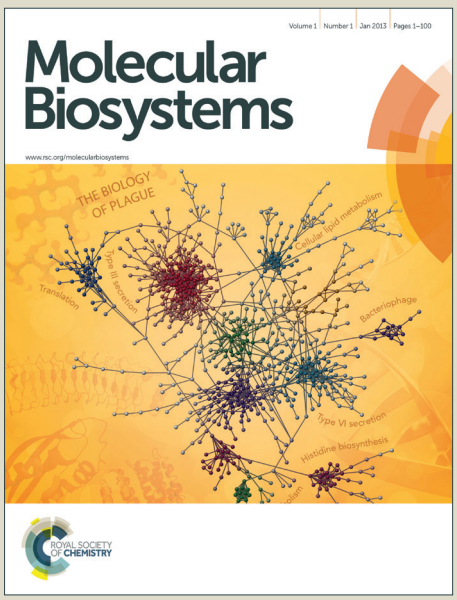
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Journal Name

RSCPublishing

ARTICLE

A multi-omics strategy resolves the elusive nature of alkaloids in *Podophyllum* species†

Cite this: DOI: 10.1039/x0xx00000x

Joaquim V. Marques, Doralyn S. Dalisay, Hong Yang, Choonseok Lee, Laurence B. Davin and Norman G. Lewis*

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Podophyllum hexandrum and, to a much lesser extent *P. peltatum*, are sources of podophyllotoxin, extensively used as a chemical scaffold for various anti-cancer drugs. In this study, integrated omics technologies (including advanced mass spectrometry/metabolomics, transcriptome sequencing/gene assemblies, and bioinformatics) gave unequivocal evidence that both plant species possess a hitherto unknown aporphine alkaloid metabolic pathway. Specifically, RNA-seq transcriptome sequencing and bioinformatics guided gene assemblies/ analyses *in silico* suggested presence of transcripts homologous to genes encoding all known steps in aporphine alkaloid biosynthesis. A comprehensive metabolomics analysis, including UPLC-TOF-MS and MALDI-MS imaging in situ, then enabled detection, identification, localization and quantification of the aporphine alkaloids, magnoflorine, corytuberine and muricinine, in the underground and aerial tissues. Interestingly, the purported presence of alkaloids in *Podophyllum* species has been enigmatic since the 19th century, remaining unresolved until now. The evolutionary and phylogenetic ramifications of this discovery are discussed.

Introduction

Podophyllum hexandrum and *P. peltatum* from the Berberidaceae family are best known as sources of the medicinal aryltetralin lignan, (–)-podophyllotoxin (PTOX **1**, Fig. 1).[1] It is, however, generally commercially extracted from *P. hexandrum* rhizomes, and is extensively used worldwide for semi-synthesis of the topoisomerase II inhibitor anticancer drugs, teniposide (**2**), *etopophos*® (**3**), and etoposide (**4**).[2]

Recently two cytochrome P450s (CyP450s) were reported from *P. hexandrum* (CYP719A23) and *P. peltatum* (CYP719A24), capable of facile methylenedioxy bridge formation, in the conversion of (–)-matairesinol (**18**) into (–)-pluviatolide (**19**) (Fig. 2b), i.e. with the latter being a proposed pathway intermediate to PTOX (**1**).[3] However, phylogenetic comparison of their corresponding proteins indicated they were more closely aligned to the CYP719A family known to catalyse an analogous reaction in isoquinoline alkaloid-producing plant species,[4] i.e. CYP719A23 and CYP719A24 share ~59% identity and ~78% similarity to *Coptis japonica* CYP719A1. Both of these CyP450s also have lower sequence identity/similarity (~26/~46%) to CYP81Q1, which is reportedly involved in sesame lignan methylenedioxy bridge formation.[5] Furthermore, in *P. hexandrum*, there are another two CyP450s, CYP81B56 and CYP81B57, with higher

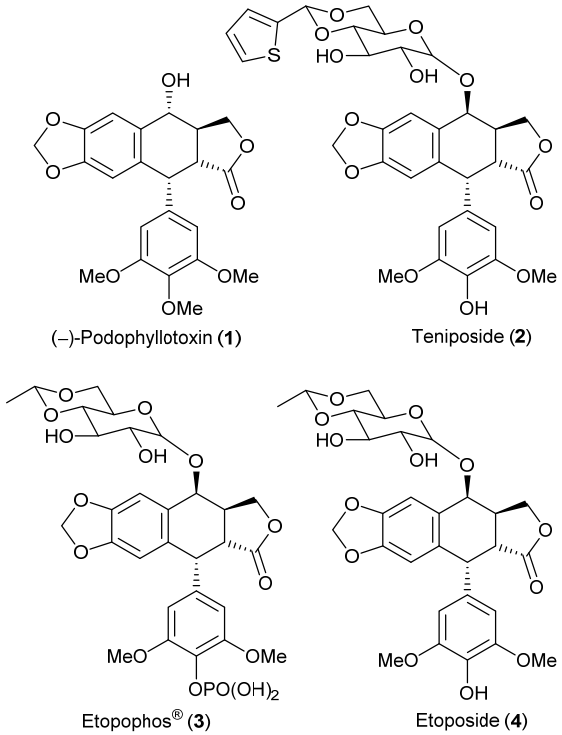


Fig. 1 Podophyllotoxin (**1**) scaffold and its anticancer drugs/derivatives

identity/similarity (46/63% and 49/66%, respectively) to the CYP81Q1 involved in the piperitol/sesamin (**21**/**22**) biosynthetic pathway that we had earlier discovered.[6] However, the corresponding recombinant proteins were catalytically inactive in our assays with matairesinol (**18**) and other putative intermediates tested.

Using a combination of transcriptomics, metabolomics and bioinformatics approaches, we discovered a hitherto undetected metabolic pathway in both *Podophyllum* species, which afford various aporphine alkaloids (Fig. 2a) – albeit those lacking a methylenedioxy bridge functionality. In addition, we report the successful use of MALDI metabolite imaging mass spectrometry, together with laser capture microdissection (for total RNA extraction), to determine spatial and temporal distribution of aporphine alkaloids and transcripts responsible for their biosynthesis. Our findings suggest evolutionary linkages between both lignan and alkaloid biosynthetic pathways.

Experimental details

Plant material

Podophyllum hexandrum and *P. peltatum* plants were obtained from Digging Dog Nursery (Albion, CA) and Companion Plants (Athens, OH), respectively, and maintained in Washington State University greenhouse facilities.

Chemicals

Dopamine (**5**) was purchased from Sigma-Aldrich. Magnoflorine (**13**), corytuberine (**12**) and norcoclaurine (**7**) authentic standards were kindly provided by Dr. Toni Kutchan (Donald Danforth Plant Science Center).

4-Hydroxyphenylacetaldehyde (**6**) was produced from 4-hydroxyphenylethanol in a one-step reaction in 95% crude yield as follows: 4-hydroxyphenylethanol (15 mmol) in dichloromethane (10 ml) was added to dichloromethane (200 ml) containing pyridinium chlorochromate (20 mmol), over ice. Following a 2 h reaction, during which the temperature reached room temperature, the reaction mixture was filtered, with the

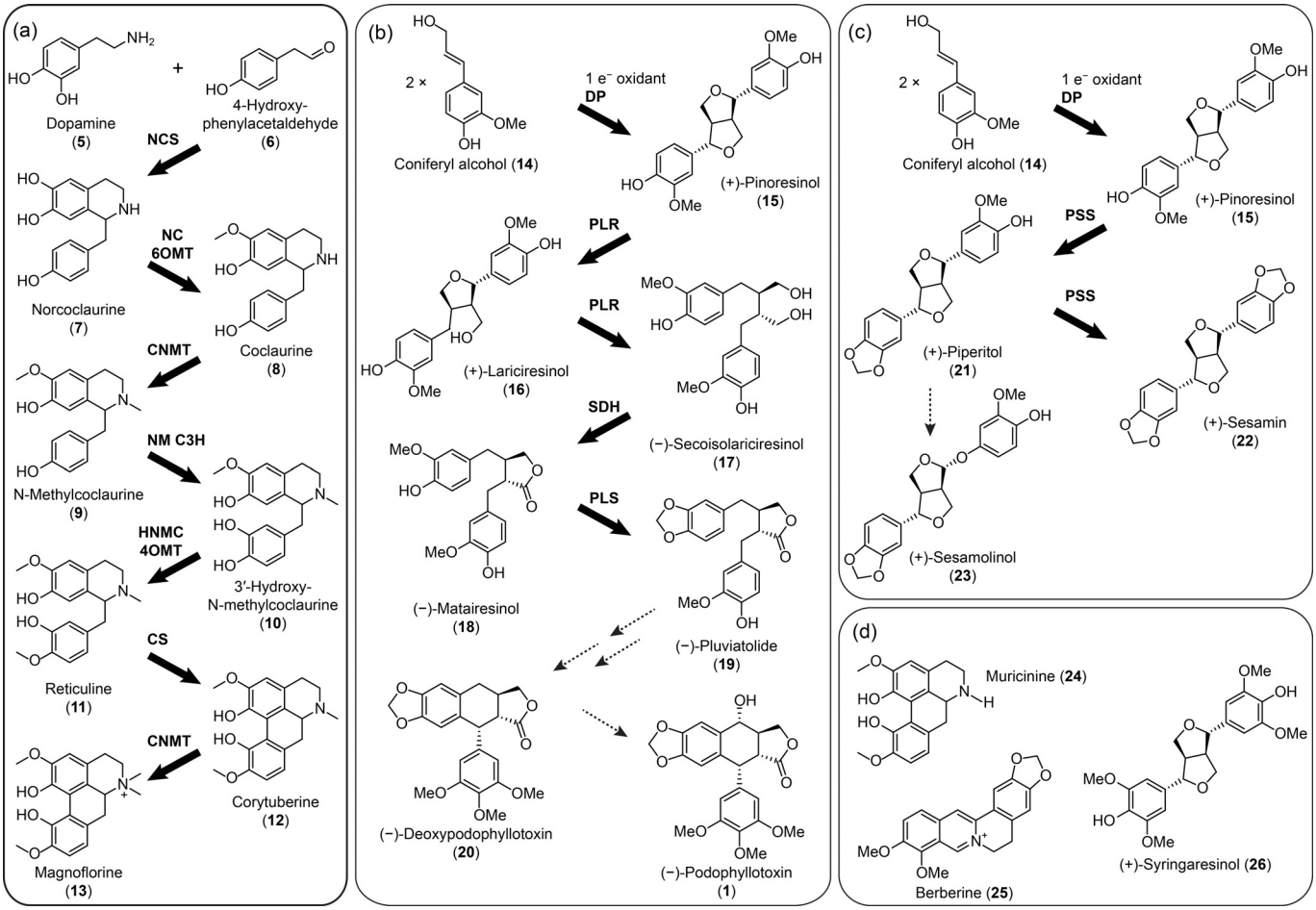


Fig. 2 Alkaloid and lignan pathways. (a) Magnoflorine (**13**), (b) podophyllotoxin (**1**), and (c) sesamin (**23**) biosynthetic pathways. (d) Structures of muricinine (**24**), berberine (**25**) and syringaresinol (**26**). Bold arrows represent known biosynthetic steps and dashed arrows represent currently unknown steps. CNMT, coclaurine *N*-methyltransferase; CS, corytuberine synthase; DP, dirigent protein, HNMC 4OMT, 3′-hydroxy-*N*-methyl-(*S*)-coclaurine 4′-*O*-methyltransferase; NC, 6OMT, norcoclaurine 6-*O*-methyltransferase; NCS, norcoclaurine synthase; NM C3H, *N*-methylcoclaurine 3′-hydroxylase; PLR, pinoresinol-lariciresinol reductase; PLS, pluviatolide synthase; PSS, piperitol/sesamin synthase; SDH, secoisolariciresinol dehydrogenase.

supernatant evaporated *in vacuo*, to afford the crude product, which was used without further purification.

Transcriptome sequencing and library assembly

Transcriptome sequencing, assemblies, and analyses were performed as described previously.[3]

Metabolite extraction and analyses

Metabolite extraction was carried out as before.[3] Briefly, different tissues (i.e. rhizomes, stems and leaves) were individually harvested, immediately frozen in liq. N_2, ground and subsequently lyophilized. After passage through a 150 μm sieve, each sample was individually extracted with 10 μl/mg methanol–water (7:3, v/v) for 5 minutes in an ice bath and sonication; the corresponding extracts were then maintained at –80 °C until analysis. For alkaloid metabolite detection and isolation, samples were subjected to an Acquity ultra performance liquid chromatography (UPLC, Waters) system, utilizing a BEH C18 column (Waters, 1.7 μm particles, 2.1 × 50 mm) with a binary mobile phase of formic acid–water (0.1%) (A) and formic acid-acetonitrile (0.1%) (B); detection was carried out at both 280 nm and by use of electrospray ionization mass spectrometry in the positive ion mode. A linear gradient (A:B) was used at a flow rate of 0.3 ml/min: from 95:5 to 75:25 in 11 min, to 60:40 in 5 min, to 0:100 in 4 min, followed by 1.5 min at 0:100. The column temperature was held at 25 °C and the sample injection volume was 5 μl. Masses were determined using a Waters Xevo G2 Q-TOF mass spectrometer using leucine–enkephalin, $C_{28}H_{37}N_5O_7$ with *m/z* 556.2771 $[M+H]^+$ as a lock-mass standard.

MALDI-imaging mass spectrometry (IMS)

Fresh plant sample tissue of a *P. hexandrum* rhizome and *P. peltatum* rhizome and root tissues were obtained by cryo-sectioning. Briefly, fresh tissues were embedded in ultra-pure agarose (3%) and cryo-sectioned at 15 μm thickness. The cryo-sections were placed on a MALDI target plate and 2,5-dihydroxybenzoic acid was applied at 40 mg/ml (in methanol–water, 1:1,v/v) by robotic sprayer (TM-sprayer™, HTX Technologies, Carrboro, NC). The acquisition mass range was then set from 100 to 1000 *m/z* with data obtained in the positive ion mode at 1,000 Hz, the laser energy set at 350, and using leucine-enkephalin as lock mass. The data was obtained at either 30 or 22 μm spatial resolution. The raw data was then processed and ion maps were visualized in HDImaging™ (v1.1) software. Three hundred of the most intense peaks were obtained after processing, with magnoflorine (**13**) readily identifiable. Collision induced dissociation MS/MS analysis using 30V collision energy gave fragmentation patterns identical to authentic magnoflorine (**13**). Ion mobility spectrometry was also used to obtain the drift time of the ions identified in MALDI-TOF IMS.

Toluidine blue O (TBO) stain for imaged sections

After MALDI imaging analysis, the MALDI plate was rinsed with ethanol to remove the matrix. The tissue section was then flooded with an aqueous 0.1% TBO solution for 3 minutes, and next successively washed with water and ethanol. The TBO stained tissue section was scanned using a standard flatbed scanner (Epson Perfection V500 photo, Japan).

Gene cloning and sequencing

Putative alkaloid biosynthetic genes, with the exception of 3′-hydroxy-*N*-methyl-(*S*)-coclaurine 4′-*O*-methyltransferase, were amplified from *P. hexandrum* cDNA using the primers described in Table 1. Amplification was performed in a thermocycler using PfuTurboDNA polymerase (Agilent Technologies) with 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 70 °C for 3 min, and a final extension at 70 °C for 10 min. PCR products were resolved in 1% agarose gels. Bands were extracted using QIAquick Gel Extraction Kit (QIAGEN) and cloned into pENTR™/D-TOPO®, following the manufacturer's instructions. Sequencing was performed by Eurofins. All sequences were deposited in the GenBank™ Data Bank under accession numbers KJ786954–KJ794165 (Table 2 and Fig. 3 for NCS).

Table 1 Primers used for sequence confirmation

Gene	Forward/reverse primers
Norcoclaurine synthase	ATGAGGATGGAAATTGTTTCCC
	TTATATTTCAATCAAGCTCTTATTCTTGT
Norcoclaurine 6-*O*-	ATGGAAGCTCAAAAGGAGAACATATC
methyltransferase	ATAAGGGAAAGCCTCAATTACAGA
Coclaurine *N*-	ATGGCCGTTACTGAAGATTCC
methyltransferase	TCATTTCCTGGTGAGTAGAAGATGA
N-Methylcoclaurine 3′-	ATGGATTCCCTGCACTGCC
hydroxylase	CTAGACAAAACATTTTGGAGAGATGC
Corytuberine synthase	ATGGAGCTAACAACAGCACTGC
	CTACACCCTGGATTTCGGAATG

***P. hexandrum* norcoclaurine synthase (NCS) heterologous expression and activity determination**

After confirmation of its DNA sequence in the pENTR/D-TOPO cloning vector, a Δ9 truncated PhNCS with addition of the MTGS[7] signal peptide was generated by PCR using forward (5′-CACCATGACCGGTTCGCTATTTTTCTTAGCA TGCGGAA-3′) and reverse (5′-TATTTCAATCAAGCTCTT ATTCTTGTT-3′) primers. This was subsequently cloned into the pET101/D expression vector (Invitrogen) by the TOPO reaction, according to the manufacturer's instructions. Δ9 PhNCS was next heterologously expressed in BL21 (DE3) (Novagen) cells after induction with 1 mM isopropyl β-D-1-thiogalactopyranoside at 30 °C.

After 19 h induction, cells were harvested and recovered after centrifugation (3900 × *g*) with pellets re-suspended in BugBuster Protein Extraction Reagent (Novagen) according to the manufacturer's instructions. Cell debris was removed by

Table 2 GenBank™ accession numbers

Gene	*P. hexandrum*		*P. peltatum*	
	Name	GenBank™ accession number	Name	GenBank™ accession number
Norcoclaurine synthase	PhNCS	KJ794165	PpNCS	KJ786964
Norcoclaurine 6-*O*-methyltransferase	PhNC6OMT	KJ786954	PpNC6OMT	KJ786955
Coclaurine *N*-methyltransferase	PhCNMT	KJ786956	PpCNMT	KJ786957
N-Methylcoclaurine 3′-hydroxylase	PhNMC3H	KJ786958	PpNMC3H	KJ786959
3′-Hydroxy-*N*-methyl-(*S*)-coclaurine 4′-*O*-methyltransferase	PhHNMC4OMT	KJ786960	PpHNMC4OMT	KJ786961
Corytuberine synthase	PhCS	KJ786962	PpCS	KJ786963

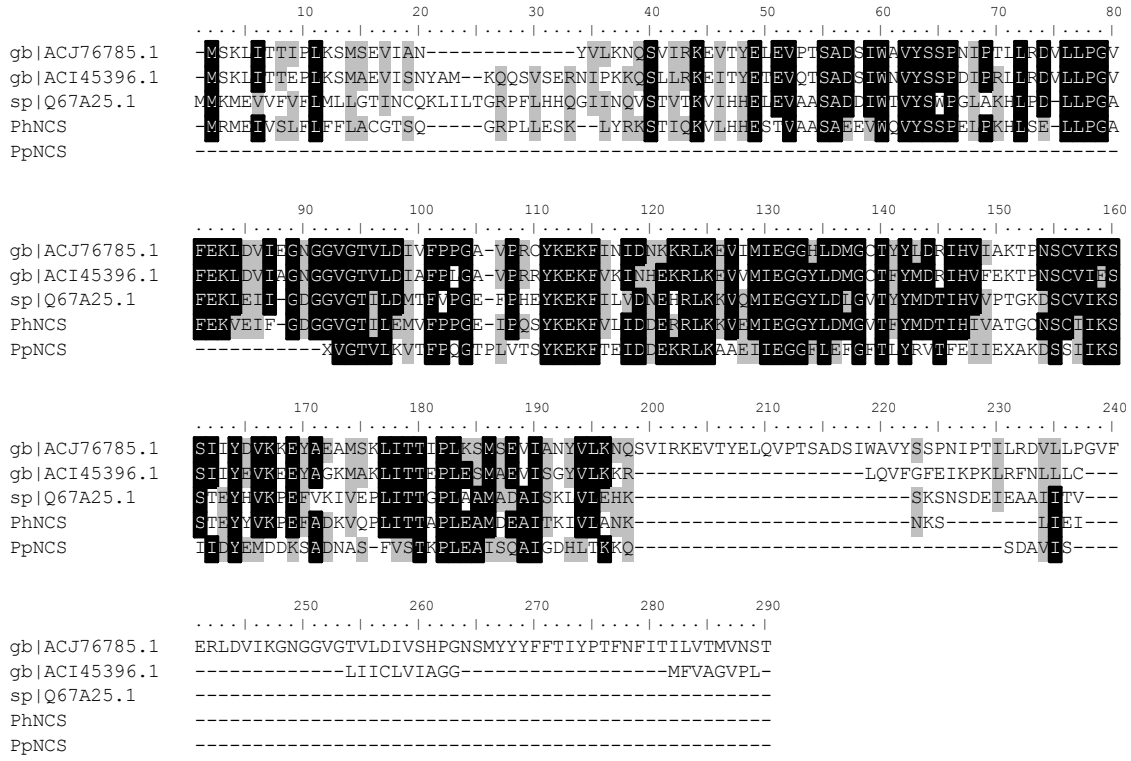


Fig. 3 Amino acid alignment of norcoclaurine synthase (NCS) genes. *Argemona mexicana* (gb|ACJ76785.1|), *Papaver somniferum* (gb|ACI45396.1|), *Thalictrum flavum* (sp|Q67A25.1|) and the homologs found in *P. hexandrum* (PhNCS, accession KJ794165) and *P. peltatum* (PpNCS, accession KJ786964), respectively

centrifugation (10,000 × *g* for 20 min). Supernatants were individually subjected to POROS MC 20 column (1 × 9 cm) chromatography with Ni^{2+} ($NiSO_4$) as metal ion.[8] After washing with buffer A (5 mM Tris-HCl, pH 7.9, 500 mM NaCl, 50 ml) containing 5 mM imidazole, recombinant NCS was eluted with a stepwise imidazole gradient (5–500 mM, 100 ml) in buffer A. Fractions containing recombinant NCS (i.e. eluting between 100 and 400 mM imidazole) were buffer-exchanged to 100 mM Tris-HCl buffer, pH 7.0, and concentrated (~0.5 ml) using a Centricon Plus-20 (Amicon, Billerica, MA). Fractions eluted with 300 mM, and containing NCS purified to apparent homogeneity (Fig. S1), were utilized in the assays. All enzyme manipulations and purifications were carried out at 4 °C.

Enzymatic assays were performed at 37 °C in 220 µl final volume of 100 mM Tris-HCl buffer, pH 7.0, containing 5 µg of purified recombinant NCS, dopamine (**5**) and 4-hydroxyphenylacetaldehyde (**6**) at a final concentration of 5 mM. Enzymatic reactions were carried out for 30 min and stopped with addition of glacial acetic acid (10 µl).

Laser microscope dissection of xylem, pith and cortex regions from rhizomes

Laser capture microdissection was used to dissect xylem, pith and cortex cells from rhizomes of both *P. hexandrum* and *P. peltatum* using a PALM MicroBeam System (Zeiss Microscopy) as described previously.[9] Rhizomes were embedded with Tissue-Tek OCT (Miles, Inc., Elkhart, IN) compound. The frozen cryocut sections (10 µm) were obtained from a Leica CM 1950 Cryostat, and then mounted on activated UV PEN slides (Applied Biosystems). Slides were placed in cold water for 1 min to remove OCT, dried *in vacuo* and stored at –20 °C. Each slide was subjected to ethanol dehydration steps using 70%, 95% ethanol in water and 100% ethanol for

Table 3 Alkaloid biosynthetic pathway homolog

P. hexandrum/ *P. peltatum* homologs	Pathway genes		Similarity/identity (%)	
	Species	Data Bank accession number	*P. hexandrum*	*P. peltatum*
Norcoclaurine synthase (NCS)	*Thalictrum flavum*	Q67A25.1	74/59	56/38
Norcoclaurine 6-O-methyltransferase (NC 6OMT)	*T. flavum*	AAU20765.1	73/41	85/59
Coclaurine N-methyltransferase (CNMT)	*T. flavum*	Q5C9L6.1	69/48	82/65
	Coptis japonica	BAB71802.1	70/48	83/66
N-Methylcoclaurine 3′-hydroxylase (NM C3H)	*T. flavum*	AAU20767.1	73/41	85/59
3′-Hydroxy-N-methyl-(S)-coclaurine 4′-O-methyltransferase (HNMC 4OMT)	*C. japonica*	Q9LEL5.1	71/40	82/63
Corytuberine synthase (CS)	*C. japonica*	BAF80448.1	74/37	85/56

1 min and air dried before laser micro-dissection. For each region (i.e. xylem, pith and cortex) approximately 8000 dissected cells were collected, combined and kept in –80 °C until further use.

RNA extraction, cDNA synthesis and gene expression analysis

Rhizomes, roots and shoots were freshly harvested from *P. hexandrum* and *P. peltatum* and immediately frozen in liquid nitrogen until use. Total RNA was isolated from each organ individually using Plant RNA Purification Reagent (Invitrogen) with contaminating genomic DNA digested with RNase-free DNase (Invitrogen). Total RNA was also extracted from dissected cells with Qiagen RNeasy Micro Kit (QIAGEN) according to the manufacture's instruction.

First strand cDNA was next synthesized using SuperScript III First-Strand Synthesis System (Invitrogen) with DNase treated RNA (1 μg for rhizome, root and shoot or 500 pg for dissected cells) as template according to the manufacturer's instructions. PCR mixtures (20 μl) for gene expression analysis consisted of cDNA (1 μl), Platinum® SYBR® Green qPCR SuperMix-UDG (Invitrogen, 10 μl), Rox dye (Invitrogen, 0.4 μl), nuclease-free water (7.8 μl) and gene specific primers (0.4 μl of each). Primers were designed using Primer3 software with melting temperature of 57–58 °C: for *P. hexandrum,* PhNCS-For (5′-GTGCTTCACCATGAGTCCAC-3′) and PhNCS-Rev (5′-TTCAACCTTCTCGAATGCAC-3′), for *P. peltatum* PpNCS-For (5′-TTTCCACAAGGAACACCTC TC-3′) and PpNCS-Rev (5′-CAAACTCTAGGAATCCTC CTTCA-3′), as well as Actin-For (5′-TCAGGCTGTTCTTTC CCTTT-3′) and (Actin-Rev 5′-GGAAGGGCATAACCTT CGTA-3′). Triplicate reactions were performed using cDNAs synthesized from independent total RNA isolations as follows: 95 °C for 15 min followed by 40 cycles of 94°C for 30 s, 58°C for 1 min, and 72°C for 30 s. Dissociation curves were obtained with products confirmed as single products by running agarose gels. Relative expression levels were normalized against *P. hexandrum* and *P. peltatum* actin genes.

Results

Our earlier *Podophyllum* sp. transcriptome profiling of aerial and underground tissues, followed by gene assembly and bioinformatics analysis,[3] established the validity of our assembly processes for previously identified genes encoding known biochemical steps to PTOX (**1**). This not only confirmed that the assemblies for known biochemical steps were identical to earlier known gene sequences,[10-12] but it also resulted in our identifying several putative gene candidates for remaining (unknown) enzymatic steps.

However, the finding that *Podophyllum* species encoded genes more closely related on sequence similarity/identity to methylenedioxy bridge forming CyP450s in alkaloid formation, rather than lignan biosynthesis, raised the question of whether our gene assemblies contained additional potential alkaloid pathway genes.

Accordingly, we re-examined the assemblies and transcripts from *P. hexandrum* and *P. peltatum* rhizome, root, stem and leaf tissues (medplants.ncgr.org). This resulted in several contigs being detected homologous to genes possibly encoding an aporphine alkaloid biosynthetic pathway (Fig. 2). That is, there were putative *P. hexandrum* and *P. peltatum* norcoclaurine synthases (NCS) with *circa* 74/56% similarity and 59/38% identity to *Thalictrum flavum* NCS (UniProtKB accession Q67A25.1) affording entry into the magnoflorine (**13**) pathway (Figs. 2 and 3). In an analogous manner, homologs for all of the remaining steps were also identified (see Table 3). These genes, with the exception of PhHNMC4OMT, were cloned from *P. hexandrum* and their sequences again demonstrated the validity of our assemblies
These analyses thus suggested the presence of a previously undetected aporphine alkaloid biosynthetic pathway operative in both *Podophyllum* species.

Norcoclaurine synthase activity

In order to further confirm the authenticity of the alkaloid biosynthetic pathway provisionally detected, the putative *P. hexandrum* NCS (Fig. 3) was heterologously expressed in *E. coli* as a Δ9 truncated [i.e. lacking a putative signal peptide[13]], C-terminus HIS-tag, recombinant protein. After purification by Ni affinity chromatography (Fig. S1), it was found capable of converting dopamine (**5**) and 4-hydroxy-phenylacetaldehyde (**6**) into norcoclaurine (**7**), with the product identified by comparison with an authentic standard (Fig . 4). That is, norcoclaurine (**7**) had an elution time of *ca.* 3.5 minutes (Fig. 4a) and a mass spectrum with *m/z* 272.1272, corresponding to $[M+H]^+$ (calc. 272.1287, $C_{16}H_{18}NO_3$) and characteristic

fragments of *m/z* 255.1007 (calc. 255.1016, $[M+H-NH_3]^+$), 237.0909 (calc. 237.0915, $[M+H-NH_3-H_2O]^+$), 161.0591 (calc. 161.0602, $[C_{10}H_9O_2]^+$) and 107.0490 (calc. 107.0497, $[C_7H_7O]^+$) (Fig. 4d). The enzymatically produced norcoclaurine (**7**) had the same elution time (Fig. 4b) and a near identical mass spectrum (Fig. 4e), whereas with a boiled enzyme assay (Fig. 4c), norclaurine (**7**) was not detectable.

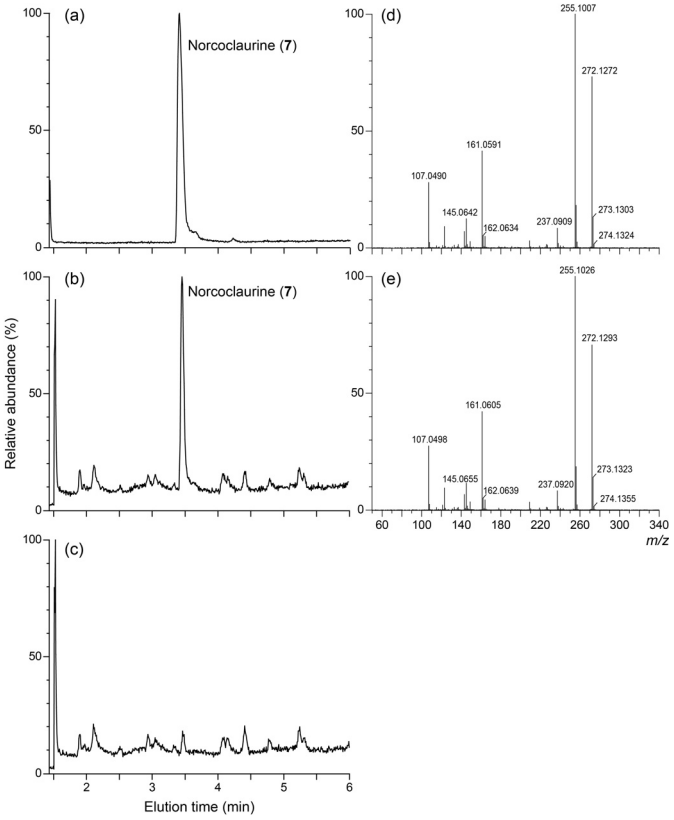


Fig. 4 Norcoclaurine (7) production by recombinant norcoclaurine synthase. (a) Chromatogram of norcoclaurine standard (7). (b and c) Enzymatic reaction products from heterologous (recombinant) PhNCS (b) and control assay with boiled enzyme (c). Data shows that norcoclaurine (7) (elution time: 3.45 min) formation occurs only in presence of PhNCS. Mass spectra of norcoclaurine (7) standard (d) and enzymatically produced 7 (e).

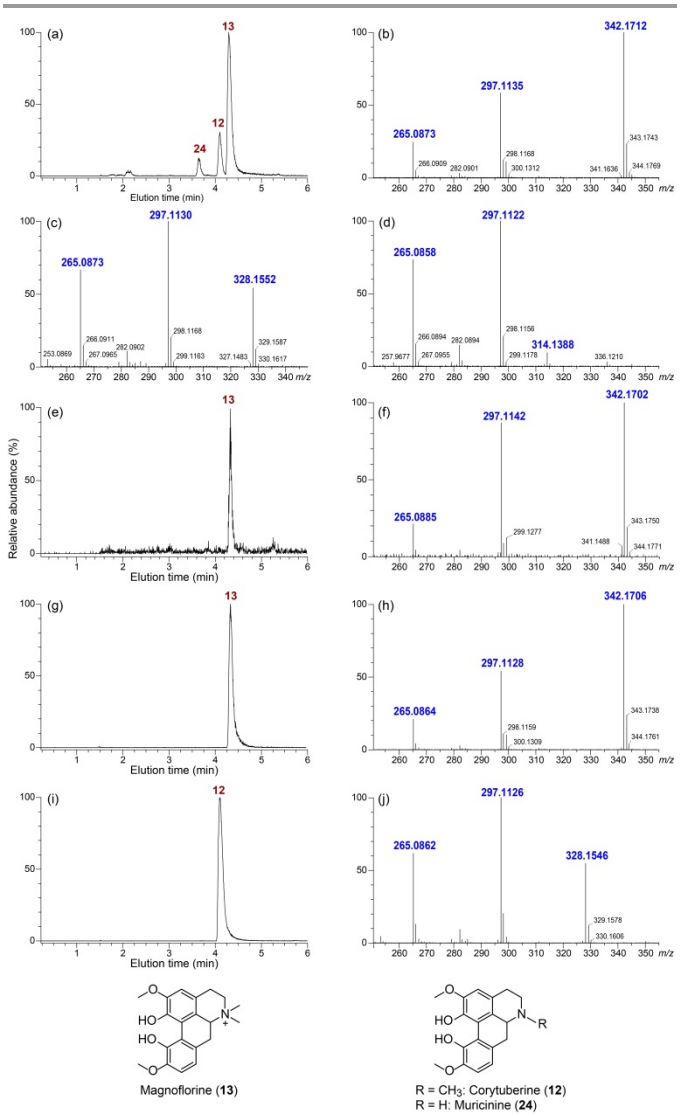


Fig. 5 Chromatograms and mass spectra of alkaloids from *Podophyllum* species and standards. (a) Extracted ion chromatogram at *m/z* 297.1±0.5 Da of *P. hexandrum* rhizome extract corresponding to alkaloid magnoflorine (13), corytuberine (12) and muricinine (24). (b) to (d) Mass spectral fragmentation patterns of magnoflorine (13, b), corytuberine (12, c), and putative muricinine (24, d) in *P. hexandrum* rhizome extract. (e) Extracted ion chromatogram at *m/z* 297.1±0.5 Da of *P. peltatum* rhizome corresponding to magnoflorine (13). (f) Mass spectral fragmentation pattern of magnoflorine (13) in *P. peltatum* rhizome extract. (g) and (i) Extracted ion chromatogram at m/z 297.1±0.5 Da of authentic standards magnoflorine (13, g) and corytuberine (12, i). (h) and (j) Mass spectral fragmentation patterns of authentic standards magnoflorine (13, h) and corytuberine (12, j).

Metabolomics-guided analysis of *Podophyllum* species

Metabolomic profiles from different tissues were next more in-depth analysed, with an emphasis – based on the provisionally annotated gene sequences above – on benzylisoquinoline alkaloids, using UPLC-TOFMS high sensitivity and accuracy. These analyses resulted in detection of 3 peaks whose mass spectral fragmentation patterns were consistent for aporphine alkaloids[14, 15] (Fig. 5). The most abundant peak in both *P. hexandrum* (Fig. 5a) and *P. peltatum* (Fig. 5e) was identified as magnoflorine (**13**) when compared with an authentic standard (Fig. 5g and h). It had mass spectra essentially identical to the standard, with a base peak of *m/z* 342.1712 (Fig. 5b), corresponding to $[M]^+$ (calc. 342.1700, $C_{20}H_{24}NO_4$), accompanied by fragments of *m/z* 297.1135 (calc. 297.1127, $C_{18}H_{17}O_4$), resulting from the loss of $NH(CH_3)_2$, and of *m/z* 265.0873 (calc. 265.0865, $C_{17}H_{13}O_3$), generated by further loss of MeOH, with these fragments being characteristic of aporphine alkaloids.[15] The magnoflorine (**13**) from *P. peltatum* also gave identical data (Fig. 5f). Interestingly, the main tissue for accumulation of magnoflorine (**13**) in *P. hexandrum* was in the rhizomes (Fig. 6a), while in *P. peltatum* it was more abundant in the leaves (Fig. 6b), respectively.

The second most abundant alkaloid detected was identified as corytuberine (**12**) by comparison with the authentic standard (Fig. 5i and j). It had a mass spectrum (Fig. 5c) identical to that of the authentic standard (Fig. 5j), consisting of an ion of *m/z*

328.1552, corresponding to $[M+H]^+$ (calc. 328.1549, $C_{19}H_{22}NO_4$), accompanied by the same fragments of *m/z* 297 and 265. The third alkaloid detected (Fig. 5a) was provisionally identified as muricinine (**24**), with a peak of *m/z* 314.1388, corresponding to $[M+H]^+$ (calc. 314.1392, $C_{18}H_{20}NO_4$) and the same fragments of *m/z* 297 and 265 (Fig. 5d). The differences in the mass spectra of the three detected alkaloids (Fig. 5b, d and f) thus reflect successive methylation of the secondary amine functionality. The different molecular ions, after their characteristic losses of $NH(CH_3)_2$ in magnoflorine (**13**), NH_2CH_3 in corytuberine (**12**) and NH_3 in muricinine (**24**)], produce the characteristic fragments of *m/z* 297 and 265.

Thus, our hypothesis that *Podophyllum* species had a complete aporphine alkaloid biosynthetic pathway was fully validated.

MALDI-imaging mass spectrometry and qPCR of laser microdissected tissues

It was instructive to next compare and contrast the localization of magnoflorine (**13**) *in situ* with expression patterns of various genes involved in its biosynthetic pathway in rhizome/root tissues and in individual cell types (Fig. 7 and 8). From an anatomical perspective, underground tissues from both *P. hexandrum* and *P. peltatum*, respectively, were stained with TBO after MALDI IMS analysis for visualization. This includes: *P. hexandrum* rhizome and its roots [Fig. 7a, with arrow heads indicating older (green) and younger (yellow) parts of the roots], as well as a *P. peltatum* rhizome with a newly developed root (green arrow head, Fig. 8a); longitudinal sections of *P. hexandrum* root tissues with emerging lateral roots (indicated with yellow arrow heads, Fig. 7c), as well as a *P. peltatum* rhizome with newly developed shoots (Fig. 9*C*, right of green broken line).

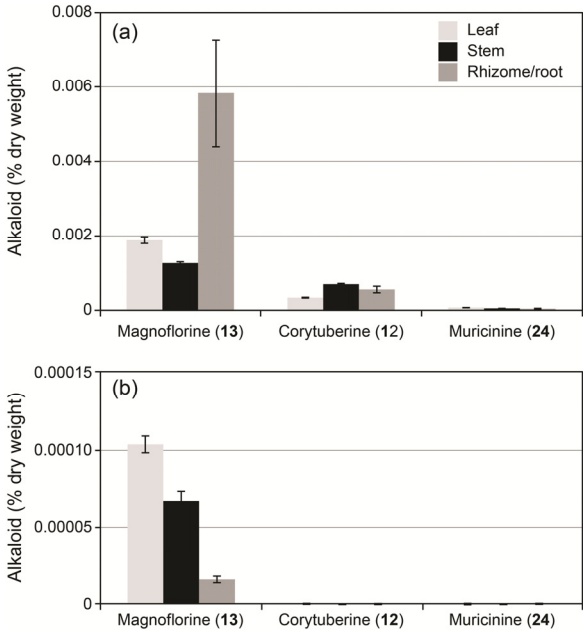


Fig. 6 Alkaloid content in *Podophyllum* species. Alkaloid accumulation (% dry weight) in extracts from different tissues of *P. hexandrum* (*A*) and *P. peltatum* (*B*).

Prior to magnoflorine (**13**) localization *in situ*, we first examined whether 2,5-dihydroxybenzoic acid was a good ionization matrix for detection of this alkaloid. With this matrix, the magnoflorine (**13**) molecular ion of *m/z* 342.1705 $[M]^+$ was readily detected in the positive ion mode (data not shown). Its spatial distribution was thus next investigated in the various *Podophyllum* tissues. High resolution MALDI-TOF IMS showed molecular ions at *m/z* 342.1707 (*P. hexandrum*) and *m/z* 342.1700 (*P. peltatum*) corresponding to $[M]^+$, with the MS/MS fragmentation patterns being in good agreement to an authentic **13** standard. The *P. hexandrum* tissues showed a higher magnoflorine (**13**) molecular ion intensity, which was ~8 times higher compared to *P. peltatum* tissues, a finding in good agreement with our UPLC-MS analysis (Fig. 6).

In the *P. hexandrum* cross-section, magnoflorine (**13**) molecular ion intensity varied, being highest in the rhizome pith and epidermal cells (Fig. 7b) as illustrated in the heat map (yellow being the most intense). One younger root cross-section (Fig. 7b, yellow arrowhead) also had a higher molecular ion abundance in the epidermal and cortical parenchyma cells. By contrast, there was no magnoflorine (**13**) molecular ion detected in the vascular bundles of either the rhizome or the roots. In the longitudinal tissue sample of a *P. hexandrum* root, magnoflorine (**13**) was also mainly highly detectable in the emerging roots (Fig. 7d, yellow arrows) with a lower ion intensity in parts of the epidermis. For the *P. peltatum* rhizome cross-section, only a quarter of the sampled cross-section was subjected to imaging analysis (Fig. 8b), which included a newly developed root (green arrowhead). Although less readily detectable in this tissue, magnoflorine (**13**) was spatially localized in pith and epidermal cells of the rhizome, and also where the root was emerging (Fig. 8b). In the longitudinal section of *P. peltatum* rhizome with newly developed shoots (Fig. 8b), magnoflorine (**13**) displayed a higher ion abundance in the region just before the budding shoot. It was also detected in the vasculature of the rhizome that it connects to (Fig. 8d, orange arrows).

Lastly, it was considered instructive to establish the location of expression of the genes in the magnoflorine (**13**) pathway. This was investigated, using qPCR from RNA extracted either from different organs and tissues or cell types excised via laser microdissection. First, in *P. hexandrum*, the emerging roots were separated from the older roots (Fig. 7c, green dotted line), whereas in *P. peltatum*, the rhizome was separated from the budding shoot (Fig. 8c, green dotted line), with each RNA sample then subjected to qPCR analysis. The expression levels of NCS (Fig. 7e and 8e) correlated very well with the pattern of alkaloid distribution as determined by MALDI-MS imaging (Fig. 7d and 8d). That is, in the case of the *P. hexandrum*,

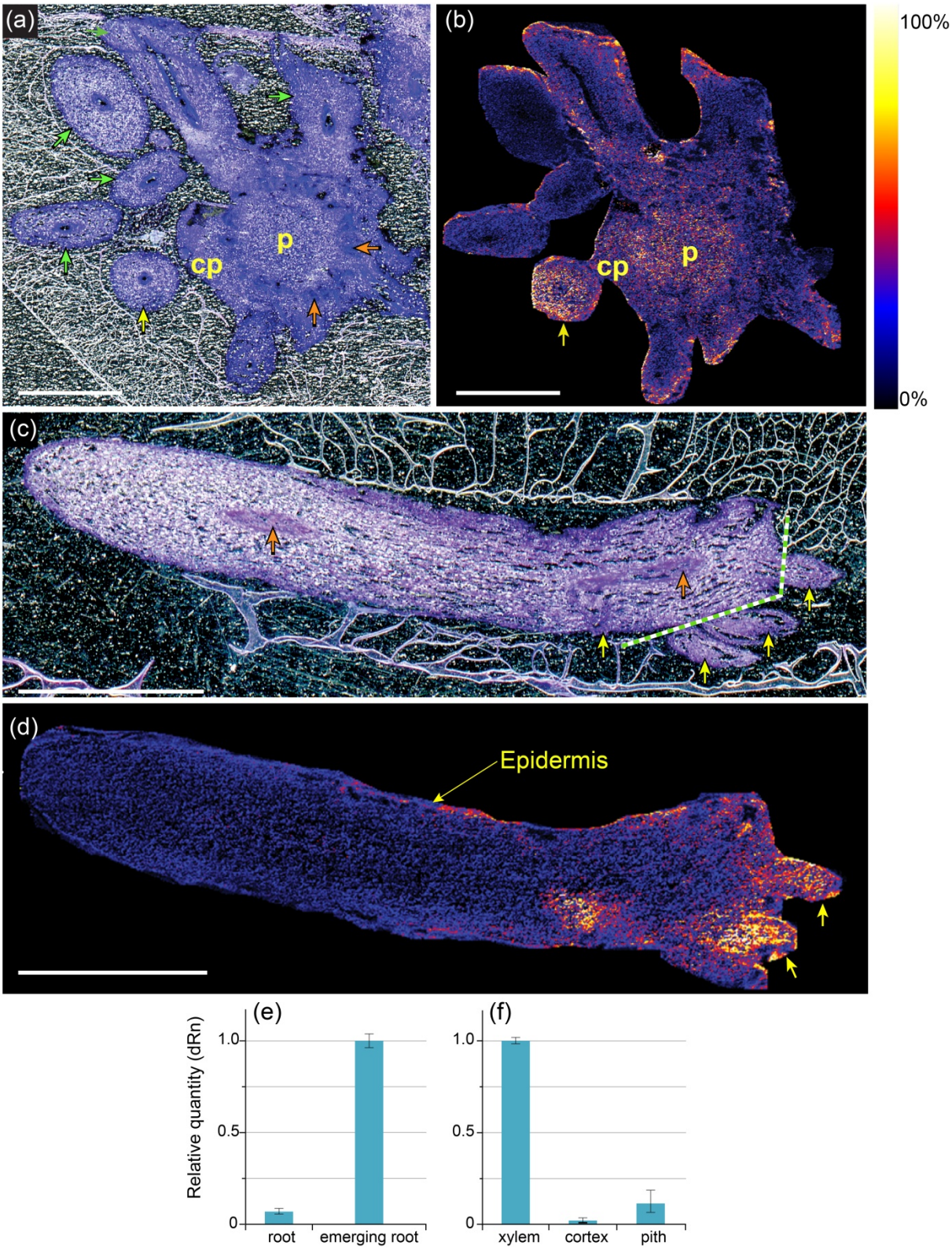


Fig. 7 Spatial localization of magnoflorine (13) and norcoclaurine synthase gene expression in *P. hexandrum*. MALDI imaging mass spectrometry images of magnoflorine (13) (b and d) and expression levels of norcoclaurine synthase (e and f). For illustration, the MALDI plate is shown stained with toluidine blue O, TBO, after MALDI-imaging (a and c). Green arrow heads indicate roots, whereas yellow arrow heads show emerging roots. Regions with high magnoflorine (13) content in emerging roots of *P. hexandrum* (d) have higher NCS expression (e). Orange arrow heads indicate vasculature. The green broken lines in *C* shows areas excised for qPCR analyses: Magnoflorine (13) accumulates in tissues (d) with higher NCS expression level (e). Abbreviations: p, pith; cp, cortical parenchyma. Relative intensity of metabolites depicted in heat map, where yellow is the highest intensity, and blue the lowest. Bars = 5 mm.

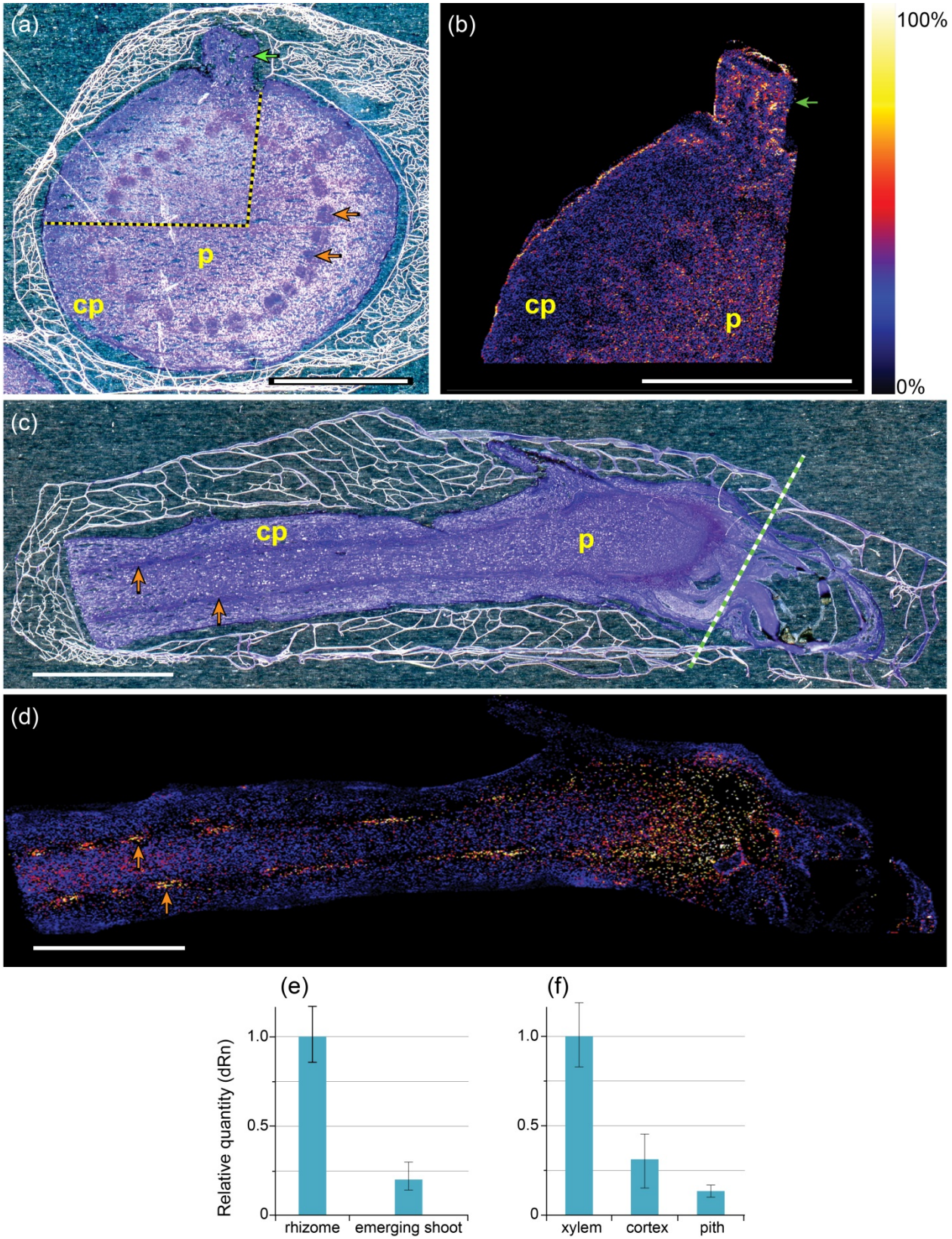


Fig. 8 Spatial localization of magnoflorine (**13**) and norcoclaurine synthase gene expression in *P. peltatum*. MALDI imaging mass spectrometry images of magnoflorine (**13**) (b and d) and expression levels of norcoclaurine synthase (e and f). For illustration, the MALDI plate is shown stained with toluidine blue O, TBO, after MALDI-imaging (a and c). Green arrow heads indicate roots, whereas yellow arrow heads show emerging roots. Orange arrow heads indicate vasculature. The green broken lines in (c) shows areas excised for qPCR analyses: Magnoflorine (**13**) accumulates in tissues (d) with higher NCS expression level (e). Abbreviations: p, pith; cp, cortical parenchyma. Relative intensity of metabolites depicted in heat map, where yellow is the highest intensity, and blue the lowest. Bars = 5 mm.

higher transcript and magnoflorine (**13**) levels were co-localized in the emerging roots, while alternatively in *P. peltatum*, highest gene expression and alkaloid (**13**) content were more evident in the rhizome, rather than the young, emerging shoots. All other genes in the pathway gave similar results, namely norcoclaurine 6-*O*-methyltransferase (NC 6OMT), coclaurine *N*-methyltransferase (CNMT) and corytuberine synthase (CS) (data not shown). Thus these data strongly correlated magnoflorine (**13**) accumulation with site of gene expression. It was next of interest to establish patterns of gene expression in individual cell types. In this regard, laser microdissection was carried out to isolate xylem, cortex and pith-enriched cell preparations from *P. hexandrum* (Fig. 7f) and *P. peltatum* (Fig. 8f), where expression of NCS was found to be mainly associated with the xylem cells, thus spatially distinct from site of magnoflorine (**13**) accumulation. Similar results were obtained with NC 6OMT, CS and CNMT genes (Fig. S2).

Discussion

The presence of alkaloids in *P. hexandrum* has been a subject of debate since the 19th century, with the first report indicating that it contained berberine (**25**, Fig. 2d). However, other researchers later concluded that no alkaloids were present [16-18] More recently, Wani and co-workers reported very preliminary colorimetric findings in terms of faintly detecting. alkaloids in *P. hexandrum* aqueous extracts using Mayer's reagent.[19] No alkaloids were, however, identified. Additionally, these same investigators failed to detect any positive colorimetric testing of the corresponding methanolic extracts using three reagents (Mayer's, Wagner's and Dragendorff) specific for alkaloids.[20] By contrast, we unambiguously established presence of a hitherto unknown aporphine alkaloid pathway being operative in both *Podophyllum* species, and the identity of the alkaloids so formed. This discovery relied initially on in depth Illumina transcriptome sequencing and bioinformatics "omics" analyses, but was then definitively established by probing UPLC-TOF-MS metabolomic data, with magnoflorine (**13**), corytuberine (**12**) and putatively muricinine (**24**) being identified.

As further proof, the deduced entry point pathway enzymatic step, catalysed by NCS, was also shown to be functionally active, by cloning the encoding Δ9 truncated gene from *P. hexandrum*, and demonstrating that recombinant NCS was capable of forming norcoclaurine (**7**) from dopamine (**5**) and 4-hydroxy-phenylacetaldehyde (**6**). The remaining downstream pathway steps were also all cloned from *P. hexandrum* (with the exception of PhHNMC4OMT), with each cloned gene having the full sequences deduced earlier from the RNA-seq assemblies.

Yet, in contrast, prior work by Facchini and co-workers[21] on the phylogeny of benzylisoquinoline alkaloid biosynthesis in various plant species, was unable to detect NCS activity in *P. peltatum*. This earlier inability to detect NCS activity may be due to the experimental conditions (and limited sensitivity) of the assay conditions employed in that investigation.

Potential evolutionary ramifications

The biosynthetic pathways to the Podophyllum aryltetralin lignans and aporphine alkaloids have many similarities, in terms of being derived from aromatic amino acids, and in undergoing somewhat analogous skeletal re-arrangements (i.e. inter- and/or intra- molecular coupling/cyclisations), hydroxylations and methylations. Indeed, the hydroxylations, methylations and the C-C intramolecular cyclisation to form PTOX (1) and related lignans are essentially mirrored in the pathway leading to the aporphine alkaloids (Fig. 2). In contrast to PTOX (1), however, alkaloids displaying methylenedioxy bridge functionalities were not detected in the present investigation.

Presence of alkaloids in *Podophyllum* species might be considered consistent with its evolutionary proximity to several benzylisoquinoline alkaloid producing species in the Berberidaceae family.[21] On the other hand, current molecular phylogeny of this family[22, 23] clearly places *Podophyllum* species in a monophyletic group (often designated *Podophyllum* group) which includes *Sinopodophyllum, Dysosma* and *Diphylleia*. All of these plant species are known to accumulate PTOX (**1**) and/or related lignans,[24, 25] but not alkaloids. The results herein, however, now clearly demonstrate how our rapidly evolving multi-'omics' approaches can radically change the way we consider natural product chemistry and chemotaxonomy, i.e. with increasingly more sensitive analytical technologies, deeper and more accurate coverage of transcriptomes, genomes, proteomes, and metabolomes, as well as more powerful bioinformatics.

The detection of alkaloid metabolism is also an indication that evolution did not completely eliminate alkaloid metabolism, but was perhaps minimized in favour of lignan biosynthesis that characterize these species. We can thus hypothesize though that one or more alkaloid biosynthetic genes have been recruited into the PTOX (**1**) and related lignan biosynthetic pathways, as exemplified by the CYP719A enzymes capable of converting matairesinol (**18**) into pluviatolide (**19**).[3]

Another interesting report was that when *Arabidopsis* was treated with the soil-borne vascular pathogen *Verticillium longisporum*, this apparently resulted in induced accumulation of lignans, piperitol (**21**), sesamolinol (**23**) and syringaresinol (**26**) in the root tissue.[26] Until then, methylenedioxy-bridge containing lignans **21** and **23** had never been reported in *Arabidopsis*. However, they are also well-known abundant food protective lignans in sesame oil, implying perhaps antioxidant or defence roles. Accordingly, while the genes in *Arabidopsis* leading to these lignans are currently unknown, formation of (**21**) and (**23**) raises the possibility that a "silent" metabolic pathway is induced on pathogen attack.

Aporphine alkaloid localization and metabolic pathway considerations

Metabolite localization at the cell, tissue and organ type level *in situ,* together with corresponding targeted gene expression

levels, are powerful tools in natural product biosynthetic studies. In this work, we thus mapped the spatial distribution of magnoflorine (**13**) in *Podophyllum* sp. mature/older rhizome and root tissues using MALDI imaging mass spectrometry and complemented these results with tissue and cell type targeted gene expression analysis. Our results demonstrated that in mature/older rhizome/root tissues, the localization of magnoflorine (**13**) and gene transcripts in its metabolic pathway were distinct in terms of cell types. That is, the pathway genes from NCS to CS are mainly detected in xylem cells, whereas magnoflorine (**13**) is mainly found in pith, epidermal cells, and newly emerging root sections (Fig. 7 and 8). This finding apparently strongly suggests transport processes are in place from one cell type for its formation to the other type(s) for storage. In the case of *Podophyllum* species, biosynthesis apparently occurs in younger tissues and vasculature, with accumulation in the rhizome pith, where transcript levels are considerably lower.

It should be noted, though, that many factors also influence metabolite imaging, such as: different compounds having different ionization behaviours and so MALDI-imaging conditions must be optimized on a case-by-case basis. Besides that; results must be interpreted keeping in mind that biosynthetic intermediates might not accumulate in sufficient amounts for detection; different steps in a biosynthetic pathway can occur in different tissues/cell types; and metabolites can be transported to different tissues/cell types for storage, making the actual site of biosynthesis more complex to decipher.

Nevertheless, these technological advancements allow for understanding metabolism at a new level, with the potential to greatly shorten the time required for complete elucidation of natural product biosynthetic pathways and their temporal and spatial organization

Conclusions

The application of multi-omics analyses (transcriptomics, bioinformatics and metabolomics) enabled discovery of the aporphine alkaloid pathway in *Podophyllum* species, and the evolutionary/phylogenetic ramifications thereof. This discovery has potentially additional benefits, both medicinally and economically. This is because *P. hexandrum* rhizome extracts are the commercial source of PTOX (**1**) for the semi-synthesis of anticancer drugs, and the alkaloids corytuberine (**12**), magnoflorine (**13**) and muricinine (**24**) can be easily recovered from methanolic and aqueous extracts. They thus may be of additional economic value, since aporphine alkaloids display a wide array of biological activities,[27-29] ranging from antioxidant[28, 29] to anti-inflammatory[30] and psychoactive[31] properties. Indeed, their presence should also be considered in the context of the use of crude extracts of such plant species in folk medicine. The discovery of this alkaloid pathway thus also raises the possibility of increasing their levels using biotechnological approaches; this could be achieved via effective transformation and/or transient silencing/upregulation (e.g. virus-induced gene silencing).

Acknowledgements

We thank the National Science Foundation (MCB-1052557), the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (DE-FG-0397ER20259) and the G. Thomas and Anita Hargrove Center for Plant Genomic Research**,** for generous financial support. MALDI-MS based imaging analysis was performed on an instrument acquired through a Major Research Instrumentation grant (DBI-1229749) from the National Science Foundation. Bioinformatics analysis and gene cloning/recombinant protein expression work was supported by the NSF MCB-1052557 grant and the G. Thomas and Anita Hargrove Center for Plant Genomic Research; metabolite isolation, identification and *in situ* localization (using MALDI-IMS) was supported by the DOE, with instrumentation being from the NSF MRI support. We thank Dr. Toni Kutchan (Donald Danforth Plant Science Center) for magnoflorine, corytuberine and norcoclaurine standards.

Notes and references

[a] Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340, USA, Fax: +1 509 335 8206; Tel: +1 509 335 2682; E-mail: lewisn@wsu.edu.

Abbreviations: CNMT, coclaurine *N*-methyltransferase; CS, corytuberine synthase; CyP450, cytochrome P450; DP, dirigent protein, HNMC 4OMT, 3′-hydroxy-*N*-methyl-(*S*)-coclaurine 4′-*O*-methyltransferase; NC 6OMT, norcoclaurine 6-*O*-methyltransferase; NCS, norcoclaurine synthase; NM C3H, *N*-methylcoclaurine 3′-hydroxylase; PLR, pinoresinol-lariciresinol reductase; PLS, pluviatolide synthase; PSS, piperitol/sesamin synthase; PTOX, podophyllotoxin; SDH, secoisolariciresinol dehydrogenase; TBO, toluidine blue O.

† Electronic Supplementary Information (ESI) available: [Purification of *Podophyllum hexandrum* norcoclaurine synthase (Figure S1) and expression levels of norcoclaurine 6-*O*-methyltransferase, corytuberine synthase and coclaurine *N*-methyltransferase in *P. hexandrum* and *P. peltatum* (Figure S2)]. See DOI: 10.1039/b000000x/

1. M. Gordaliza, P. A. Garcia, J. M. del Corral, M. A. Castro and M. A. Gomez-Zurita, *Toxicon*, 2004, **44**, 441-459.
2. C. Canel, R. M. Moraes, F. E. Dayan and D. Ferreira, *Phytochemistry*, 2000, **54**, 115-120.
3. J. V. Marques, K.-W. Kim, C. Lee, M. A. Costa, G. D. May, J. A. Crow, L. B. Davin and N. G. Lewis, *J. Biol. Chem.*, 2013, **288**, 466-479.
4. N. Ikezawa, K. Iwasa and F. Sato, *Plant Cell Rep.*, 2009, **28**, 123-133.
5. E. Ono, M. Nakai, Y. Fukui, N. Tomimori, M. Fukuchi-Mizutani, M. Saito, H. Satake, T. Tanaka, M. Katsuta, T. Umezawa and Y. Tanaka, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 10116-10121.
6. Y. Jiao, L. B. Davin and N. G. Lewis, *Phytochemistry*, 1998, **49**, 387-394.
7. A. Pasquo, A. Bonamore, S. Franceschini, A. Macone, A. Boffi and A. Ilari, *Acta Crystallogr. Sect. Sect. F*, 2008, **64**, 281-283.

8. E. K. M. Ueda, P. W. Gout and L. Morganti, *J. Chromatogr.*, 2003, **988**, 1-23.
9. A. M. Patten, M. Jourdes, C. L. Cardenas, D. D. Laskar, Y. Nakazawa, B.-Y. Chung, V. R. Franceschi, L. B. Davin and N. G. Lewis, *Mol. BioSyst.*, 2010, **6**, 499–515.
10. A. T. Dinkova-Kostova, D. R. Gang, L. B. Davin, D. L. Bedgar, A. Chu and N. G. Lewis, *J. Biol. Chem.*, 1996, **271**, 29473-29482.
11. D. R. Gang, M. A. Costa, M. Fujita, A. T. Dinkova-Kostova, H.-B. Wang, V. Burlat, W. Martin, S. Sarkanen, L. B. Davin and N. G. Lewis, *Chem. Biol.*, 1999, **6**, 143-151.
12. Z.-Q. Xia, M. A. Costa, H. C. Pélissier, L. B. Davin and N. G. Lewis, *J. Biol. Chem.*, 2001, **276**, 12614-12623.
13. L. Y. P. Luk, S. Bunn, D. K. Liscombe, P. J. Facchini and M. E. Tanner, *Biochemistry*, 2007, **46**, 10153-10161.
14. C. Stevigny, J. L. H. Jiwan, R. Rozenberg, E. de Hoffmann and J. Quetin-Leclercq, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 523-528.
15. W. Wu and C. Huang, *Chin. Pharm. J.*, 2006, **58**, 41-55.
16. W. R. Dunstan and T. A. Henry, *J. Chem. Soc., Transactions*, 1898, **73**, 209-226.
17. F. B. Power, *Am. J. Pharm.*, 1878, **50**, 369-370.
18. F. B. Power, *Chem. News*, 1898, **78**, 26-27.
19. S. A. Wani, K. W. Shah and M. A. Ahmad, *Int. J. Biol. Pharm. Res.*, 2012, **3**, 684-686.
20. S. A. Wani, M. Ashfaq, K. W. Shah and D. Singh, *J. Curr. Chem. Pharm. Sc.*, 2012, **2**, 125-128.
21. D. K. Liscombe, B. P. MacLeod, N. Loukanina, O. I. Nandi and P. J. Facchini, *Phytochemistry*, 2005, **66**, 1374-1393.
22. Y.-D. Kim, S.-H. Kim, C. H. Kim and R. K. Jansen, *Biochem. Syst. Ecol.*, 2004, **32**, 291-301.
23. W. Wang, Z.-D. Chen, Y. Liu, R.-Q. Li and J.-H. Li, *Syst. Bot.*, 2007, **32**, 731-742.
24. A. J. Broomhead and P. M. Dewick, *Phytochemistry*, 1990, **29**, 3831-3837.
25. R. W. Jiang, J. R. Zhou, P. M. Hon, S. L. Li, Y. Zhou, L. L. Li, W. C. Ye, H. X. Xu, P. C. Shaw and P. P. But, *J. Nat. Prod.*, 2007, **70**, 283-286.
26. S. König, K. Feussner, A. Kaever, M. Landesfeind, C. Thurow, P. Karlovsky, C. Gatz, A. Polle and I. Feussner, *New Phytol.*, 2014, **202**, 823-837.
27. S. M. Mohamed, E. M. Hassan and N. A. Ibrahim, *Nat. Prod. Res.*, 2009, **24**, 1395-1402.
28. T. M. Hung, J. P. Lee, B. S. Min, J. S. Choi, M. K. Na, X. F. Zhang, T. M. Ngoc, I. Lee and K. H. Bae, *Biol. Pharm. Bull.*, 2007, **30**, 1157-1160.
29. T. M. Hung, M. K. Na, B. S. Min, X. F. Zhang, I. S. Lee, T. M. Ngoc, P. T. Thuong, D. E. Sok and K. H. Bae, *Planta Med.*, 2007, **73**, 1281-1284.
30. E. Küpeli, M. Kosar, E. Yesilada, K. Husnu and C. Baser, *Life Sci.*, 2002, **72**, 645-657.
31. J. B. de la Peña, H. L. Lee, S. Y. Yoon, G. H. Kim, Y. S. Lee and J. H. Cheong, *J. Nat. Med.*, 2013, **67**, 814-821.

