

Electronic Supplementary Information

Regio-specific prenylation of pterocarpanes by a membrane-bound prenyltransferase from *Psoralea corylifolia*

Junbin He, Zeyuan Dong, Zhi-min Hu, Yi Kuang, Jingran Fan, Xue Qiao* and Min Ye*

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Beijing 100191, China

* Corresponding authors. Tel.: +86 10 82801516. Fax: +86 10 82802024. Email address: qiaoxue@bjmu.edu.cn (X. Qiao), or yemin@bjmu.edu.cn (M. Ye).

Experimental

Species identification of *Psoralea corylifolia* by DNA barcoding

Dried seeds of *Psoralea corylifolia* L. were collected from Yunnan province in China and were identified by DNA barcoding analysis of the chloroplast *psbA-trnH* spacer. Genomic DNA was extracted using the plant genomic DNA Kit (Tiangen Biotech, China). The primers and the reaction conditions for the *psbA-trnH* barcode used for PCR amplification were as follows. Primers for *trnH-psbA* sequence were *psbA-F* and *trnH-R* (Table S1). PCR amplification was performed in 25 µL reaction mixtures containing approximately 50 ng of genomic DNA template (0.5 µL), 1×TransStart KD Plus Buffer, 0.4 µM of each primer, 200 µM dNTPs, and 1.0 unit *TransStart*[®] KD Plus DNA Polymerase (TransGen Biotech, Beijing, China). The PCR conditions were initial denaturation at 94 °C for 3 min, followed by 35 cycles at 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 45 s, and a final extension at 68 °C for 10 min. The products were then checked on 1.5% agarose gel electrophoresis, and purified using TIANquick Midi Purification Kit (Tiangen Biotech). Samples were sequenced using a 3730XL sequencer (Applied Biosystems, USA), and sequence alignment and analysis were performed with DNAMAN (version 8.0, lynn bio soft, USA) software. The sample was finally identified as *P. corylifolia* based on the sequence alignment (GenBank accession number GU396699, GU396701, GU396700 and GQ434957). The *psbA-trnH* spacer sequence was given below.

***psbA-trnH* sequence**

ATAATTTCCCTCTAGATCTAGCTGCGATCGACGCTCCATCTATTAATGGATAAAATTTTATC

TTAAATAAAAGTAGATGACTTTTTGAAAGTAAAGGAGCAATATCAAAAAAGTTGATATTGCT
 CCTTTACTTCTATTTGATTAGTAATCTTTTACTTTTATTTGATTAAAAATCTAACTATTTGAAAA
 TATTTATACAGACGTTTTTTATTTGTTTCACGATTCTTTAGCATTTCCTATCTTAGAAGGAA
 AAAAAAAAAAGAAGGCTAACGAAAGCGTAAAAATGGAGATAGAGATCATTTTTACATAGTAT
 AAGGGCGGATGTAGCCAAGTGGATCAAGGCAGT

ESI-MS, ¹H and ¹³C NMR data for prenylated product 1a

4-dimethylallylmaackiain (**1a**) (1.52 mg; amorphous powder, white): [α]_D²⁵ −120.0 (*c* 0.05, MeOH). ¹H NMR (DMSO-*d*₆, 400 MHz): δ 9.49 (1H, s, 3-OH), 7.09 (1H, d, *J* = 8.4 Hz, H-1), 6.97 (1H, s, H-7), 6.52 (1H, d, *J* = 8.4 Hz, H-2), 6.51 (1H, s, H-10), 5.94 (2H, two d, *J* = 0.9 Hz, −O−CH₂−O−), 5.51 (1H, d, *J* = 6.8 Hz, H-11a), 4.27 (1H, dd, *J* = 9.0, 6.3 Hz, H-6 α), 3.62–3.48 (2H, m, H-6 β , H-6a), 5.12 (1H, t, *J* = 7.2 Hz, H-2'), 3.19 (2H, d, *J* = 6.8 Hz, H-1'), 1.69 (3H, s, H-4'), 1.60 (3H, s, H-5'). ¹³C NMR (DMSO-*d*₆, 100 MHz): δ 155.8 (C-3), 154.0 (C-4a), 153.7 (C-10a), 147.4 (C-9), 140.9 (C-8), 128.5 (C-1), 118.5 (C-7a), 115.1 (C-4), 111.4 (C-1a), 108.9 (C-2), 105.4 (C-7), 100.9 (−O−CH₂−O−), 93.07 (C-10), 78.6 (C-11a), 66.2 (C-6), 40.1 (C-6a), 130.0 (C-3'), 122.9 (C-2'), 25.4 (C-5'), 21.8 (C-1'), 17.6 (C-4'). ESI-MS *m/z*: 353.0 [M+H]⁺. All of these data are in good agreement with the previously reported data (Figure S4, S5 and S9).¹

Table S1. PCR primers used in this experiment.

Primer name	Sequences (5'-3')
<i>psbA</i> -F	GTTATGCATGAACGTAATGCTC
<i>trnH</i> -R	CGCGCATGGTGGATTCAATCC
PcPT-F	AAATWGACAAGATAAAC
PcPT-R	GTCAGGTATATCCTTGYAC
PcPT-5'GSP	CAATCATGCACATTGCCGCAACCAATGG
PcPT-3'GSP	ATGGAAGAGATACCCATTGGTTGCGGCA
PcM4DT- <i>Bam</i> H I-F*	GGAGAAAAAACC CGATCC ATGGATTGGGGATTAGCAATAT
PcM4DT- <i>Xho</i> I-R*	CGGTACCAAGCTT ACTCGAG TTATCTGACAAAAGCCATGAGC
Δ TP ₄₁ -PcM4DT- <i>Bam</i> H I-F*	GGAGAAAAAACC CGATCC ATGGCTTCACAACACAAAAAGA

* Bold text represents restriction sites for subcloning.

Table S2. Accession numbers of plant PTs used for phylogenetic analysis.

Name	Accession number	Definition	Organism
PcM4DT	MH626730	pterocarpan prenyltransferase	<i>Psoralea corylifolia</i>
AhR4DT-1	AQM74172.1	resveratrol-4-dimethylallyltransferase	<i>Arachis hypogaea</i>
AhR3'DT-1	AQM74173.1	resveratrol-3'-dimethylallyltransferase	<i>Arachis hypogaea</i>
AhR3'DT-2	AQM74174.1	resveratrol-3'-dimethylallyltransferase	<i>Arachis hypogaea</i>
AhR3'DT-3	AQM74175.1	resveratrol-3'-dimethylallyltransferase	<i>Arachis hypogaea</i>
AhR3'DT-4	AQM74176.1	resveratrol-3'-dimethylallyltransferase	<i>Arachis hypogaea</i>
AaVTE2-1	ABB70124.1	homogentisate phytyltransferase VTE2-1	<i>Allium ampeloprasum</i>
AtHPT1	AAM10489.1	homogentisate phytylprenyltransferase	<i>Arabidopsis thaliana</i>
AtPPT1	BAB20818.2	p-hydroxybenzoate polyprenyltransferase	<i>Arabidopsis thaliana</i>
AtVTE2-1	AAM10489.1	homogentisate phytylprenyltransferase	<i>Arabidopsis thaliana</i>
AtVTE2-2	ABB70127.1	homogentisate phytyltransferase VTE2-2	<i>Arabidopsis thaliana</i>
CIP1a	BAP27988.1	umbelliferone 8-geranyltransferase	<i>Citrus limon</i>
CpVTE2-1	ABB70125.1	homogentisate phytyltransferase VTE2-1	<i>Cuphea pulcherrima</i>
CtiDT	AJD80983.1	isoliquiritigenin 3'-dimethylallyltransferase	<i>Cudrania tricuspidata</i>
GmG2DT	BAW32578.1	pterocarpan 2-dimethylallyltransferase	<i>Glycine max</i>
GmG4DT	BAH22520.1	pterocarpan 4-dimethylallyltransferase	<i>Glycine max</i>
GmVTE2-1	ABB70126.1	homogentisate phytyltransferase VTE2-1	<i>Glycine max</i>
GmVTE2-2	ABB70128.1	homogentisate phytyltransferase VTE2-2	<i>Glycine max</i>
GuA6DT	AIT11912.1	flavone prenyltransferase	<i>Glycyrrhiza uralensis</i>
HIPT-1	BAJ61049.1	aromatic prenyltransferase	<i>Humulus lupulus</i>
HIPT-2	AJD80255.1	aromatic prenyltransferase HIPT	<i>Humulus lupulus</i>
HvHGGT	AAP43911.1	homogentisic acid geranylgeranyl transferase	<i>Hordeum vulgare</i>
LaPT1	AER35706.1	genistein 3'-dimethylallyltransferase	<i>Lupinus albus</i>
LePGT-1	BAB84122.1	4-hydroxybenzoate geranyltransferase	<i>Lithospermum erythrorhizon</i>
LePGT-2	BAB84123.1	4-hydroxybenzoate geranyltransferase	<i>Lithospermum erythrorhizon</i>
HcPT	ALD84371.1	xanthone-specific prenyltransferase	<i>Hypericum calycinum</i>
MalDT	AJD80982.1	isoliquiritigenin 3'-dimethylallyltransferase	<i>Morus alba</i>
OsHGGT	AAP43913.1	homogentisic acid geranylgeranyl transferase	<i>Oryza sativa</i>
OsPPT1	BAE96574.1	p-hydroxybenzoate polyprenyltransferase	<i>Oryza sativa</i>
PcPT	BAO31627.1	umbelliferone 6-dimethylallyltransferase	<i>Petroselinum crispum</i>
PsPT1	AJW31563.1	umbelliferone dimethylallyl transferase	<i>Pastinaca sativa</i>
PsPT2	AJW31564.1	umbelliferone dimethylallyl transferase	<i>Pastinaca sativa</i>
SfFPT	AHA36633.1	flavonoid prenyltransferase	<i>Sophora flavescens</i>
SfG6DT	BAK52291.1	genistein 6-dimethylallyltransferase	<i>Sophora flavescens</i>
SfILD	BAK52290.1	isoliquiritigenin dimethylallyltransferase	<i>Sophora flavescens</i>
SfN8DT-1	BAG12671.1	naringenin 8-dimethylallyltransferase	<i>Sophora flavescens</i>
SfN8DT-2	BAG12673.1	naringenin 8-dimethylallyltransferase	<i>Sophora flavescens</i>
SfN8DT-3	BAK52289.1	naringenin 8-dimethylallyltransferase	<i>Sophora flavescens</i>
TaHGGT	AAP43912.1	homogentisic acid geranylgeranyl transferase	<i>Triticum aestivum</i>
TaVTE2-1	ABB70123.1	homogentisate phytyltransferase VTE2-1	<i>Triticum aestivum</i>
ZmVTE2-1	ABB70122.1	homogentisate phytyltransferase VTE2-1	<i>Zea mays</i>
AgGPPS2	AAN01134.1	geranyl diphosphate synthase	<i>Abies grandis</i>
AtGPPS	CAC16849.1	geranyl diphosphate synthase	<i>Arabidopsis thaliana</i>

PgFPS	AAV87903.1	farnesyl diphosphate synthase	<i>Panax ginseng</i>
AtFPS	AAB07248.1	farnesyl diphosphate synthase	<i>Arabidopsis thaliana</i>
GbGGPPS	AAQ72786.1	geranylgeranyl diphosphate synthase	<i>Ginkgo biloba</i>
SmGGPPS	ACR19637.1	geranylgeranyl diphosphate synthase	<i>Salvia miltiorrhiza</i>
CsPT1	Page J, Boubakir Z. 2010. Aromatic prenyltransferase from <i>cannabis</i> . Patent WO 2011017798 A1.		

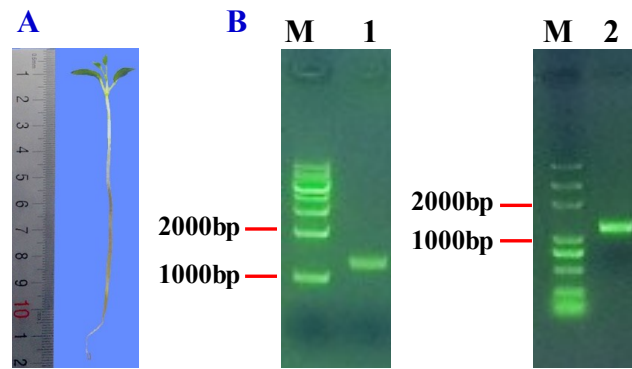


Figure S1. Four-week-old seedlings of *Psoralea corylifolia* (A) and agarose gel electrophoresis of RT-PCR products (B). M: DNA marker; 1: *PcM4DT*; 2: $\Delta TP41$ -*PcM4DT*.

PcM4DT	MDWCLAISSSSKAYSVTTGNIWRSKHVNNIHYFSSCITKASCPKRRKTCIEHNVLEFQRFESIGGCHNFSGCGGSGTYQENRKYA	85
GmG4DT	MDWCLAISSHPKPYSVTTGNIWRSKHVNNIHYFSSCITKASCPKRRKTCIEHNVLEFQRFESIDHHYKCIKIRGG.STYQENRKYA	84
GmG2DT	MDSCSVISST.SACSIITGNIWRSKHVNNIHYFSSCASKASKYKRRKTCIEHNVLEFQRFESIDHHYKRIKIRGG.ATYQENRKYA	83
Consensus	rd g iss s ttg nlwr kh t ni ss kas k tqie n lr q sl h y gg tyq c k	
PcM4DT	VKAVYDQPKDFELBASNPKNILLASAKKFIAPFYMSYPYIMICITLCAFCSSALAVEKILSDISLHFIGVLQAVIECLFIETIYLS	170
GmG4DT	VKATSKQPLGFELBASNPKNILLASAKKFIAPFYMSYPYIMICITLCAFCSSALAVEKILSDISLHFIGVLQAVIECLFIETIYIC	169
GmG2DT	VKATIFEFTELSLFCASNFEFNVDSAKKFIAPFYMSYPYIMICITLCAFCSSALAVEKILSDISLHFIGVLQAVIECLFIETIYIC	168
Consensus	svka e asnp n ds k l fy f ypy ri i lca ss lavek sdls f igvlq p lf	
PcM4DT	GVNCLVDFEIDKINKFELFASQCFSEKTCWIMSPAFLAISFGFTWITGSWFLIWNIVVIASSWTAYSIDVFFMRWRKMEIVFAM	255
GmG4DT	GVNCLVDFEIDKINKFELFASQCFSEKTCWIMSPAFLAISFGFTWITGSWFLIWNIVVIASSWTAYSIDVFFMRWRKMEIVFAM	254
GmG2DT	VVNQVDFEIDKINKFELFASQCFSEKTCWIMSPAFLAISFGFTWITGSWFLIWNIVVIASSWTAYSIDVFFMRWRKMEIVFAM	253
Consensus	vng d eickinkp lp asgg sf t v a l sf gswpli n v s w ysi vp rwrk p a	
	NQxxDxxxD	
PcM4DT	CMATLGLALPISFHHMCQTFVLQRFICGFERSLGHVAFMTFYSGIALSKDITLVEGDAEHGINSFSVILGCKRIEHWICVSIFF	340
GmG4DT	CMISTGLALPISFHHMCQTFVLQRFICGFERSLGHVAFMTFYSGIALSKDITLVEGDAEHGINSFSVILGCKRIEHWICVSIFF	339
GmG2DT	CTISVAFILPITFELHMCQTFVLQRFICGFERSLGHVAFMTFYSGIALSKDITLVEGDAEHGINSFSVILGCKRIEHWICVSIFF	338
Consensus	usc l w lpi f hrqt vl rpi fprsl f v fm fys g als kdi gv qk gi lgqk fwic fe	
	KD(I/L)xDx(E/D)GD	
PcM4DT	MAFGVGLLGCVISPEHWIKITITGLGNVILCSIIHYCAKSFDTITKASTGCFYMYIWKLIYAFILIMAFV	409
GmG4DT	MAFGVGLLGCASCSEHWIKITITGLGNVILCSIIHYCAKSFDTITKASTGCFYMYIWKLIYAGFFIMALI	408
GmG2DT	MAFGVALLGCATSSYLWIKITITGLGNVILCSIIHYCAKSFDTITKASTGCFYMYIWKLIYAYFFIMALI	407
Consensus	mafvg l g w ki tg g l sil yqaks l k st sfym iwkllya lra	

Figure S2. A multiple alignment of the amino acid sequences of PcM4DT and other plant pterocarpan prenyltransferases. GmG4DT: pterocarpan 4-dimethylallyltransferase from *Glycine max* (Genebank accession number BAH22520); GmG2DT: pterocarpan 2-dimethylallyltransferase from *Glycine max* (Genebank accession number BAW32578). Identical amino acid are shown in white on a black background. The conserved aspartate-rich motifs NQxxDxxxD and KD(I/L)xDx(E/D)GD are highlighted with red box.

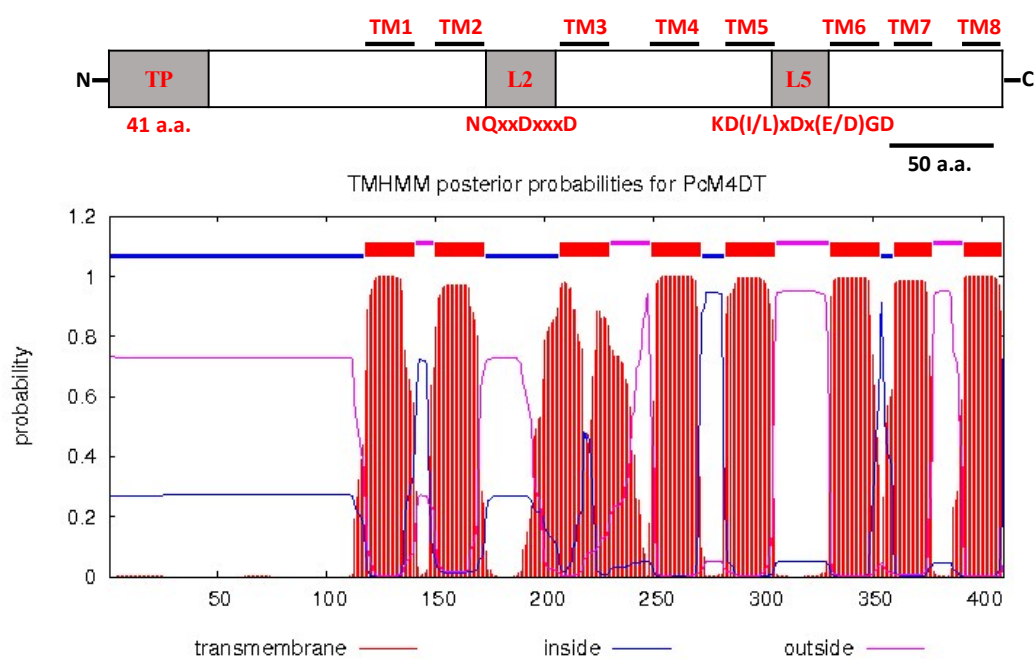


Figure S3. Putative transmembrane domains of PcM4DT were predicted by TMHMM Server program version 2.0. TP: Transit peptide; TM: transmembrane α -helix; L2: loop number 2; L5: loop number 5.

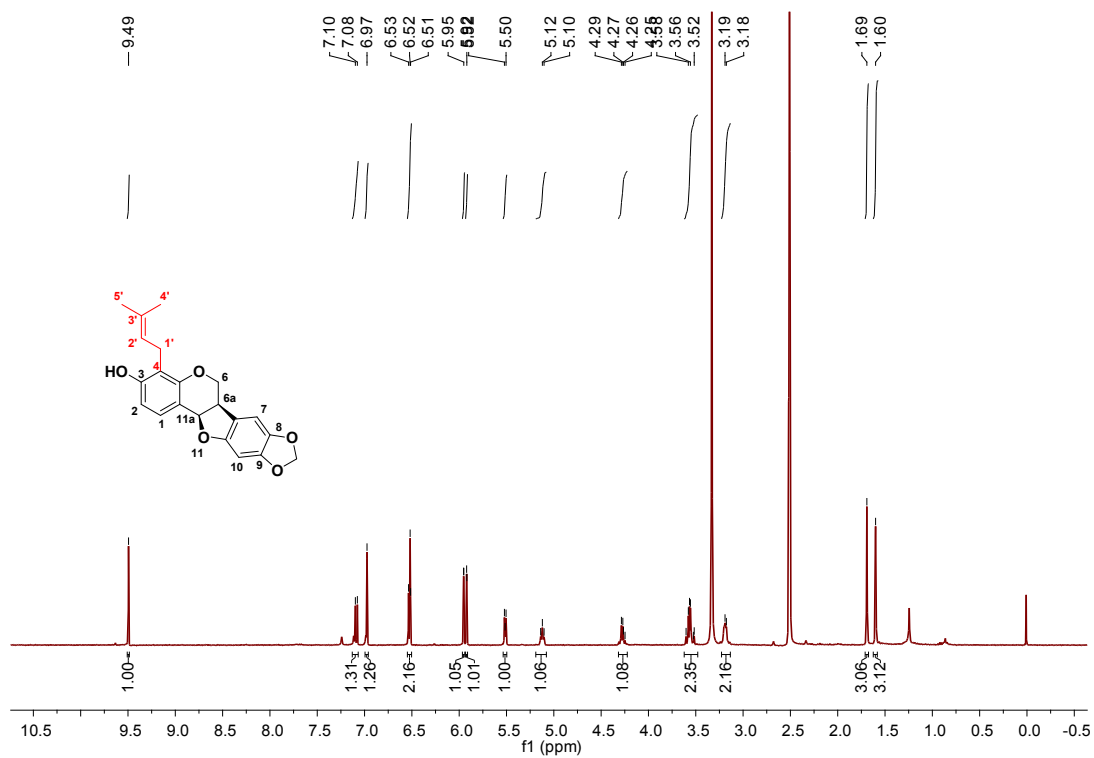


Figure S4. ^1H NMR spectrum of **1a** in $\text{DMSO}-d_6$ (400 MHz).

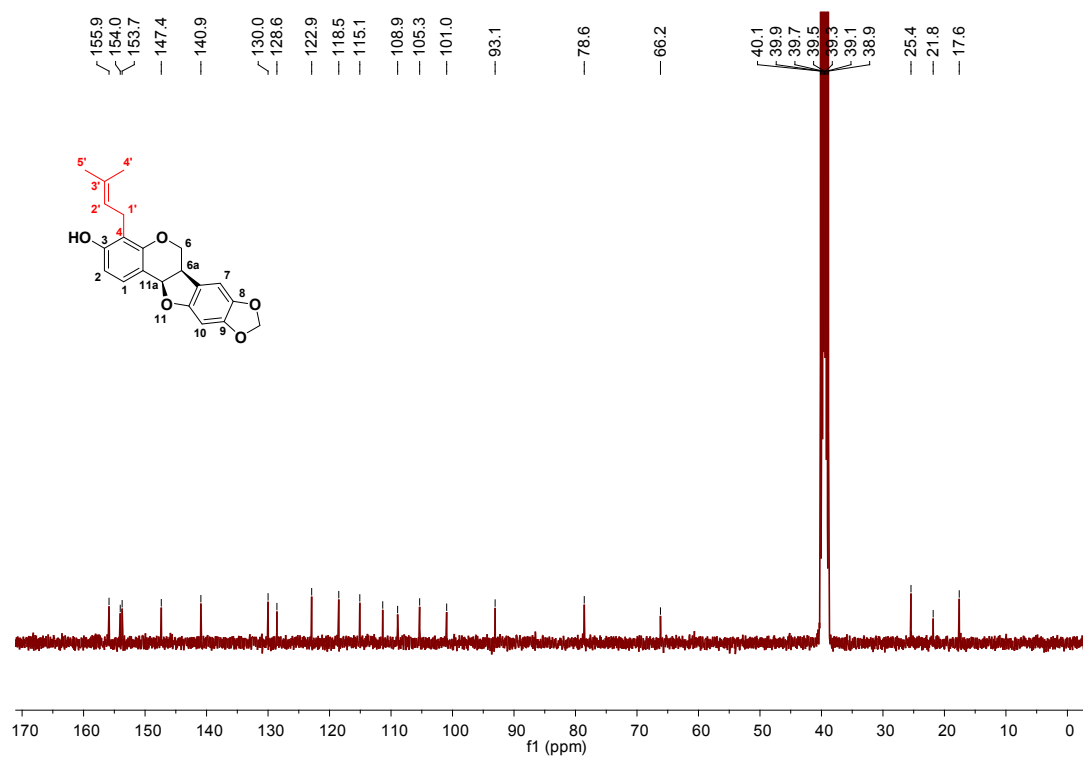


Figure S5. ¹³C NMR spectrum of **1a** in DMSO-*d*₆ (100 MHz).

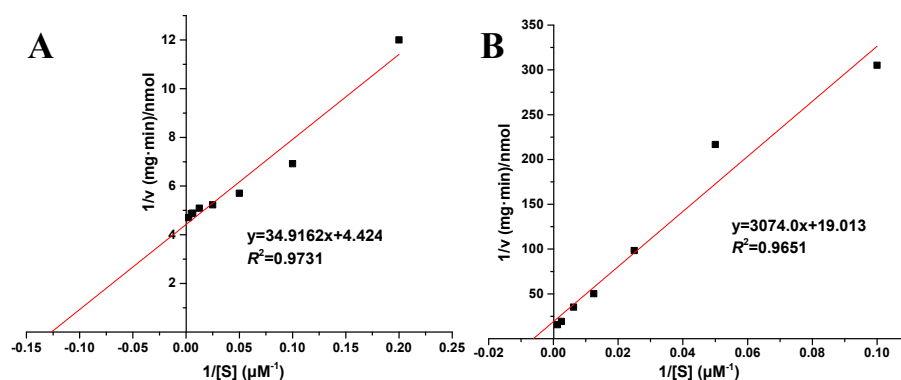


Figure S6. Determination of kinetic parameters for recombinant PcM4DT. The apparent K_m values for maackiain (**1**) were determined with varying concentrations (5–400 μM) using 400 μM DMAPP as prenyl donor (A), while that for DMAPP were determined with various concentrations (10–800 μM) using 400 μM **1** as prenyl acceptor (B). The values are averages of three parallel experiments.

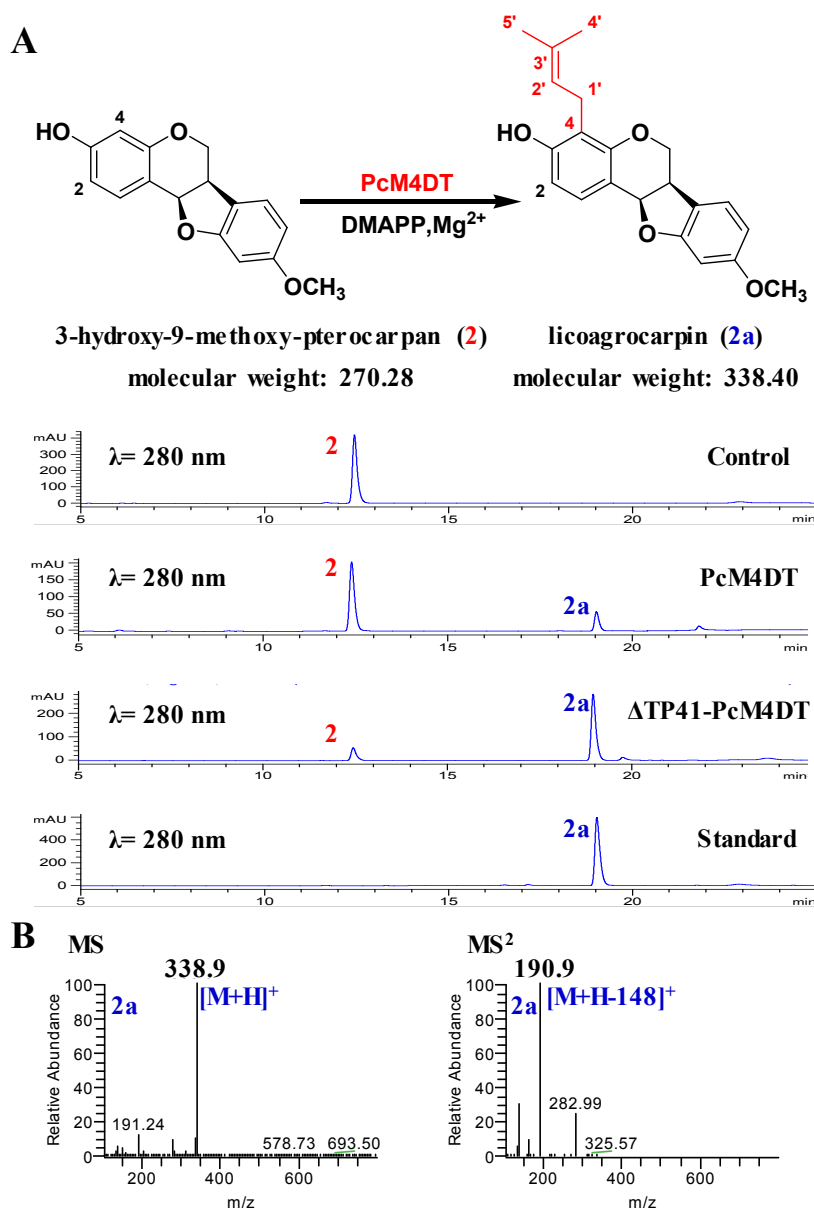


Figure S7. Enzymatic characterization of recombinant PcM4DT. (A) HPLC analysis of the enzymatic reaction mixture from microsomal fractions containing 3-hydroxy-9-methoxy-pterocarpan (**2**), DMAPP, and Mg²⁺. (B) MS analysis of prenylated product **2a** from ΔTP₄₁-PcM4DT-expressing microsomes.

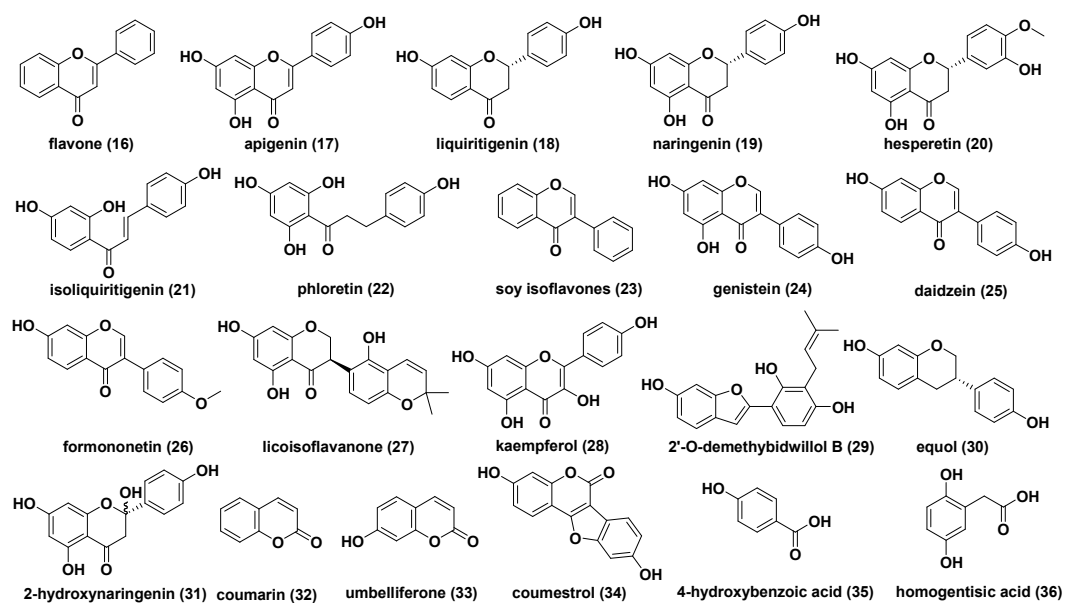


Figure S8. Aromatic compounds that are not prenylated by PcM4DT.

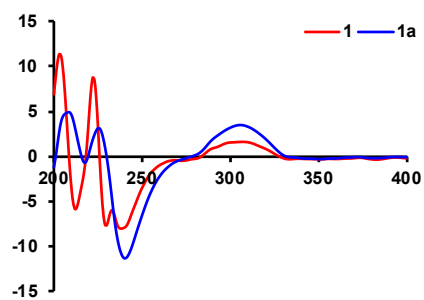


Figure S9. The CD spectra of compound **1** (red) and **1a** (blue).

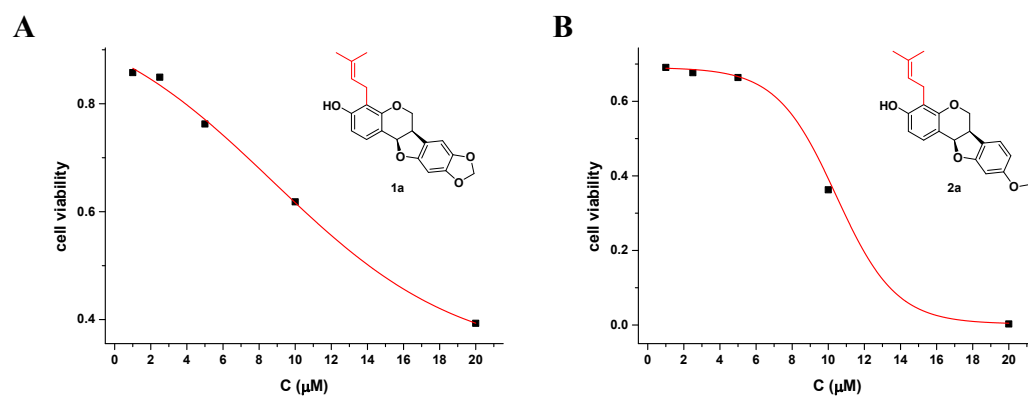


Figure S10. IC_{50} curves of compounds **1a** and **2a**.

References

- 1 G. L. DaSilva, F. J. D. Matos and E. R. Silveira, *Phytochemistry*, 1997, 46, 1059-1062.