Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supporting information

Characterization of a coumarin *C*-/*O*-prenyltransferase and a quinolone *C*-prenyltransferase from *Murraya exotica*

Na Li, [a,b] Xiao Liu,[a] Mingliang Zhang,[a,b] Zekun Zhang,[a,b] Beibei Zhang,[a,b] Xiaohui Wang,[a] Juan Wang,[a] Pengfei Tu,[a,c] Yong Jiang*[c] and She-po Shi*[a]

[a]Modern Research Center for Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100102, China.

[b]School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 100102, China.

[c]State Key Laboratory of Natural and Biomimetic Drugs, Peking University, Beijing 100191, China.

*Corresponding authors: E-mail: shishepo@163.com, yongjiang@bjmu.edu.cn

Table of contents

Table S1 Plant PTs used for local tBlastn analysis.

Enzyme	Species	Accession number	Acceptor	Donor	Regio-specificity
ClPT1a	*Citrus limon*	BAP27988.1	Umbelliferone Esculetin 5,7-Dihydroxycoumarin 5-Methoxy-7-hydroxycoumarin	GPP	C-8, C-6, C-5
PsPT1	*Pastinaca sativa*	AJW31563.1	Umbelliferone	DMAPP	C-8, C-6
PsPT2	*Pastinaca sativa*	AJW31564.1	Umbelliferone	DMAPP	C-8, C-6
PcPT	*Petroselinum crispum*	BAO31627.1	Umbelliferone	DMAPP	C-6, C-8
CpPT1	*Citrus paradisi*	LC557129.1	5,7-Dihydroxycoumarin 5-Hydroxy-7-methoxycoumarin Bergaptol Xanthotoxol 8-Hydroxybergapten	GPP	*O*-prenyltransferase 5-*O*/7-*O*/8-*O*
CpPT3	*Citrus paradisi*	LC557131.1	Umbelliferone	GPP	C-8, C-6 (supposed)

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

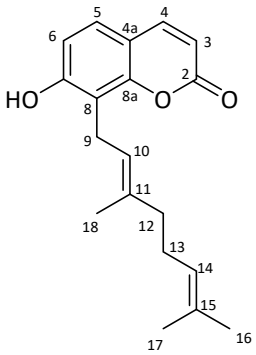
Species	**Enzyme**	**Accession number** （GenBank）
Murraya exotica	MePT1	ON652845
Murraya exotica	MePT2	ON652846
Arabidopsis thaliana	AtPPT1	BAB20818.2
Arabidopsis thaliana	AtVTE2-1	AAM10489.1
Arabidopsis thaliana	AtVTE2-2	ABB70127.1
Artemisia capillaris	AcPT1	LC425153.1
Arachis hypogaea	AhR4DT-1	AQM74172.1
Arachis hypogaea	AhR3'DT-1	AQM74173.1
Arachis hypogaea	AhR3'DT-2	AQM74174.1
Arachis hypogaea	AhR3'DT-3	AQM74175.1
Arachis hypogaea	AhR3'DT-4	AQM74176.1
Arnebia euchroma	AePGT	ABD59796.2
Arnebia euchroma	AePGT4	ANC67957.1
Arnebia euchroma	AePGT6	ANC67959.1
Angelica keiskei	AkPT1	LC557134.1
Cannabis sativa	CsPT1	DAC76711.1
Cannabis sativa	CsPT3	DAC76713.1
Cannabis sativa	CsPT4	DAC76710.1
Cannabis sativa	CsPT8	XP_030491742 .1
Chlamydomonas reinhardtii	CrHST (CrVTE2-2)	CAL01105.1
Citrus limon	ClPT1a	BAP27988.1
Citrus limon	ClPT1b	BAP27989.1
Clitorea ternatea	CtHPT	ALR81009.1
Cudriana tricuspidata	CtIDT	AJD80983.1
Citrus paradisi	CpPT1	LC557129.1
Citrus paradisi	CpPT2	LC557130.1
Citrus paradisi	CpPT3	LC557131.1
Citrus micrantha	CmiPT1a	LC557132.1
Citrus micrantha	CmiPT1b	LC557133.1
Ficus carica	FcPT1a	LC369744.1
Glycine max	GmC4DT	BAW32575.1
Glycine max	GmG2DT	BAW32578.1
Glycine max	GmG4DT	BAH22520.1
Glycine max	GmIDT1	BAW32576.1
Glycine max	GmVTE2-1	ABB70126.1
Glycine max	GmVTE2-2	ABB70128.1
Glycyrrhiza uralensis	GuA6DT	AIT11912.1
Glycyrrhiza uralensis	GuILDT	AMR58303.1

Species	Enzyme	Accession number (GenBank)
Hordeum vulgare	HvHGGT	AAP43911.1
Humulus lupulus	HlPT-1	BAJ61049.1
Humulus lupulus	HlPT2	AJD80255.1
Hypericum calycinum	HcPT8px	AZK16226.1
Hypericum calycinum	HcPTpat	AZK16227.1
Hypericum sampsonii	HsPT8px	AZK16224.1
Hypericum sampsonii	HsPTpat	AZK16225.1
Ipomoea batatas	IbHPT	ALG62646.1
Lactuca sativa	LsHPT	ACN78585.1
Lithospermum erythrorhizon	LePGT1	BAB84122.1
Lithospermum erythrorhizon	LePGT2	BAB84123.1
Lotus japonicus	LjG6DT (LjPT1)	ARV85585.1
Lupinus albus	LaPT1	AER35706.1
Lupinus albus	LaPT2	MH298812.1
Malus domestica	MdHPT	ACT75571.1
Morus alba	MaIDT	AJD80982.1
Morus alba	MaOGT	AXN57307.1
Oryza sativa	OsHGGT	AAP43913.1
Oryza sativa	OsHPT (RTD1)	BAS98961.1
Oryza sativa	OsPPT1	BAE96574.1
Oberonia myosurus	ObPPT1	n.a.
Oberonia myosurus	ObPPT2	n.a.
Pastinaca sativa	PsPT1	AJW31563.1
Pastinaca sativa	PsPT2	AJW31564.1
Petroselinum crispum	PcPT	BAO31627.1
Psoralea corylifolia	PcM4DT	AYV64464.1
Rhododendron dauricum	RdPT1	BBD96134.1
Sophora flavescens	SfFPT	AHA36633.1
Sophora flavescens	SfG6DT	BAK52291.1
Sophora flavescens	SfiLDT	BAK52291.1
Sophora flavescens	SfN8DT-1	BAG12671.1
Sophora flavescens	SfN8DT-2	BAG12673.1
Sophora flavescens	SfN8DT-3	BAK52289.1
Triticum aestivum	TaHGGT	AAP43912.1
Zea mays	ZmHST	ONM24667.1
Zea mays	ZmHPT	ABB70122.1

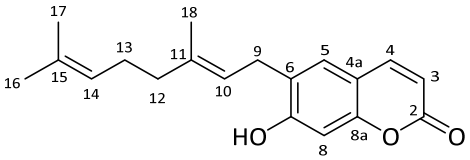
n.a., not available

Table S3 ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR data of compounds **P1a** and **P1b** (δ in ppm, in $CDCl_3$).

Position	**P1a**		**P1b**	
	δH (*J* in Hz)	δC	δH (*J* in Hz)	δC
2		161.8		162.0
3	6.23 (d, 9.6)	113.5	6.23 (d, 9.6)	112.9
4	7.62 (d, 9.6)	144.4	7.62 (d, 9.6)	144.0
4a		112.9		112.6
5	7.22 (d, 8.4)	126.9	7.20 (s)	128.6
6	6.79 (d, 8.4)	112.7		125.3
7		158.8		158.6
8		114.8	6.86 (s)	103.6
8a		153.2		154.5
9	3.62 (d, 7.2)	22.2	3.39 (d, 7.2)	29.0
10	5.26 (t, 7.2)	120.3	5.32 (t, 7.2)	121.0
11		140.3		139.5
12	2.08 (m)	39.9	2.14 (m)	39.9
13	2.06 (m)	26.5	2.12 (m)	26.6
14	5.02 (t, 6.6)	123.8	5.09 (t, 6.0)	123.9
15		132.3		132.3
16	1.57 (s)	25.9	1.61 (s)	26.0
17	1.83 (s)	17.9	1.77 (s)	18.0
18	1.65 (s)	16.5	1.70 (s)	16.5



P1a



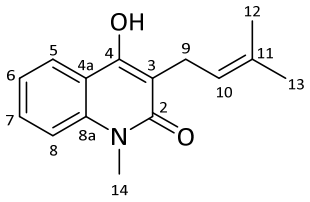
P1b

Table S4 Selectivity of MePT1 and MePT2 for 20 accepters and 4 prenyl donors.

NO.	Compound	MePT1				MePT2			
		DMAPP	GPP	FPP	GGPP	DMAPP	GPP	FPP	GGPP
1	Umbelliferone	+	+	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
2	5,7-Dihydroxycoumarin	N.D.	+	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
3	Coumarin	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
4	7-Methoxycoumarin	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
5	7-Hydroxy-6-methoxcoumarin	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
6	5,7-Dimethoxycoumarin	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
7	Xanthotoxol	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
8	Bergaptol	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
9	4-Hydroxy-*N*-methylquinolone	N.D.	N.D.	N.D.	N.D.	+	N.D.	N.D.	N.D.
10	2,4-Dihydroxy-quinolone	N.D.	N.D.	N.D.	N.D.	+	N.D.	N.D.	N.D.
11	1-methoxy-3-Methyl-carbazole	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
12	2-hydroxy-3-Methyl-carbazole	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
13	3-Methyl-carbazole	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
14	Indole	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
15	4-Hydroxybenzaldehyde	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
16	*p*-Coumaric acid	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
17	3-Hydroxybenzoic acid	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
18	4-Hydroxybenzoic acid	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
19	Kaempferol	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
20	Kaempferide	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

Table S5 ^{1}H (700 MHz) and ^{13}C (175 MHz) NMR data of compound **P9** (δ in ppm, in CD_3OD).

Position	**P9**	
	δH (*J* in Hz)	δC
2		167.0
3		Not found
4		Not found
4a		110.4
5	8.16 (dd, 7.7, 0.7)	122.0
6	7.17 (t, 7.7)	125.9
7	7.49 (t, 7.7)	130.6
8	7.42 (d, 7.7)	115.1
8a		140.3
9	3.38 (d, 7.0)	26.1
10	5.26 (t, 7.0)	115.1
11		130.8
12	1.65 (s)	24.7
13	1.81 (s)	18.4
14	3.67 (s)	29.9



P9

Table S6 PCR primers used in this experiment.

Primer name	**Sequences (5'–3')**
MePT1-Fw	ATTCAAATGCATTCGAGTTCAAGC
MePT1-Rv	TCATCGTAAGAAATGAATAAGGAGATATTCAG
MePT2-Fw	ATGCTGCTTCAAATGAATTCTTCTGC
MePT2-Rv	TCAACGTATGAAGTGCATAAGGA
ΔTP$_{49}$-*MePT1-Apa*I-Fw	ATACGACTCACTATA**GGGCCC**ATGTGCACTCAAAGTGATTCATTTT
*MePT1-Xho*I-Rv	GCGGTACCAAGCTTA**CTCGAG**TCATCGTAAGAAATGAATAAGGAGATATT
ΔTP$_{89}$-*MePT2-Apa*I-Fw	ATACGACTCACTATA**GGGCCC**ATGGCGTCACAAGATGATAATATTAAA
*MePT2-Xho*I-Rv	GCGGTACCAAGCTTA**CTCGAG**TCAACGTATGAAGTGCATAAGGAAG
MePT1-qpcr-Fw	ATTCAAATGCATTCGAGTTCAAGCT
MePT1-qpcr-Rv	TCATCGTAAGAAATGAATAAGGAGATATTCAG
MePT2-qpcr-Fw	ATGCTGCTTCAAATGAATTCTTCTGC
MePT2-qpcr-Rv	TCAACGTATGAAGTGCATAAGGA
Actin-Fw	TACGTCGCCCTTGACTTCG
Actin-Rv	AGACTCCATACCTAAGAAG

* Bold text represents restriction sites for subcloning.

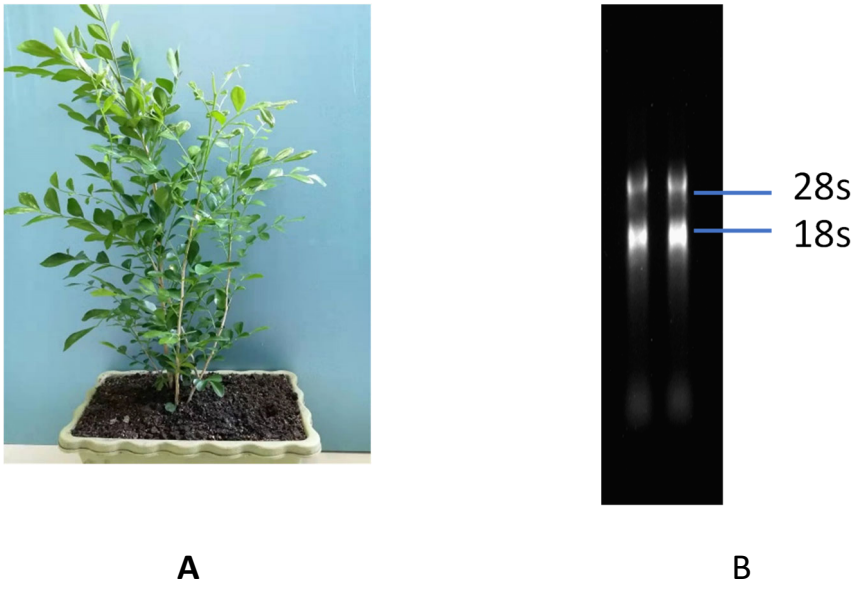


Figure S1 *M. exotica* used in this study (**A**) and agarose gel electrophoresis of RNA (**B**).

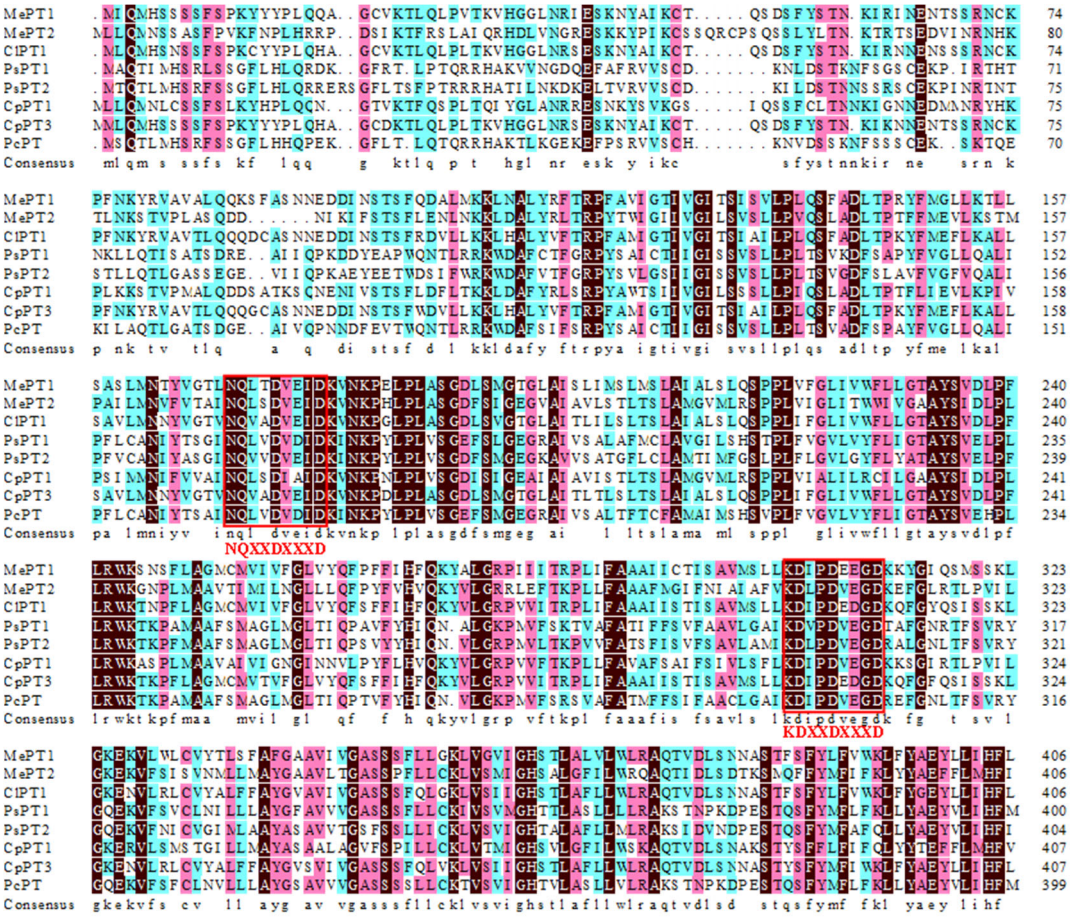
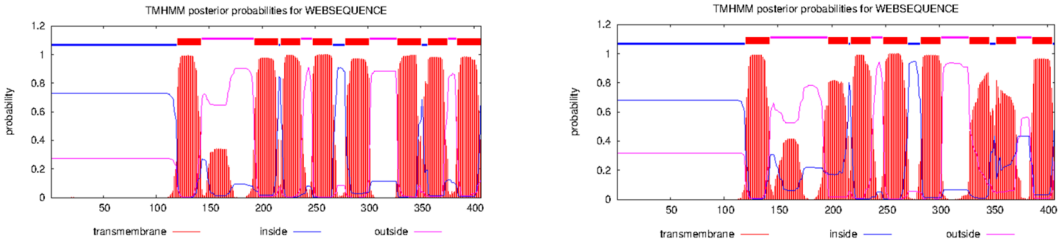


Figure S2 ClustalW multiple sequence alignments of MePT1, MePT2 and related PTs. Identical amino acids were shown in white on a black background. The conserved aspartate-rich motifs NQxxDxxxD and KDxxDx(E/D) GD are highlighted with red box.

A



MePT1 **MePT2**

B

Name	Score	TP	TP length
MePT1	0.475	-	49
MePT2	0.508	Y	89

C

PT	ClPT1	PsPT1	PsPT2	PcPT	CpPT1	CpPT3
MePT -1	85.50%	38.82%	35.87%	38.82%	49.52%	84.07%
MePT-2	49.15%	39.42%	40.15%	41.22%	64.92%	49.39%

Figure S3 In silico analysis of MePT1 and MePT2 polypeptides. (**A**) Transmembrane (TM) regions of MePT1 and MePT2 predicted by TMHMM. (**B**) Transit peptides (TPs) of MePT1 and MePT2 predicted by ChloroP. (**C**) Comparison of the amino acid identity of MePT1, MePT2, and six known PTs from *Rutaceae* and *Apiaceae* plants.

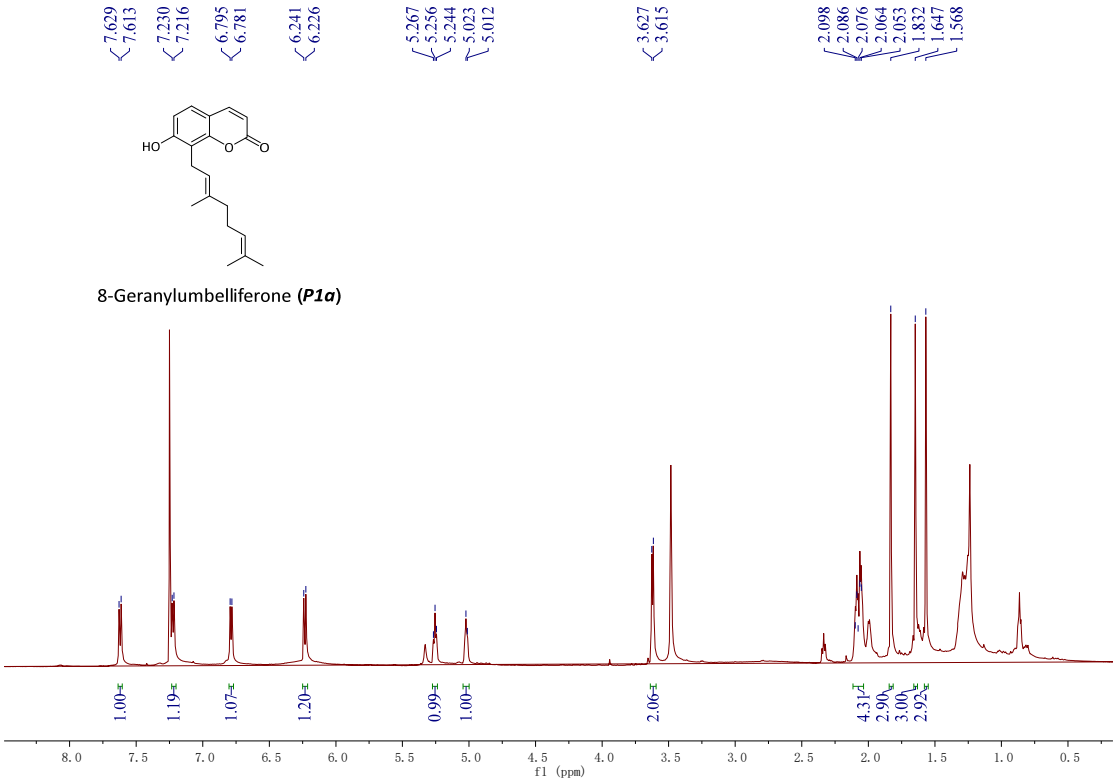


Figure S4 The ^{1}H NMR spectrum of **P1a** (600 MHz, in $CDCl_3$)

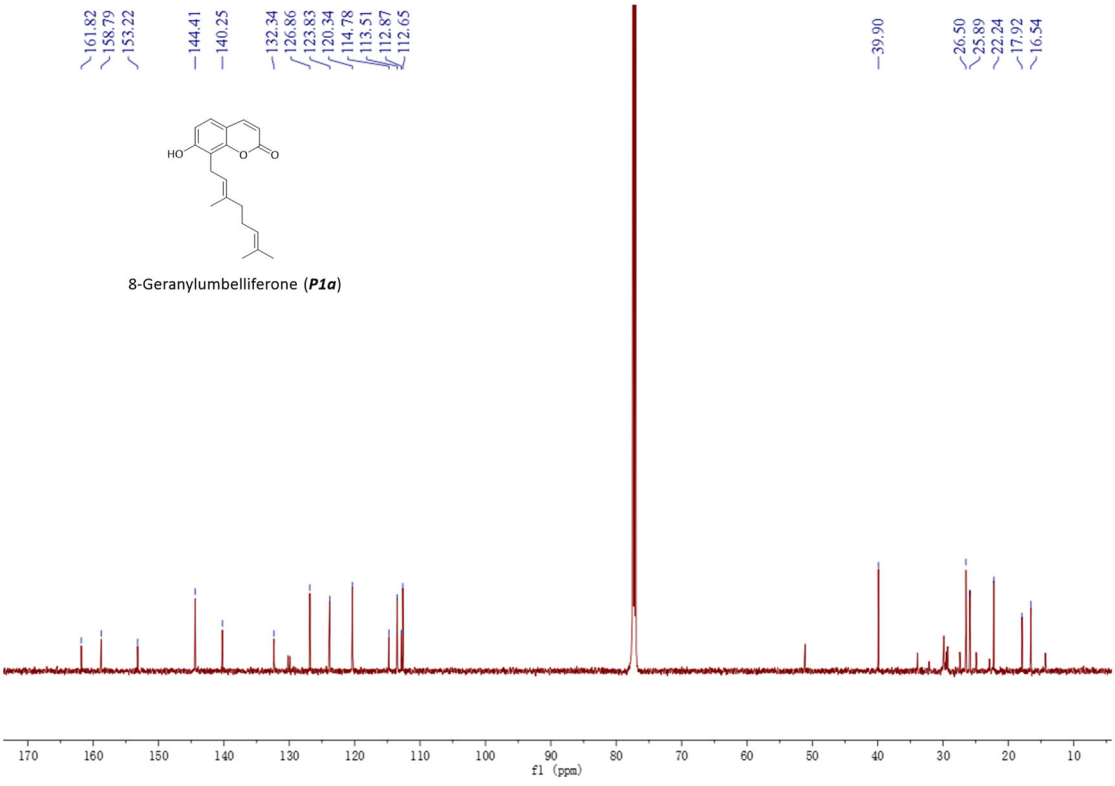


Figure S5 The ^{13}C NMR spectrum of **P1a** (150 MHz, in $CDCl_3$)

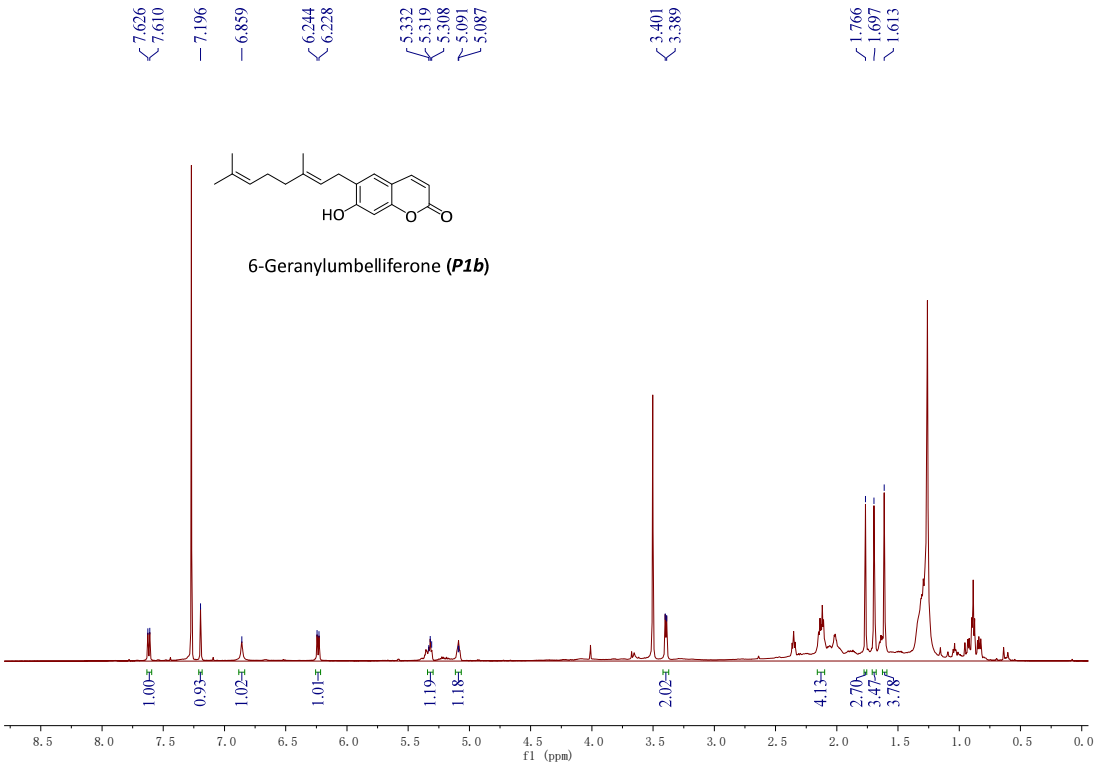


Figure S6 The ^{1}H NMR spectrum of **P1b** (600 MHz, in $CDCl_3$)

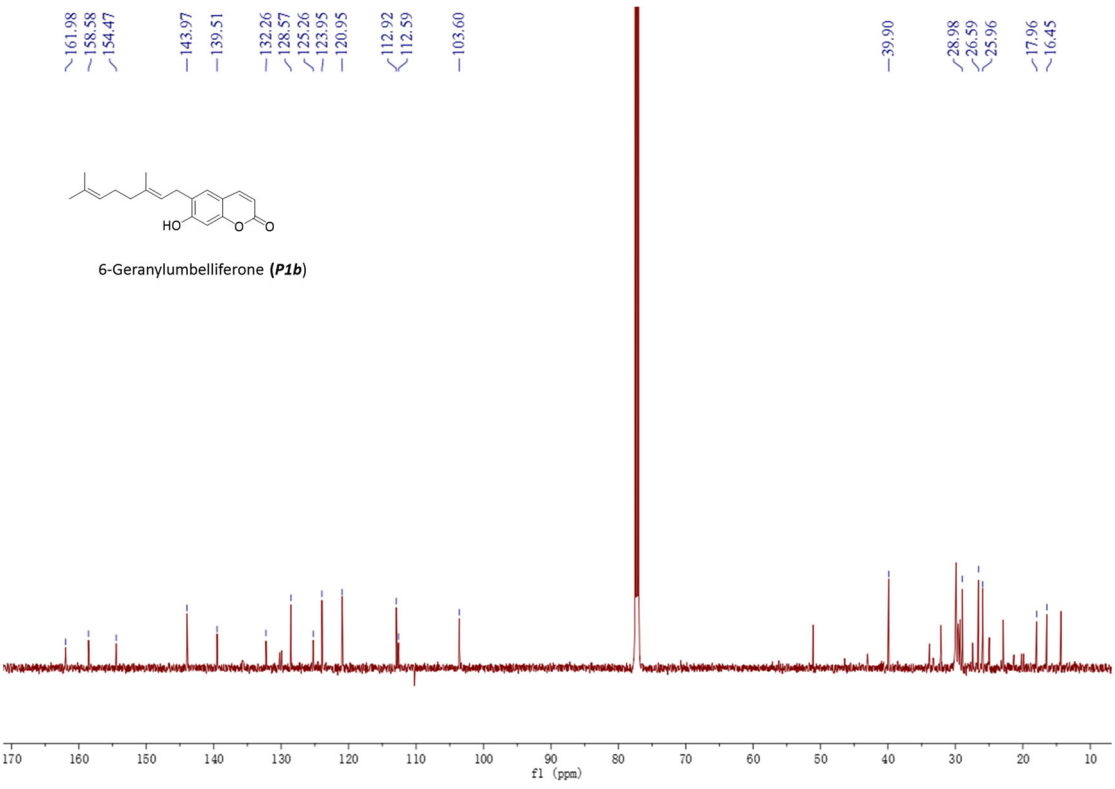


Figure S7 The ^{13}C NMR spectrum of **P1b** (150 MHz, in $CDCl_3$)

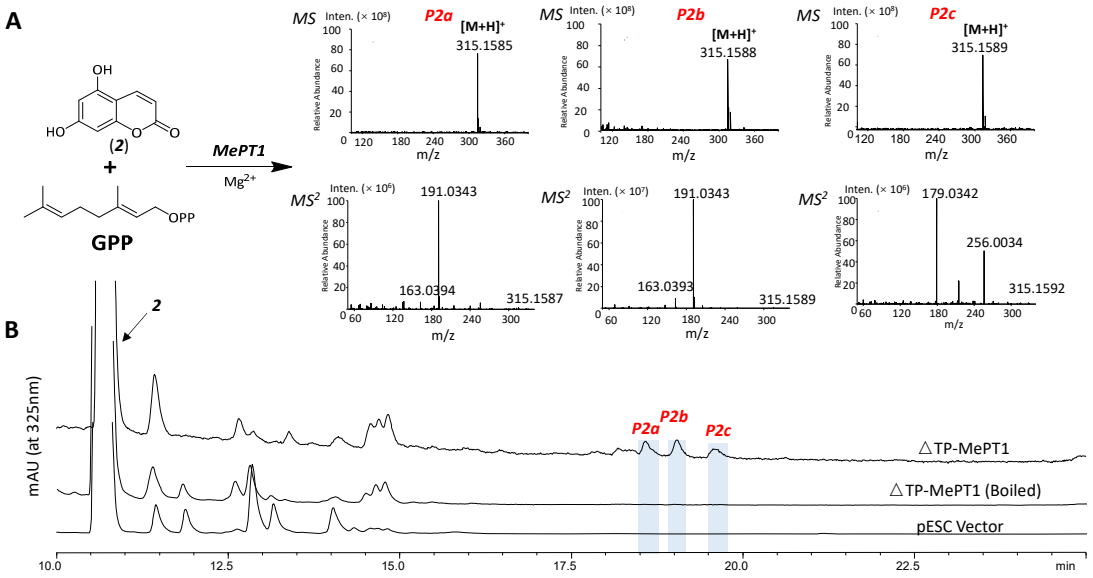


Figure S8 Enzymatic formation of three geranylated products (**P2a**, **P2b**, and **P2c)** by MePT1 from 5,7-dihydroxycoumarin (**2**) and GPP. (**A**) MS analysis of the reaction products **P2a**, **P2b** and **P2c**. (**B**) HPLC chromatograms for the formation of **P2a**, **P2b** and **P2c**. Limited by the low yields, further preparation of enough amount of samples for NMR analysis are challenging, which prevent us from the unambiguous assignments of the exact prenylated positions in **P2a**, **P2b**, and **P2c**. However, In the MS^2 spectra of the products **P2a** and **P2b**, the observation of the typical fragment ion peak at *m*/*z* 191 contributing to the loss of *C*-geranyl moiety (one carbon left on the backbone) suggested that **P2a** and **P2b** are two *C*-geranylated products. In contrast, the presence of the typical fragment ion peak at *m*/*z* 179 contributing to the loss of the entire geranyl moiety suggested that **P2c** is an *O*-geranylated product from 5,7-dihydroxycoumarin. ΔTP-MePT1: the truncated protein of MePT1 with the *N*-terminal transit peptides removed.

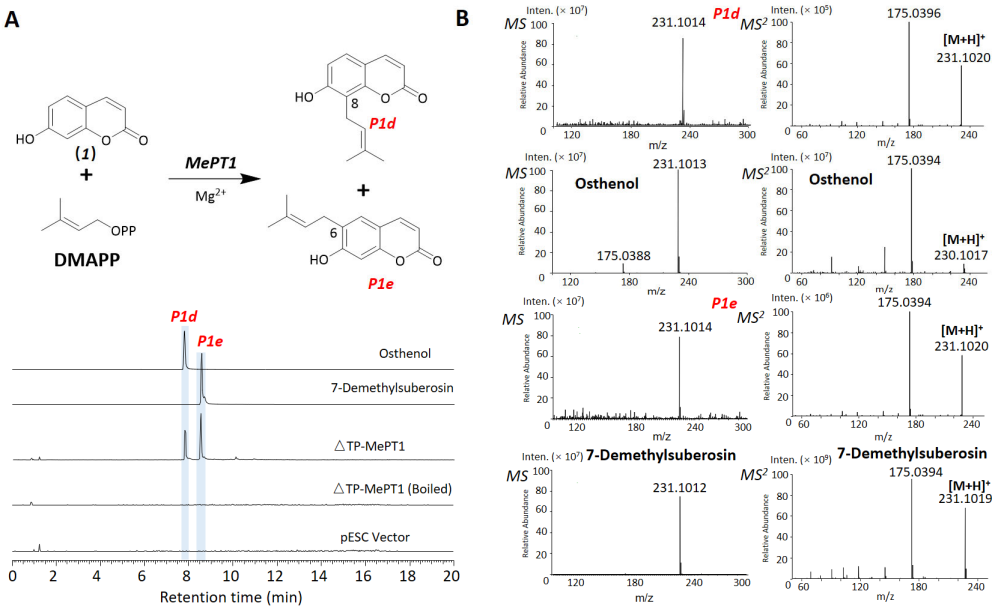


Figure S9 Enzymatic formation of 8-*C*-dimethyallylumbelliferone (**P1d**) and 6-*C*-dimethyallylumbelliferone (**P1e**) by MePT1 from umbelliferone (**1**) and DMAPP. (**A**) Structures and extract ion chromatograms (EICs). (**B**) The MS and MS^2 spectra of the reaction products **P1d** and **P1e**, the authentic compounds 7-demethylsuberosin and osthenol. ΔTP-MePT1: truncated protein of MePT1 with the *N*-terminal transit peptides removed.

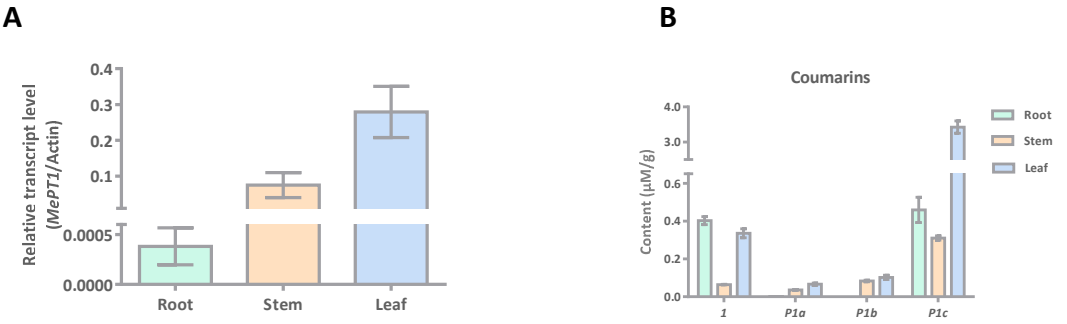


Figure S10 Real-time PCR analysis of *MePT1* in the different tissues of *M. exotica* (**A**) and the contents of substrate **1**, enzymatic products **P1a**, **P1b**, and **P1c** in the different tissues of *M. exotica* (**B**). Relative expression levels of *MePT1* were normalized by the actin gene as a reference (*n* = 3). Compounds were quantified by LC/MS. Values are means ± SD of three biological replicates.

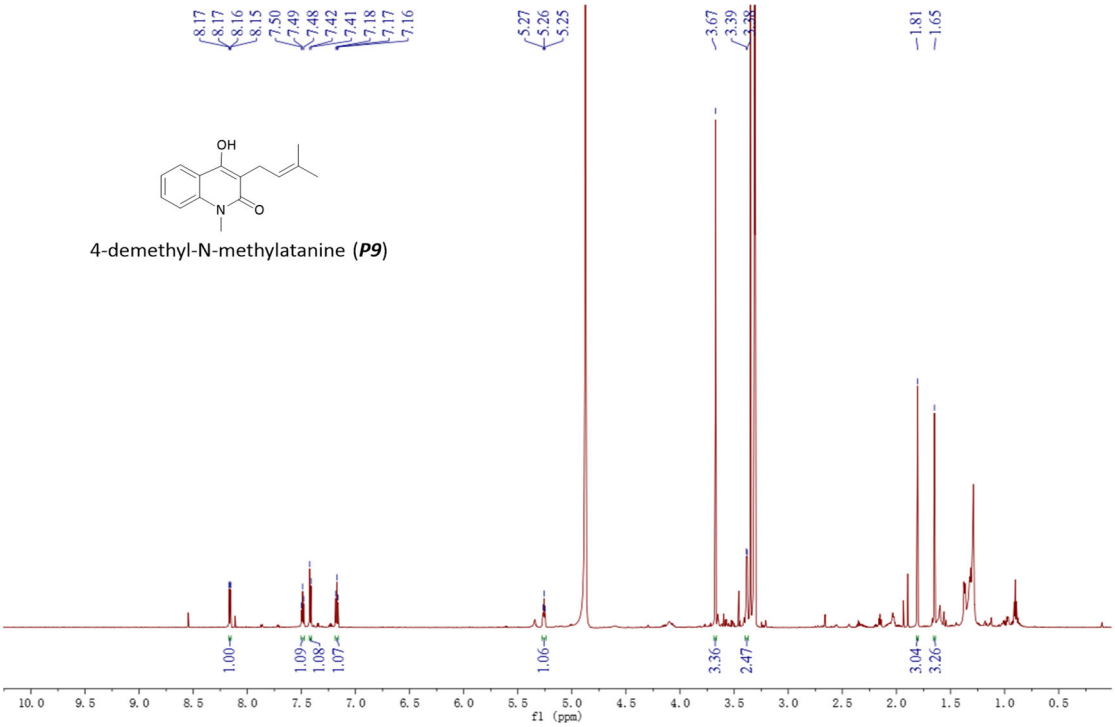


Figure S11 The ^{1}H NMR spectrum of **P9** (700 MHz, in CD_3OD)

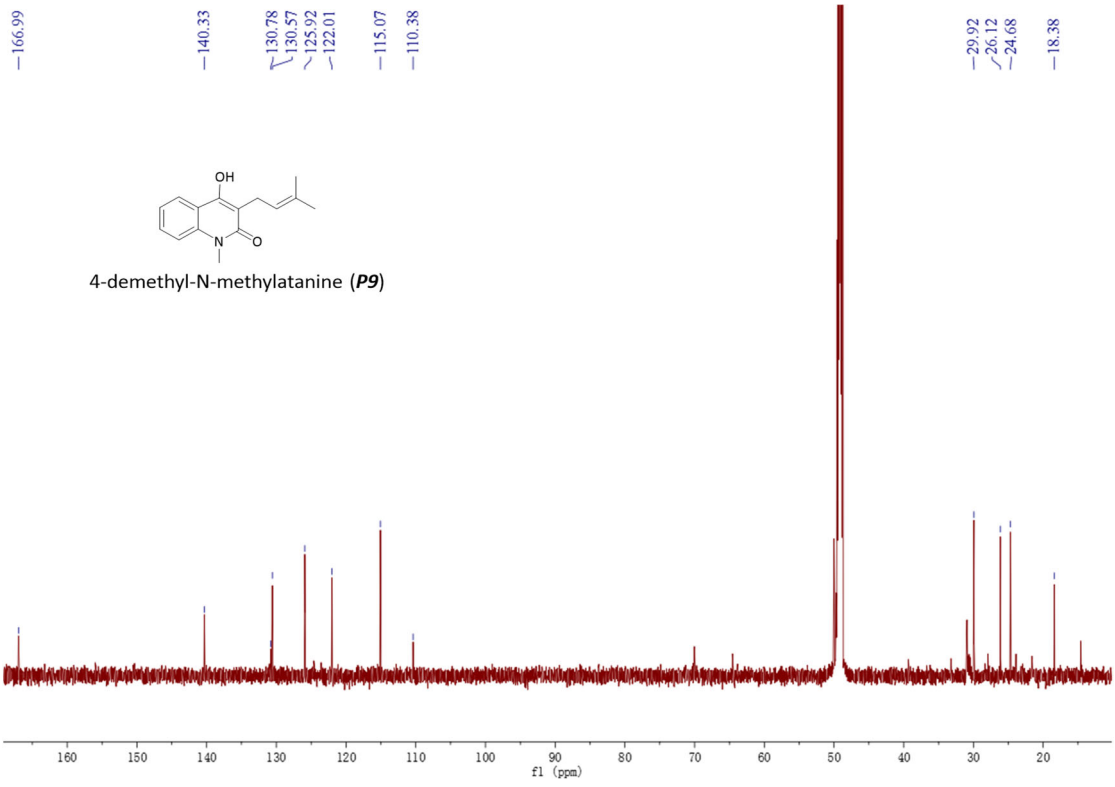


Figure S12 The ^{13}C NMR spectrum of **P9** (175 MHz, in CD_3OD)

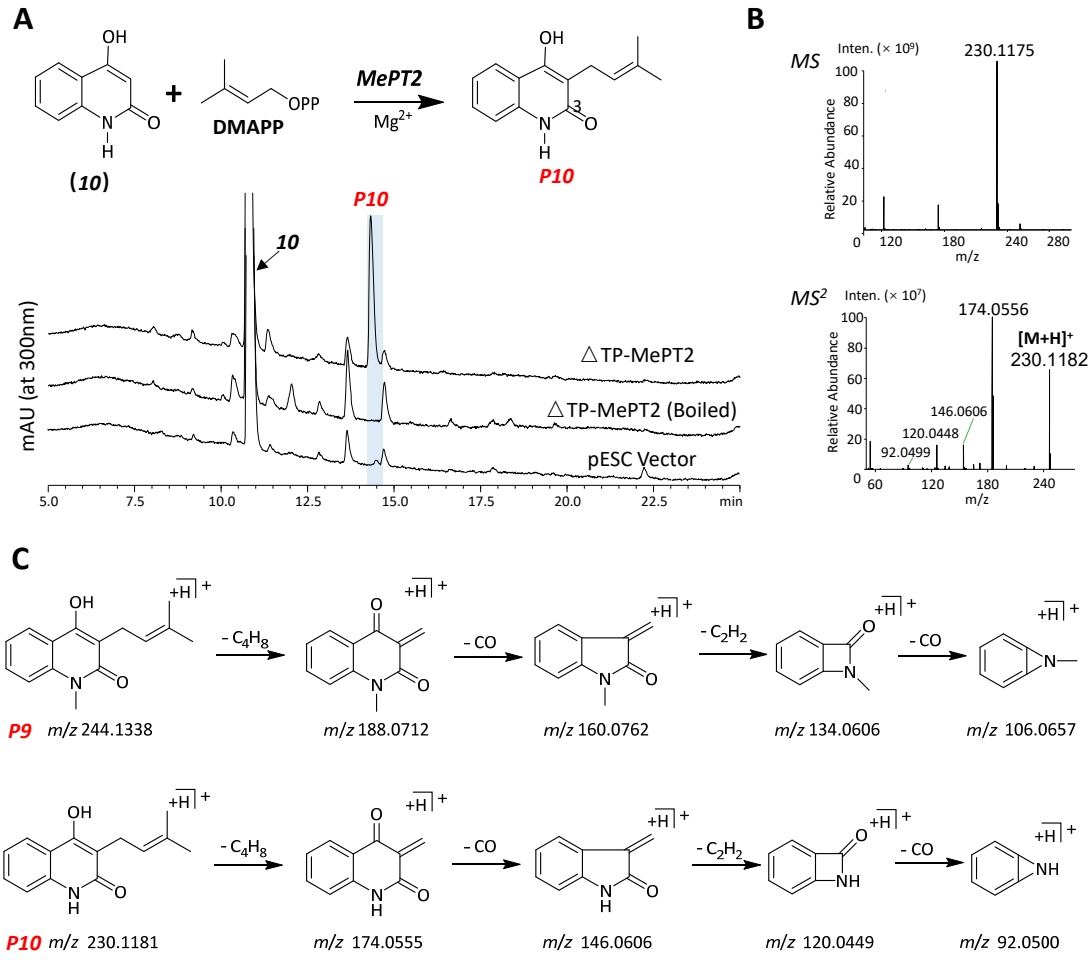


Figure S13 Enzymatic formation of 3-*C*-dimethylallylquinolone (**P10**) by MePT2 from quinolone (**10**) and DMAPP. (**A**) Structures and HPLC chromatograms. (**B**) The MS and MS^2 spectra of the enzymatic product **P10**. (**C**) The comparison of the MS^2 fragments of the enzymatic products **P9** and **P10.** The same MS cleavage behaviors of **P9** and **P10** observed in their MS^2 spectra suggested that **P10** shares a similar skeleton and the same dimethylallylated position (at C-3) with **P9**. ΔTP-MePT2: truncated protein of MePT2 with the *N*-terminal transit peptides removed.

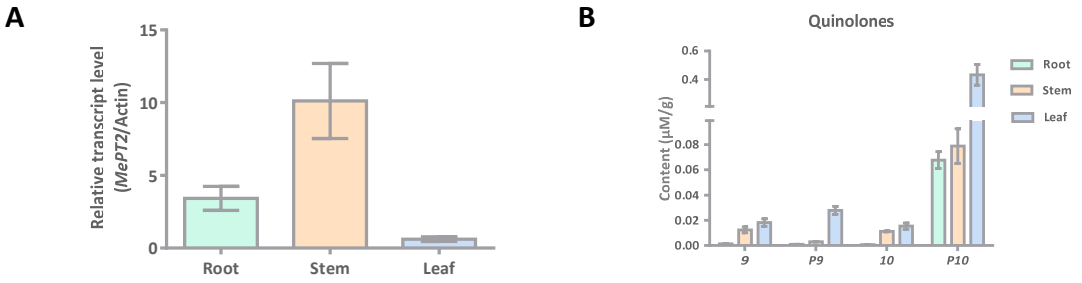


Figure S14 Real-time PCR analysis of *MePT2* in the different tissues of *M. exotica* (**A**), and the contents of substrates **9** and **10**, enzymatic products **P9** and **P10** in the different tissues of *M. exotica* (**B**). Relative expression levels of *MePT2* were normalized by the actin gene as a reference. Compound contents were quantified by LC/MS. Values are means ± SD of three biological replicates.