGCTACTATCTTCA
				R	CACTACTCTAAGCTTTACACCATGCCT
				UF	GGCTTAAXATGGAGTTAGTGGGTGTAGAGTTTTTC
				UR	GGTTAAXTCATTCTAGAAAGATCACTACTCTAAGCTTTACA
	CYP81AM1	CYP81	71 clan	F	CATGGAAACCCTTCACTACTTGGGTC
				R	AGTGTGCACACACACGCAGATA
				UF	GGCTTAAXATGGAAACCCTTCACTACTTGGGTCTCT
				UR	GGTTAAXTTATAGGTGGGAAAGTGC GCGGTGAGAGGCTG
	CYP82C61	CYP82	71 clan	F	GCTCTCCAGATGGAGTTTCTTC
				R	ACAAACCAGCAACACAAAGTAAAC
				UF	GGCTTAAXATGGAGTTTCTTCTTCACTTCCAACAA
				UR	GGTTAAXTTACCCCATATAGAGATGACTTGAAGGCGTGG
	CYP82AS1	CYP82	71 clan	F	GTCACATGCTTTGCTTCCGTA
				R	CTCCTTAAGCAGAAAACAGTAACA
				UF	GGCTTAAXATGCTTTGCTTCCGTAGCCCCAAGCCCT
				UR	GGTTAAXTTAGTGTACGTAAAGCTTAGGATCAAGACGCGG
	CYP82D166	CYP82	71 clan	F	CTGTTACAATGATTCCCCTCTGTC
				R	TGACGGTAAGGAATAGCCCTAGGT
				UF	GGCTTAAXATGATTCCCCTTCTGCCATACCTGAATA
				UR	GGTTAAXTTATTCATAGGCAGTAGGAGACAAGCGAGGTCT
	CYP82D169	CYP82	71 clan	F	GCTGCAACCATGGAGTACTCAA
				R	AACACAAATGGCAAGTGGCAACA
				UF	GGCTTAAXATGGAGTACTCAATTACCACCATTGGAG
				UR	GGTTAAXTTAATTATAAAGGTGGGTAGGGAGGCGAGGCGA
	CYP82C60	CYP82	71 clan	F	TAACCATCAGCACTCCTGT
				R	CTCCTCTCTCAAACCACAT
				UF	GGCTTAAXATGTCCATACACTTACAGCTCACAACAA
				UR	GGTTAAXTTATCCATATAAATGAGGTGGGAGACGTGGACT
	CYP82J13	CYP82	71 clan	F	CTCACGATCATGGATTTCCC
				R	TTTTCCCTTCCAATTTAGCTT
				UF	GGCTTAAXATGGATTTCCCTTCCAATCTAACAGCAA
				UR	GGTTAAXTTAATAGAGATCAGAACAAAGTCTTGACTGAC
	CYP82D168	CYP82	71 clan	F	CTAGGCAATAGAAACCATGGAATC
				R	TCCTAACGTAGGCATATTGACAC
				UF	GGCTTAAXATGGAATCTTCTCTCTCCTCAGGTAC
				UR	GGTTAAXTTAATCATACAGGTGGGCTGGGTGGCGTGGACT
leaf	CYP82AQ1	CYP82	71 clan	UF	GGCTTAAXATGGATTTTGCCTACCATTGCCACAA
				UR	GGTTAAXTTAAGCATAAAGGTGGGCAGGAAGGCGTGGAGT

(Table S1 continued)

Type	Gene	Family	Clan	Primer type	Sequence
leaf	CYP82D167	CYP82	71 clan	UF	GGCTTAAXATGGCAGACAAGTGC GGCCCAATCTCA
				UR	GGTTTAAAXTTAGGTGCTAGTACTCCAATTATCCATATAGAG
	CYP701A58	CYP701	71 clan	F	ATGGCGGCTCCAACCTCTTCT
				R	TCAATTTCTTGGCTTTATCATTGCGTGCAAC
				UF	GGCTTAAXATGGCGGCTCCAACCTCTTCT
				UR	GGTTTAAXTCAATTTCTTGGCTTTATCATTGCGTGCAAC
	CYP72D19	CYP72	72 clan	F	ACATGGAAGGCTTGATTGTG
				R	CCTCAAAGCGGGTGAAGGATAA
				UF	GGCTTAAXATGGAAGGCTTGATTGTGCAAACGATGT
				UR	GGTTTAAAXTTAAAGCGGGTGAAGGATAATGGGAGCTCCATG
	CYP76A55	CYP72	72 clan	F	TCCGCTTGCTCCAGAATGGTTT
				R	TGGGGTAGGCCACACAATCAT
				UF	GGCTTAAXATGGTTTGCCTCTGTAGTTGGTTTGT
				UR	GGTTTAAAXTTACATGATTCGTTTCTTTGGTATTAGTTTCAA
	CYP711A61	CYP711	711 clan	F	CATGGAATTTGTGGGTGTAACGTG
				R	TCATGTTCTATTCACTACTCTAAGCTTTACAC
				UF	GGCTTAAXATGGAATTTGTGGGTGTAACGTGTA
				UR	GGTTTAAAXTCATGTTCTATTCACTACTCTAAGCTTTACAC
	CYP749A88	CYP749	72 clan	F	CCAGTCATGAGCTTAGTAACCC
				R	CAGAGTGACCAATTTAGCCCTTTTA
				UF	GGCTTAAXATGAGCTTAGTAACCCCTAATTGTTGGTC
				UR	GGTTTAAAXTTAATTCCCATATCCAACCGTTGAAGGATAAT
	CYP749A89	CYP749	72 clan	F	GCAATGGCCAGTTTAGGGA
				R	CCCAATACAACTCAAATGCAG
				UF	GGCTTAAXATGGCCAGTTTAGGGACCCTAGTTCTTA
				UR	GGTTTAAAXTTACTCTACCAGTGGGCGAAGGATTATCTGAAC
	CYP88H2v2	CYP88	85 clan	F	CAAAGAGCAAACAATGGGGTTGG
				R	CTACTCAATTTTGTGATTTTGACCAGATAA
				UF	GGCTTAAXATGGAATTGACCAACGTCATAGGAT
				UR	GGTTTAAAXCTACTCAATTTTGTGATTTTGACCAGATAA
	CYP88H12	CYP88	85 clan	F	ATGGAGTTGTTGAGAGTGTTAGGATG
				R	CAATATCGGTTATTTTGGGAGTTTAGTGA
				UF	GGCTTAAXATGGAGTTGTTGAGAGTGTTAGGATGGG
				UR	GGTTTAAAXTTATTTTGGGAGCTTAGTGATTTTCACGAGATA
	CYP88A89	CYP88	85 clan	F	ATGGAATTGACCAACGTCATAGGAT
				R	GATGAACATTTTGTGATCCTTGCGAC
				UF	GGCTTAAXATGGGGTTGGGGTGGATGGTT
				UR	GGTTTAAAXTTACGCAGATGATGAACATTTTGTGATCCTT
	CYP88H7	CYP88	85 clan	F	ATGGAGTACTCAAGCATGTTAGGATT
				R	CACTCGAGAGCTTGATGATTTTGG
				UF	GGCTTAAXATGTTAGGATTGGCTAGCACTCTCA
				UR	GGTTTAAAXCTAGAGATCACTCGAGAGCTTGATGATT

(Table S1 continued)

Type	Gene	Family	Clan	Primer type	Sequence
	CYP88H8	CYP88	85 clan	F	CGCAATTAAGACTAGAGCGAGAGAGAAATAC
				R	CCACACGGCTTAATTATTATGATGACACT
				UF	GGCTTAAXATGGAGTTATTCAACAACAAGCTAGATTT
				UR	GGTTAAXTTATGATGACACTTTGGTGAGTTTCACTA
	CYP88H11	CYP88	85 clan	F	ATGGAGTTGTTGAGAGTGTTAGGATG
				R	CAATATCGGTTATTTGGGAGTTTAGTGA
				UF	GGCTTAAXATGGAGTTGTTGAGAGTGTTAGGATGGG
				UR	GGTTAAXTTATTTGGGAGCTTAGTGATTTTCACGAGATA
	CYP88A88	CYP88	85 clan	UF	GGCTTAAXATGGAGTTAGGGTCTGTGCTTATATGCA
				UR	GGTTAAXTCATGACGAACATTTATTGATCCTCGCC
	CYP88A44v1	CYP88	85 clan	UF	GGCTTAAXATGGAGTTAGGGTCTGTGCTTATATGCA
				UR	GGTTAAXTCATGACGAACATTTATTGATCCTCGCC
	CYP88A44v2	CYP88	85 clan	UF	GGCTTAAXATGGAGTTAGGGTCTGTGCTTATATGCA
				UR	GGTTAAXTCATGACGAACATTTATTGATCCTCGCC
leaf	CYP88H6	CYP88	85 clan	UF	GGCTTAAXATGGAGTATTCAAGCGTGTAGGATTGG
				UR	GGTTAAXTTAGAGCTTGATGATTTTGGCAAGAAAGTTGTC
leaf	CYP88H1	CYP88	85 clan	UF	GGCTTAAXATGGAGTATTCAACGACAAGCTAGATT
				UR	GGTTAAXTTATGATGATGATACTTTGTGAGTTTCACCAA
leaf	CYP88H10	CYP88	85 clan	UF	GGCTTAAXATGGAGTTGTTGAGAGTGTTAGGATGGG
				UR	GGTTAAXTTATTTGGGAGCTTAGTGATTTTCACGAGATA
leaf	CYP88A43	CYP88	85 clan	UF	GGCTTAAXATGGAGTTAGGGTCTGTGCTTATATGCA
				UR	GGTTAAXTTATGACGAACATTTCTTGATCCTCGCCAAGCA
	CYP88H9	CYP88	85 clan	F	GATGGCAGATATCTTAGAATTGGTC
				R	AGTCCAAGTGGAGCACTTT
				UF	GGCTTAAXATGGCAGATATCTTAGAATTGGTCTTG
				UR	GGTTAAXTTATTTGGGAGTTTTGTGATTTTGACGAGATA
	CYP716C52	CYP716	85 clan	F	GTTGTGACAAGAATGGGAAAG
				R	AGATTAATTGGCTGAGTGAAGG
				UF	GGCTTAAXATGGGAAAGTACTCCAATGAAGTGTCA
				UR	GGTTAAXTTAATTGGCTGAGTGAAGGCGGATTG
	CYP716E53	CYP716	85 clan	F	CTCGTCATTCTTAGCCAT
				R	CTTAATTGTCATGAGGCTGT
				UF	GGCTTAAXATGGATCTCCTGTCTCCATACCTTTTTTC
				UR	GGTTAAXTTAATTGTCATGAGGCTGTAACGAATTGGAAG
	CYP86B39	CYP86	86 clan	F	ACACCATAACATGCTCTAC
				R	GATGAAAGACCATAGTTATTAACA
				UF	GGCTTAAXATGCTCTACAACAACACCACCATGATCA
				UR	GGTTAAXTTAATAGTTCATCTAAGGTACTTATGAAGCTC
	CYP96A145	CYP96	86 clan	F	CTTAAGTCATGGCCTTATTAGGTC
				R	GAAGGATGATAAGGGACAGCTA
				UF	GGCTTAAXATGGCCTTATTAGGTCCTTCTTAAATCC
				UR	GGTTAAXTTAGGCCATCTCTTACTGACCTTCACCTCAA

(Table S1 continued)

Type	Gene	Family	Clan	Primer type	Sequence
Others	P <sub>GAL1</sub>			UF	CACGCGAXAGTACGGATTAGAAGC
				UR	ACCCGTTGAXGGGTTTTTCTCCT
	UBC7 pUG-F			F	GAGATTATCCTAAAAGGAACCTCCCTAGTAATAGTGTAATTTGGAAGGGCATA GCCAGCTGAAGCTTCGTACGC
	pUG-R			R	GTATATAGAGAACAGTTAAAAGGAAGACCAAATGATCATTAACTGCTACCT GCTTTCAGCATAGGCCACTAGTGGATCTG
	UBC7 prom			F	GAAGAACTTACCAGACTGTTTCAAGT

**Table S2. Yeast expression constructs**

<b>name</b>	<b>backbone</b>	<b>promoter*</b>	<b>insert</b>	<b>source</b>
pESC TRP EV	pESC TRP	P <sub>GAL1</sub>		Agilent Technologies, Cat. #217453
pESC TRP CYP712K1	pESC TRP	P <sub>GAL1</sub>	<i>TwCYP712K1</i>	this study
pESC TRP CYP712K2	pESC TRP	P <sub>GAL1</sub>	<i>TwCYP712K2</i>	this study
pESC TRP CYP712K3	pESC TRP	P <sub>GAL1</sub>	<i>TwCYP712K3</i>	this study
pESC LEU EV	pESC LEU	P <sub>GAL1</sub>		Agilent Technologies, Cat. #217452
pESC LEU TwCPR1	pESC LEU	P <sub>GAL1</sub>	<i>TwCPR1</i>	this study
pUUS EV	pUUS			Jensen et al., 2014 <sup>75, a</sup>
pUUS KdFRS	pUUS	P <sub>GAL1</sub>	P <sub>GAL1</sub> , <i>KdFRS</i>	this study
pHTDH EV	pHTDH	P <sub>TDH3</sub>		Triikka et al., 2015 <sup>50</sup>
pHTD-tHMGR2	pHTDH	P <sub>TDH3</sub>	<i>tHMGR2</i>	Triikka et al., 2015 <sup>50</sup>

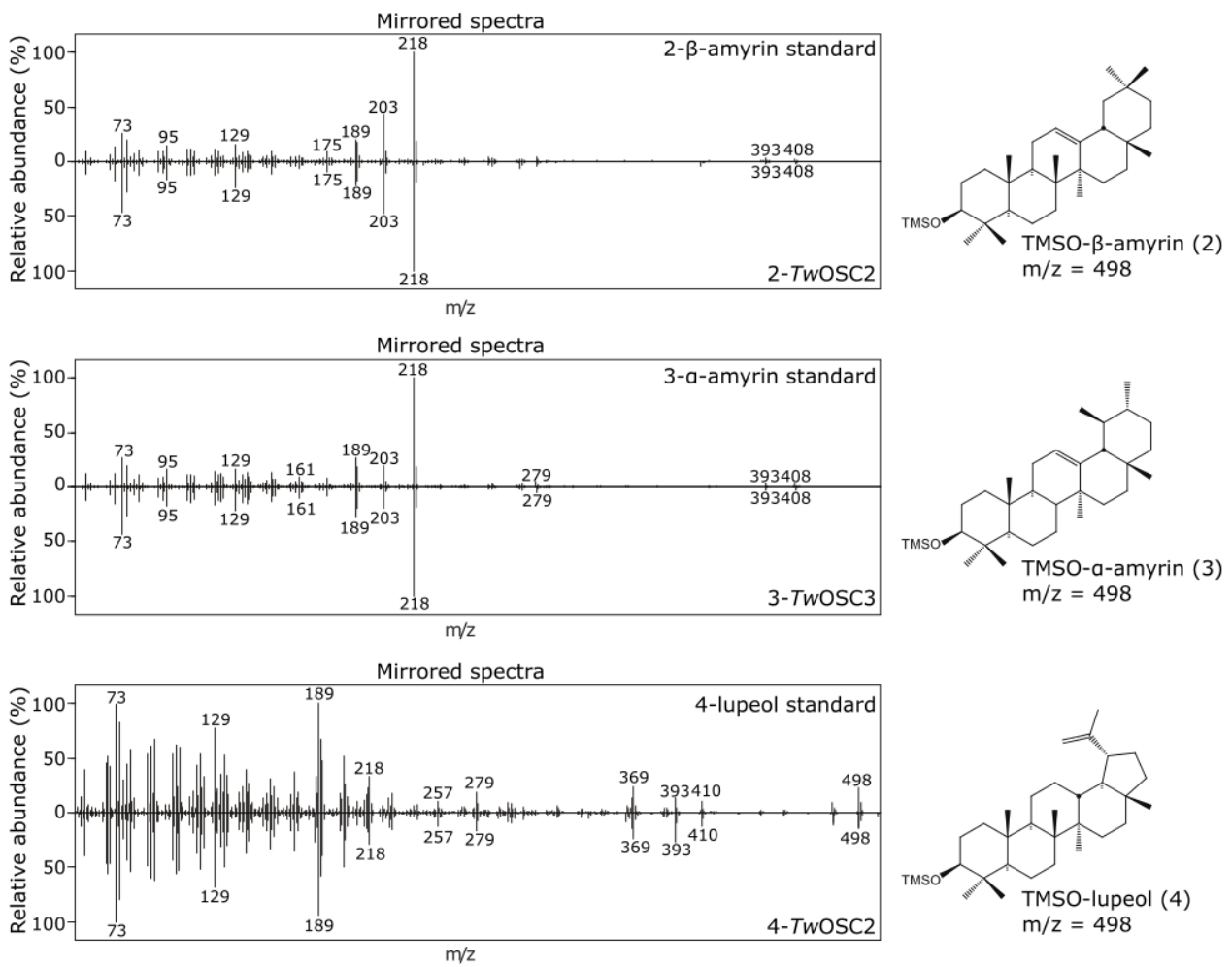
\*Promoter included in backbones except for pUUS.

<sup>a</sup>Citation numbers correspond to references in the main article.

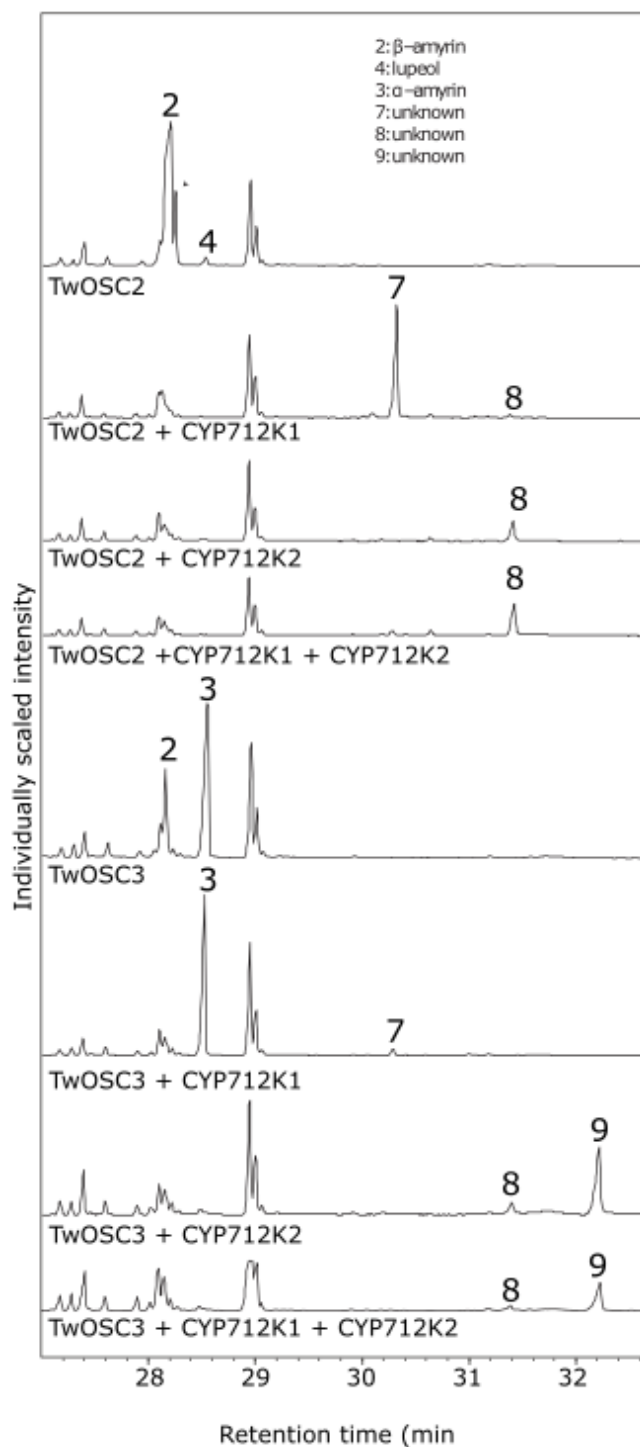
Table S3. Assigned <sup>1</sup>H and <sup>13</sup>C NMR peaks for compound 5 and 6

Position	Compound 5		Compound 6	
	<sup>1</sup> H NMR	<sup>13</sup> C NMR	<sup>1</sup> H NMR	<sup>13</sup> C NMR
	$\delta_H$ ( <i>J</i> in Hz)	$\delta_C$	$\delta_H$ ( <i>J</i> in Hz)	$\delta_C$
1	1.66 (m, H- $\beta$ ) 1.94 (m, H- $\alpha$ )	22,4	1.64 (m, H- $\beta$ ) 1.92 (m, H- $\alpha$ )	22,4
2	2.28 (m, H- $\alpha$ ) 2.36 (m, H- $\beta$ )	41,4	2.24 (m, H- $\alpha$ ) 2.31 (m, H- $\beta$ )	41,5
3		213,3		214,4
4	2.22 (q, <i>J</i> = 6.5)	58,4	2.2 (q, <i>J</i> = 6.5)	58,2
5		42,3		42,2
6	1.27 (m, H- $\beta$ ) 1.72 (m, H- $\alpha$ )	41,7	1.23 (m, H- $\beta$ ) 1.68 (m, H- $\alpha$ )	41,2
7	1.38 (m, H- $\beta$ ) 1.46 (m, H- $\alpha$ )	18,4	1.28 (m, H- $\beta$ ) 1.43 (m, H- $\alpha$ )	18,2
8	1.38 (m)	53,5	1.33 (m)	50,6
9		37,6		37,5
10	1.51 (m)	59,6	1.49 (m)	59,5
11	1.25 (m, H- $\beta$ ) 1.45 (m, H- $\alpha$ )	35,8	1.25 (m, H- $\beta$ ) 1.40 (m, H- $\alpha$ )	35,2
12	1.13 (m, H- $\beta$ ) 1.45 (m, H- $\alpha$ )	29,9	1.16 (m, H- $\beta$ ) 1.48 (m, H- $\alpha$ )	30,4
13		39,6		39,3
14		38,4		39,3
15	1.26 (m, H- $\beta$ ) 1.51 (m, H- $\alpha$ )	32,9	1.28 (m, H- $\beta$ ) 1.43 (m, H- $\alpha$ )	29,8
16	1.30 (m, H- $\beta$ ) 1.51 (m, H- $\alpha$ )	36	1.34 (m, H- $\beta$ ) 1.61 (m, H- $\alpha$ )	36,2
17		29,9		30,1
18	1.59 (m)	42	1.5 (m)	44,3
19	1.34 (m)	30,7	1.45 (m)	29,5
20		33,2		40,4
21	1.34 (m)	27,9	1.28 (m)	29,4
22	0.97 (m, H- $\beta$ ) 1.40 (m, H- $\alpha$ )	39,6	0.91 (m, H- $\beta$ ) 1.92 (m, H- $\alpha$ )	36,8
23	0.85 (d, <i>J</i> = 6.5 Hz)	6,9	0.80 (d, <i>J</i> = 6.5 Hz)	6,7
24	0.70 (s)	14,8	0.65 (s)	14,6
25	0.86 (s)	18	0.93 (s)	17,7
26	1.03 (s)	18,6	0.81 (s)	18,4
27	1.00 (s)	20,9	0.81 (s)	16,4
28	1.20 (s)	32,2	1.03 (s)	31,8
29	3.24 (sl)	74,9		182,3
30	1.01 (s)	25,9	1.15 (s)	31,6

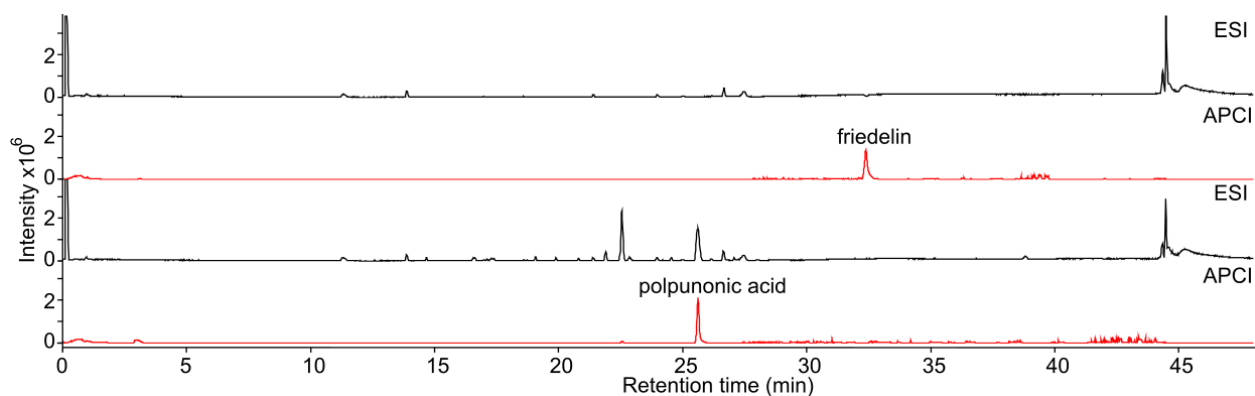




**Figure S1.** Comparison of the mass spectra of the TMS derivatized products of the *TwOSC* in tobacco leaf infiltration experiments with the spectra of TMS derivatized authentic standards.



**Figure S2.** Transient expression of *T. wilfordii* CYP712K1 and CYP712K2 to monitor their ability to oxygenate the core structures produced by TwOSC2 and TwOSC3. Two leaf discs ( $\varnothing$ 3 cm) of leaves co-expressing the genes encoding the indicated enzymes were extracted in 1 mL ethyl acetate. A 0.15 mL aliquot of this extract was dried, TMS derivatized and analyzed by GC-MS. None of the novel peaks 7-9 had mass spectra or retention times matching any of the analyzed authentic standards (betulinic acid, hederagenin, or oleanolic acid; not shown).



**Figure S3.** Comparison of LC-ESI-MS and LC-APCI-MS (base peak chromatogram) for the reliable quantification of friedelin and polpunonic acid standards. The same standards (30  $\mu\text{g}/\text{mL}$  of either friedelin or polpunonic acid) were analyzed using the same LC-MS equipment and program with the exception of ionization probe (ESI; black trace and APCI; red trace).