Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 66 (2005) 1374–1393
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms

David K. Liscombe [a], Benjamin P. MacLeod [a], Natalia Loukanina [a], Owi I. Nandi [b], Peter J. Facchini [a,*]

[a] *Department of Biological Sciences, University of Calgary, Calgary, Alta., Canada T2N 1N4*
[b] *Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, CH-8008 Zürich, Switzerland*

Received 14 March 2005; received in revised form 1 April 2005
Available online 31 May 2005

Abstract

Benzylisoquinoline alkaloids (BIAs) consist of more than 2500 diverse structures largely restricted to the order Ranunculales and the eumagnoliids. However, BIAs also occur in the Rutaceae, Lauraceae, Cornaceae and Nelumbonaceae, and sporadically throughout the order Piperales. Several of these alkaloids function in the defense of plants against herbivores and pathogens – thus, the capacity for BIA biosynthesis is expected to play an important role in the reproductive fitness of certain plants. Biochemical and molecular phylogenetic approaches were used to investigate the evolution of BIA biosynthesis in basal angiosperms. The occurrence of (*S*)-norcoclaurine synthase (NCS; EC 4.2.1.78) activity in 90 diverse plant species was compared to the distribution of BIAs superimposed onto a molecular phylogeny. These results support the monophyletic origin of BIA biosynthesis prior to the emergence of the eudicots. Phylogenetic analyses of NCS, berberine bridge enzyme and several *O*-methyltransferases suggest a latent molecular fingerprint for BIA biosynthesis in angiosperms not known to accumulate such alkaloids. The limited occurrence of BIAs outside the Ranunculales and eumagnoliids suggests the requirement for a highly specialized, yet evolutionarily unstable cellular platform to accommodate or reactivate the pathway in divergent taxa. The molecular cloning and functional characterization of NCS from opium poppy (*Papaver somniferum* L.) is also reported. Pathogenesis-related (PR)10 and Bet v 1 major allergen proteins share homology with NCS, but recombinant polypeptides were devoid of NCS activity.

Keywords: Benzylisoquinoline alkaloid; Berberine bridge enzyme; Metabolic diversity; Methyltransferase; Norcoclaurine synthase; Opium poppy; Secondary metabolism

1. Introduction

The benzylisoquinoline alkaloids (BIAs) are a complex and diverse group of natural products consisting of more than 2500 known structures. (*S*)-Norcoclaurine, is the ultimate precursor to all BIAs and is formed by the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) by (*S*)-norcoclaurine synthase (NCS; EC 4.2.1.78). The structural diversity of BIAs results from the elaboration of (*S*)-norcoclaurine by a highly branched metabolic pathway consisting several enzyme types including *N*-methyltransferases, *O*-methyltransferases, cytochromes P450, and the FAD-dependent berberine bridge enzyme (BBE; Steffens et al., 1985). Several BIA biosynthetic genes have been isolated and the corresponding enzymes functionally characterized (Facchini, 2001). Skeletal structures representing the major BIA classes are illustrated in Fig. 1. The accumulation of specific BIAs has been documented in diverse taxa throughout the angiosperms. BIAs are most common among the Ranunculales and eumagnoliids,

* Corresponding author. Tel.: +1 403 220 7651; fax +1 403 289 9311.
E-mail address: pfacchin@ucalgary.ca (P.J. Facchini).

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.04.029

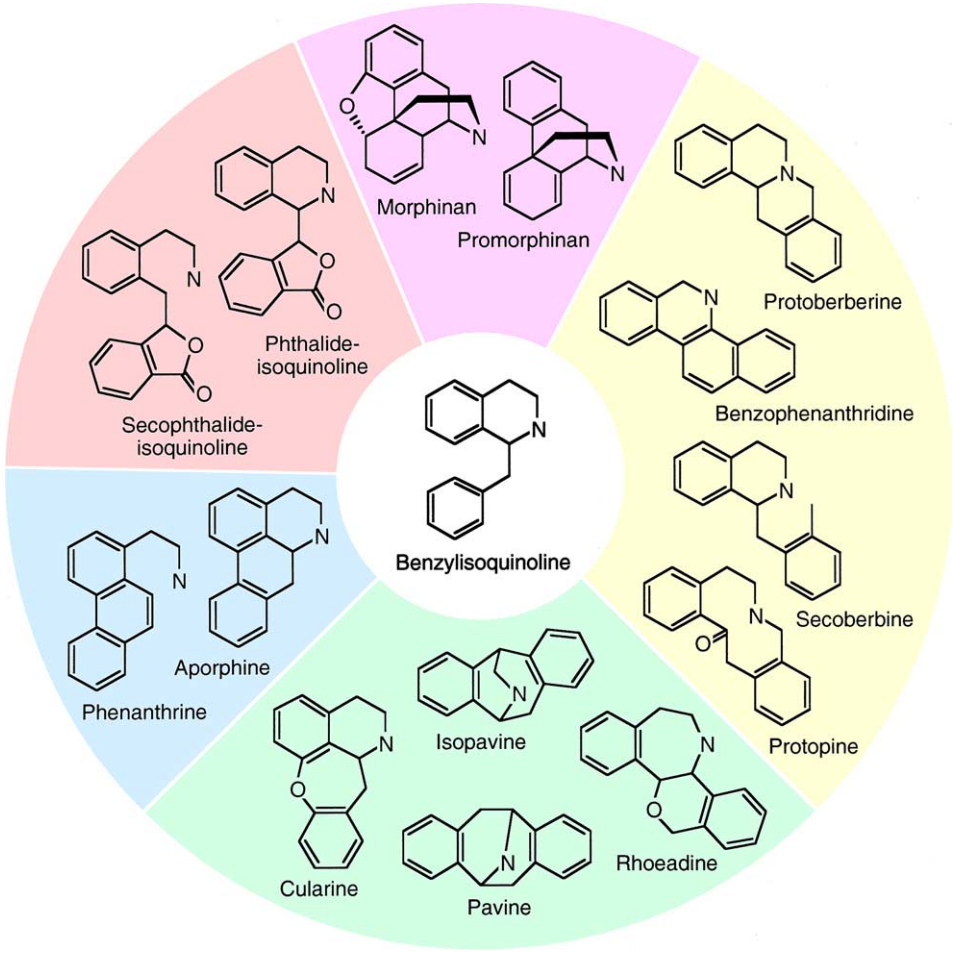

Fig. 1. Benzylisoquinoline alkaloids (BIAs) are all based on the elaboration of a simple skeletal structure. Skeletal representations of the known stereochemical conformations and functional group substitutions found in major BIA families are shown. Alkaloid families were grouped according to a common biosynthetic origin with the exception of the isopavine, cularine, pavine, and rhoeadine group, which is composed of small BIA classes found only in a restricted number of taxa.

but are also found in the distantly related families Rutaceae, Lauraceae, Cornaceae and Nelumbonaceae, and occur sporadically throughout the order Piperales (Shulgin and Perry, 2002). The general role of alkaloids in the chemical defense of plants against herbivores and pathogens suggests that BIAs contribute to the reproductive fitness of plants with the ability to produce these compounds (Schmeller et al., 1997; Wink, 2003). The ability to produce BIAs represents a trait that has been subjected to natural selection through evolution (Wink, 2003).

The evolution of secondary metabolism has long been the subject of considerable speculation (Firn and Jones, 2003; Theis and Lerdau, 2003; Wink, 2003), but few empirical studies have been reported. Elegant studies of the evolution of pyrrolizidine alkaloid biosynthesis have demonstrated that the recruitment of the enzyme catalyzing the first committed step in the pathway, homospermidine synthase (HSS), from deoxyhypusine synthase (DHS), which is involved in transcriptional activation, has occurred independently in at least three distant plant families, including twice in the Asteraceae (Ober and Hartmann, 2000; Anke et al., 2004). These data support a model for the polyphyletic evolution of pyrrolizidine alkaloid biosynthesis, which may explain the sporadic distribution of this alkaloid class. In contrast, Wink (2003) has suggested that the widespread phylogenetic distribution of other alkaloid groups results from the differential expression of biosynthetic pathways with a monophyletic origin. This model supports the taxonomic distribution of tropane and quinolizidine alkaloids (Wink, 2003), and the terpenoid indole alkaloid, camptothecin (Lorence and Desster, 2004).

Sequence analysis of specific biosynthetic enzymes has revealed unique features about the evolution of the corresponding secondary metabolic pathways. In phenylpropenoid biosynthesis the eugenol *O*-methyltransferases from *Ocimum basilicum* are more closely related to flavonoid *O*-methyltransferases than to (iso)eugenol *O*-methyltransferase from *Clarkia breweri*, indicating that similarities in substrate recognition have resulted from convergent evolution (Gang et al., 2002; Wang and

Pichersky, 1999). Analysis of tropinone reductase sequences from *Hyoscyamus niger* and *Datura stramonium* has revealed homologs from several diverse taxa, indicating that they occur ubiquitously throughout the angiosperms (Nakajima et al., 1999). Phylogenetic analysis of plant terpene synthases revealed a monophyletic origin for these enzymes (Bohlmann et al., 1998). Terpene synthases involved in secondary metabolism apparently diverged from those participating in primary metabolism prior to the separation of angiosperms and gymnosperms. BIA biosynthetic enzymes have not been subjected to phylogenetic analysis and little is known about their origin.

The gene encoding NCS, the gateway to BIA biosynthesis (Rueffer et al., 1981; Stadler et al., 1987, 1989; Samanani and Facchini, 2002), was cloned from *Thalictrum flavum* and shown to share homology with three major protein families – the pathogenesis-related (PR)10, Bet v 1 major allergen (MAP) and ripening-related (RRP) proteins (Samanani et al., 2004). Although *T. flavum* NCS was recruited from a PR10/Bet v 1 ancestor, it is not known whether NCS is derived from the same ancestral gene in all BIA-producing plants. Sequence data for other BIA biosynthetic genes is available for some restricted members of Papaveraceae, Ranunculaceae and Berberidaceae, but the evolutionary origins of these enzymes have not been investigated.

In this paper, we report an empirical investigation into the evolution of BIA biosynthesis using a combination of biochemical and molecular phylogenetic approaches. We also identify and functionally characterize the gene encoding NCS from opium poppy (*Papaver somniferum* L.). Molecular phylogenetic analysis of BIA biosynthetic enzymes provides additional insight into the evolution of the pathway.

2. Results

2.1. Phylogeny of BIA distribution among selected angiosperms

Known classes of BIAs were grouped according to similarities in carbon skeleton and biosynthetic origin (Fig. 1). According to the scheme presented in Fig. 1, the distribution of alkaloids derived from the benzylisoquinoline nucleus was superimposed on a molecular phylogeny reconstructed for genera used in this study (Fig. 2). A small number of gymnosperms also investigated were not included to improve the accuracy of the phylogenetic analysis. Published accounts of the accumulation of specific alkaloids in each genus were compiled, and the occurrence of relevant alkaloids was mapped according to family on the molecular phylogeny presented in Fig. 2. Data presented in Fig. 2 should be interpreted with some degree of caution since phytochemical analysis remains incomplete for several genera used in this study and some published accounts could, in fact, be incorrect. Thus, genera shown not to contain BIAs include those that have been thoroughly investigated, such as *Podophyllum*, and those on which little work has been done, such as *Chloranthus*. Genera reported to produce any class of complex BIA can be assumed to also contain simple BIAs, from which they are derived (Fig. 1). Proaporphine alkaloids and aristolochic acid derivatives (Castedo and Tojo, 1990; Kumar et al., 2003) were clustered with the aporphine and phenanthrene alkaloids. The protoberberine, benzophenanthridine, protopine and secoberbine alkaloids were grouped together based on their common biogenesis (Shamma et al., 1978; Rozwadowska, 1988; Facchini, 2001). The promorphinan and morphinan alkaloids and the phthalideisoquinoline and secophthalideisoquinoline alkaloids represent two additional and distinct groups of complex BIAs sharing a common biosynthetic origin. The pavine, isopavine, cularine, and rhoeadine alkaloids were clustered as a miscellaneous group representing BIAs found only in restricted taxa (Gozler et al., 1983; Sariyar, 2002).

Among the genera shown in Fig. 2, the production of simple BIAs alone is restricted to the genus *Nymphaea* (Shulgin and Perry, 2002). The aporphine and phenanthrene alkaloids are the most widely distributed (Fig. 2). These alkaloids occur in all reported genera of the Papaveraceae (except for *Eschscholzia*), Menispermaceae, Ranunculaceae (except for certain genera in which BIA production has not been investigated) and Berberidaceae (except for *Podophyllum*, *Vancouveria* and *Jeffersonia* species) (Hocquemiller et al., 1984; Castedo and Tojo, 1990; Sariyar, 2002; Shulgin and Perry, 2002). Aporphine and phenanthrene alkaloids are the only BIAs produced by members of the Lauraceae and Nelumbonaceae (Shulgin and Perry, 2002), and occur sporadically throughout the Rutaceae (e.g. *Zanthoxylum* and *Melicope*), Magnoliaceae (e.g. *Liriodendron*), Saururaceae (e.g. *Houttuynia*) and Aristolochiaceae (e.g. *Aristolochia*) (Castedo and Tojo, 1990; Kim et al., 2001; Shulgin and Perry, 2002).

Protoberberine and benzophenanthridine alkaloids also show extensive taxonomic distribution. Among genera used herein, alkaloids of protoberberine/benzophenanthridine group are produced in all plants of the Papaveraceae and Menispermaceae, and several members of the Berberidaceae, Ranunculaceae and Rutaceae (Fig. 2: Shamma et al., 1978; Rozwadowska, 1988; Sariyar, 2002; Shulgin and Perry, 2002). Protoberberine and benzophenanthridine alkaloids also show sporadic distribution throughout the angiosperms in genera such as *Alangium*, *Aristolochia* and *Liriodendron* (Pakrashi et al., 1983; Lopes and Humpfer, 1997; Rastrelli et al., 1997; Shulgin and Perry, 2002). Interestingly, 8-benzylberbine has also been isolated from the gymnosperm *Gnetum parvifolium* (Xu and Lin, 1999).

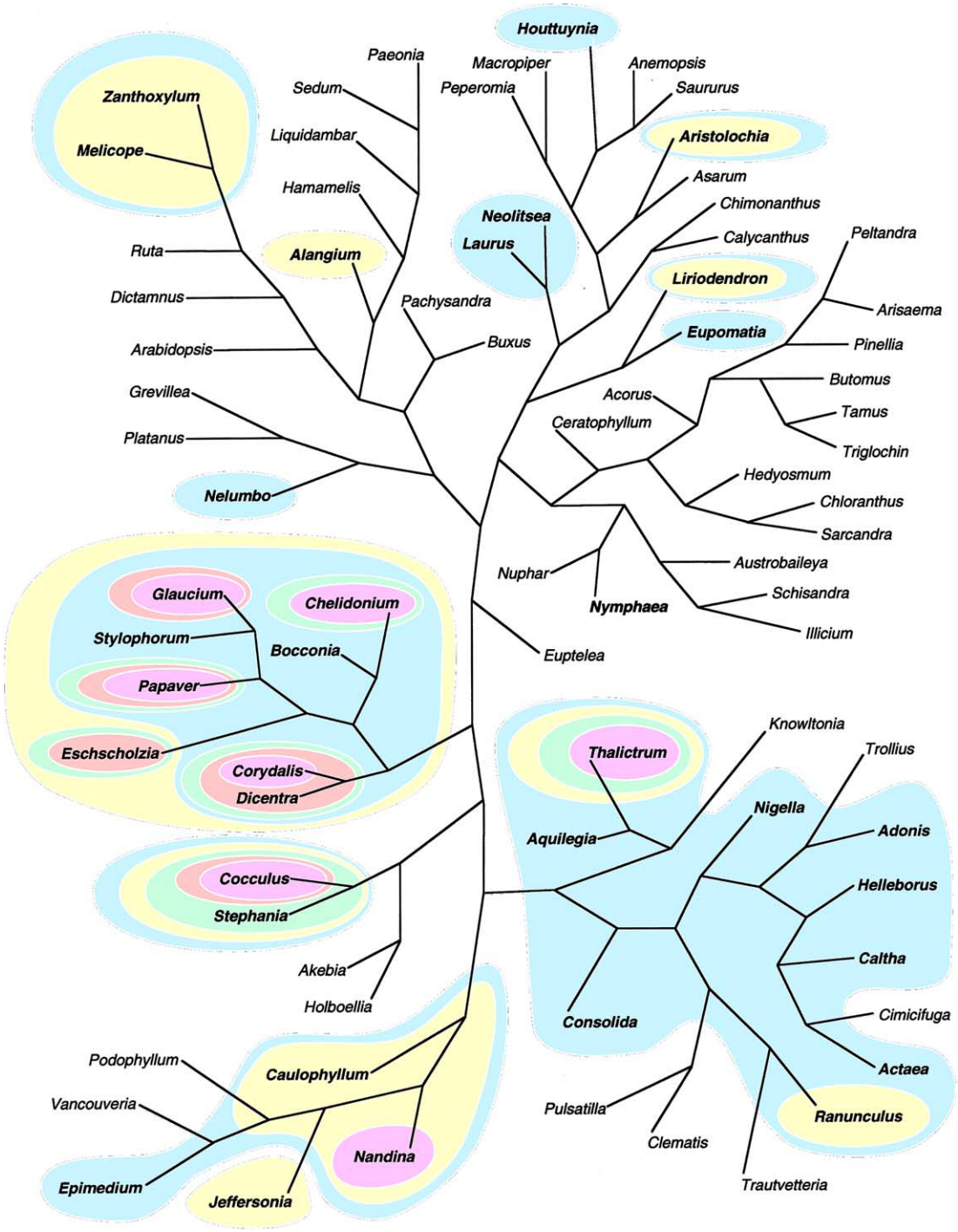

Fig. 2. The distribution of major groups of benzylisoquinoline alkaloids (BIAs) in selected angiosperm genera. The overlay of chemotaxonomic data for each genus is merged at the family level. The chemotaxonomic distribution is color-coded to match the BIA biosynthetic groups shown in Fig. 1. Nymphaea, shown in bold, has been reported to contain only simple BIAs. The tree is representative of several 50% majority-rule consensus trees from MrBayes analyses of combinations of *5.8S*, *18S*, *26S*, *atpB*, *ndhF*, *phyA*, *phyC* and *rbcL* sequences using a GTR + Γ + I model. One to eight molecular markers were available for each genus, often from different species. Branches are not drawn to scale. Accessions are listed in Section 5.

Morphinan and promorphinan alkaloids are restricted to the Ranunculales, where they are produced by certain genera in the Papaveraceae (e.g. *Chelidonium*, *Corydalis*, *Glaucium* and *Papaver*), Menispermaceae (e.g. *Cocculus*) and Berberidaceae (e.g. *Thalictrum* and *Nandina*) (Hocquemiller et al., 1984; Sariyar, 2002; Shulgin and Perry, 2002). Among the genera shown in Fig. 2, the phthalide- and seco-phthalideisoquinolines are restricted to *Eschscholzia*, *Papaver*, *Dicentra* and *Glaucium* in the Papaveraceae, and *Cocculus* in the Menispermaceae

(Rozwadowska, 1988; Shulgin and Perry, 2002). The pavine, isopavine, cularine and rhoeadine alkaloids represent small, relatively unrelated groups of BIAs found in all genera of the Papaveraceae except *Stylophorum*, *Glaucium* and *Bocconia* (Gozler et al., 1983; Sari and Sariyar, 1997; Sariyar, 2002; Shulgin and Perry, 2002). Alkaloids of this miscellaneous group are also found in the Menispermaceae and Ranunculaceae, where they are limited to the genus *Thalictrum* (Gozler et al., 1983; Shulgin and Perry, 2002).

2.2. NCS activity screening of related species

Root, stem and leaf material collected from 85 angiosperm and five gymnosperm species were assayed for NCS activity in total soluble protein extracts (Table 1). NCS activity was detected in almost all genera reported to produce BIAs, and was particularly high in members of the Papaveraceae, Ranunculaceae and Berberidaceae known to accumulate an array of complex BIAs. However, NCS activity was detected in several plants not reported to produce BIAs including the gymnosperms *Cephalotaxus fortunei* and *Taxodium distichum*, several members of the Chloranthaceae, *Vancouveria hexandra* (Berberidaceae), *Clematis vitalba* (Ranunculaceae) and *Sedum reflexum*. Several other species displayed low levels of NCS activity including the gymnosperm *Ephedra fragilis*, the basal angiosperms *Schisandra* spp., *Austrobaileya scandens*, *Tamus communis*, *Illicium lanceolatum*, *Eupomatia* spp., *Chimonanthus nitens*, *Calycanthus chinensis*, *Peperomia reflexa*, *Saururus cernuus*, in addition to *Paeonia mlokosewitschii*, *Trautvetteria carolinensis*, *Akebia trifoliata*, and *Ruta graveolens*.

NCS activity was not detected in six species despite reports of BIA production in members of the same genus. These include *Gnetum gnemon*, *Nymphaea alba*, *Ranunculus acris*, *Actaea rubra*, *Alangium platanifolium* and *Aristolochia rotunda* (Pakrashi et al., 1983; Mukherjee et al., 1986; Rastrelli et al., 1997; Xu and Lin, 1999; Shulgin and Perry, 2002). NCS activity was not detected in most species in which BIAs have also not been reported including *Knowltonia filia* and *Pulsatilla vulgaris* (Ranunculaceae), *Holboellia coriacea* and *Arabidopsis thaliana*.

2.3. Identification, isolation and functional characterization of opium poppy NCS

The recently reported molecular cloning and characterization of a NCS cDNA (Samanani et al., 2004) provided a query sequence to identify a putative, partial cDNA encoding opium poppy NCS by in silico mining of an opium poppy expressed sequence tag (EST) database. Screening of the cDNA library with the partial NCS cDNA led to the isolation of six clones, including three full-length cDNAs. Sequence analysis showed that the isolated cDNAs represented two distinct putative NCS isoforms, PsNCS1 and PsNCS2. The PsNCS1 cDNA contained a 696-bp open-reading frame (ORF) flanked by a 47-bp 5′-untranslated region (UTR), and a 160-bp 3′-UTR followed by a poly(A) tract. The PsNCS2 cDNA contained a 696-bp ORF flanked by a 4-bp 5′-UTR, and followed by a 195-bp 3′-UTR and a poly(A) tract. Both ORFs encoded predicted translation products of 232 amino acids with molecular masses of 26 kDa. The amino acid identity between PsNCS1 and PsNCS2 is 89%, with most of the divergence occurring in the first 50 residues. Two full-length cDNAs encoding pathogenesis-related proteins, PsPR10-1 and PsPR10-2, were also isolated from the opium poppy EST database as the closest homologs of NCS. Sequences for all reported clones have been deposited in the GenBank database under the accession numbers AY860500 (NCS1), AY860501 (NCS2), AY861682 (PsPR10-1), and AY861683 (PsPR10-2).

Heterologous expression of full-length PsNCS1 or PsNCS2 ORFs produced recombinant protein of the predicted molecular mass in total protein extracts from bacterial cell cultures induced with IPTG (data not shown). Recombinant proteins were not detected in extracts from cultures harvested prior to the addition of IPTG or in extracts of cells containing the empty pET29b vector in the absence or presence of IPTG (data not shown). NCS activity was detected in protein extracts containing $(His)_6$-tagged, recombinant PsNCS1 (data not shown) and PsNCS2 (Fig. 3), but was absent from control extracts of cells containing the empty pET29b vector.

An alignment of predicted amino acid sequences for PsNCS1, TfNCS from *T. flavum* (Samanani et al., 2004), and related PR10 and MAP proteins from a variety of plant species reveals considerable sequence identity (Fig. 4). NCS from *P. somniferum* and *T. flavum* are unique among known MAP and PR10 proteins due to oligopeptide extensions at both the N- and C-termini of the predicted translation products. A putative N-terminal signal peptide was predicted in TfNCS using the SignalP Prediction Server (http://www.cbs.dtu.dk/services/signalp) (Neilsen et al., 1997; Samanani et al., 2004). However, similar analysis of PsNCS1 did not predict an N-terminal signal peptide. Indeed, the first 20 N-terminal residues in PsNCS1 share little identity with the corresponding region of TfNCS. Moreover, PsNCS1 and PsNCS2 possess a prominent C-terminal oligopeptide that extends beyond the C-terminus of TfNCS.

NCS proteins also share homology with MLP and RRP, although to a lesser degree than to MAP and PR10 proteins (Fig. 4). Percentage amino acid identities determined from pair-wise global alignments of all sequences aligned in Fig. 4 are compiled in Table 2. Although the identity between PsNCS1 and TfNCS is

Table 1
Norcoclaurine synthase activity in total soluble protein extracts from various plant species

Family	Species	NCS activity, pmol s^{-1} mg $protein^{-1}$		
		Root ($\times 10^{-1}$)	Stem ($\times 10^{-1}$)	Leaf ($\times 10^{-1}$)
Ephedraceae	*Ephedra distachya*	0	0	2
Ephedraceae	*Ephedra fragilis*	0	nd	9
Gnetaceae	*Gnetum gnemon*	0	nd	0
Cupressaceae	*Taxodium distichum*	13	nd	nd
Cephalotaxaceae	*Cephalotaxus fortunei*	70	21	3
Dioscoreaceae	*Tamus communis*	4	4	2
Acoraceae	*Acorus calamus*	0	0	nd
Acoraceae	*Acorus gramineus var. variegatus*	0	0	1
Araceae	*Lysichiton camtschatcense*	nd	0	0
Araceae	*Peltandra virginica*	0	0	0
Araceae	*Arisaema flavum*	nd	0	0
Araceae	*Pinellia tuberifera* var. *pedatidecta*	0	1	0
Butomaceae	*Butomus umbellatus*	nd	0	0
Juncaginaceae	*Triglochin maritima*	0	0	0
Nymphaeaceae	*Nuphar lutea*	0	0	0
Nymphaeaceae	*Nymphaea alba*	0	0	0
Schisandraceae	*Schisandra sphaenanthera*	nd	2	1
Schisandraceae	*Schisandra chinensis*	nd	5	nd
Austrobaileyaceae	*Austrobaileya scandens*	2	4	5
Illiaceae	*Illicium lanceolatum*	nd	0	9
Ceratophyllaceae	*Ceratophyllum demersum*	nd	0	0
Chloranthaceae	*Chloranthus serratus*	26	2	12
Chloranthaceae	*Chloranthus* sp.	26	37	8
Chloranthaceae	*Hedyosmum racemosum*	8	0	7
Chloranthaceae	*Sarcandra chloranthoides*	0	0	0
Chloranthaceae	*Sarcandra glabra*	7	20	10
Magnoliaceae	*Liriodendron chinense*	8	1	5
Eupomatiaceae	*Eupomatia bennettii*	0	3	0
Eupomatiaceae	*Eupomatia laurina*	3	6	1
Calycanthaceae	*Chimonanthus nitens*	nd	4	0
Calycanthaceae	*Calycanthus chinensis*	0	3	7
Lauraceae	*Neolitsea sericea*	39	11	59
Lauraceae	*Laurus azorica*	3	2	2
Piperaceae	*Macropiper excelsum*	11	3	0
Piperaceae	*Peperomia reflexa*	nd	6	1
Saururaceae	*Anemopsis californica*	0	nd	0
Saururaceae	*Houttuynla cordata* var. *flora plena*	24	nd	0
Saururaceae	*Saururus cernuus*	0	6	0
Aristolochiaceae	*Aristolochia rotunda*	0	0	nd
Aristolochiaceae	*Asarum europaeum*	3	2	0
Eupteleaceae	*Euptelea polyandra*	0	0	0
Papaveraceae	*Chelidonium majus*	222	84	30
Papaveraceae	*Papaver rhoeas*	89	24	17
Papaveraceae	*Papaver somniferum*	206	122	22
Papaveraceae	*Dicentra formosa*	100	46	44
Papaveraceae	*Stylophorum diphyllum*	115	68	5
Papaveraceae	*Glaucium flavum*	137	93	36
Papaveraceae	*Bocconia frutescens*	82	66	55
Papaveraceae	*Eschscholzia californica*	158	48	33
Papaveraceae	*Corydalis cheiranthifolia*	42	58	81
Ranunculaceae	*Adonis vernalis*	27	7	9
Ranunculaceae	*Aquilegia vulgaris*	8	7	16
Ranunculaceae	*Clematis vitalba*	4	0	1
Ranunculaceae	*Helleborus viridis*	5	1	25
Ranunculaceae	*Nigella damascena*	0	7	2
Ranunculaceae	*Pulsatilla vulgaris*	3	2	2
Ranunculaceae	*Ranunculus acris*	0	0	0
Ranunculaceae	*Actaea rubra*	1	0	0
Ranunculaceae	*Caltha palustris*	17	nd	0
Ranunculaceae	*Cimicifuga racemosa*	0	0	nd
Ranunculaceae	*Aquilegia ecalcarata*	66	14	1

(*continued on next page*)

Table 1 (*continued*)

Family	Species	NCS activity, pmol s^{-1} mg $protein^{-1}$		
		Root ($\times 10^{-1}$)	Stem ($\times 10^{-1}$)	Leaf ($\times 10^{-1}$)
Ranunculaceae	*Trautvetteria carolinensis* var. *occidentalis*	2	5	0
Ranunculaceae	*Trollius chinensis*	nd	1	nd
Ranunculaceae	*Consolida ajacis*	17	32	22
Ranunculaceae	*Knowltonia filia*	0	1	0
Ranunculaceae	*Thalictrum flavum*	31	47	3
Ranunculaceae	*Thalictrum aquilegifolium*	14	10	16
Berberidaceae	*Epimedium pinnatum*	43	20	nd
Berberidaceae	*Podophyllum peltatum*	0	0	0
Berberidaceae	*Caulophyllum thalictroides*	69	46	47
Berberidaceae	*Jeffersonia diphylla*	66	12	14
Berberidaceae	*Vancouveria hexandra*	87	21	4
Berberidaceae	*Nandina domestica*	165	78	0
Lardizabalaceae	*Holboellia coriacea*	nd	2	1
Lardizabalaceae	*Akebia trifoliata*	5	0	0
Menispermaceae	*Stephania japonica*	16	14	13
Menispermaceae	*Cocculus trilobus*	10	6	4
Nelumbonaceae	*Nelumbo nucifera*	0	0	5
Platanaceae	*Platanus orientalis*	0	0	nd
Proteaceae	*Grevillea banksii*	0	0	4
Buxaceae	*Pachysandra terminalis*	0	0	0
Buxaceae	*Buxus sempervirens*	0	1	0
Hamamelidaceae	*Liquidambar styraciflua*	nd	0	0
Paeoniaceae	*Paeonia mlokosewitschii*	9	0	nd
Crassulaceae	*Sedum reflexum*	0	15	8
Brassicaceaceae	*Arabidopsis thaliana*	0	0	0
Rutaceae	*Dictamnus albus*	nd	0	1
Rutaceae	*Ruta graveolens*	2	6	4
Rutaceae	*Zanthoxylum simulans*	nd	23	nd
Rutaceae	*Melicope hupehensis*	1	0	0
Cornaceae	*Alangium platanifolium* var. *macrophyllum*	nd	0	0

Values represent the mean of duplicate experiments. The detection limit for norcoclaurine formation was 540 pmol above background. Nd = not determined.

only 39% when compared by global alignment, the identity increases to 50% when a local alignment excluding divergent N- and C-terminal domains is performed. Although the overall identity between NCS and MAP/PR10 proteins, and between NCS and MLP/RRP sequences are similar, the conserved motifs are distinct in each case (Fig. 4). For example, a highly conserved G(D/N)GGXGT(V/I) motif in PsNCS1 is absent in MLP and RRP sequences. Only Gly_{91} and Gly_{128} in PsNCS1 are absolutely conserved among all known NCS, MLP, RRP, MAP and PR10 proteins.

Six recombinant proteins, PsPR10-1, PsPR10-2, OsPR10, VrCSBP, HpHYP1, BpMAP (Bet v 1a), or PmPR10, that share the most extensive sequence identity with PsNCS1 were assayed for NCS activity. None of the recombinant proteins showed any detectable NCS activity, in comparison with the high levels of activity displayed by recombinant PsNCS2 and TfNCS (Fig. 3).

2.4. Phylogenetic analysis of BIA biosynthetic enzymes

A Bayesian inference, implemented using Markov chain Monte Carlo (MCMC) in MrBayes under a GTR + Γ + I 4by4 nucleotide model allowing rates to vary within and between partitions, was applied to NCS sequences combined with representative PRs, MLP and CBP sequence(s). Fig. 5 is a representative 50% majority rule consensus tree from a MCMC run of 1×10^6 cycles with a burn in value of 1.5×10^5 cycles. The overall topology includes three major groups of approximately equal phylogenic divergence and representing MAP/PR10, MLP/RRP and NCS sequences. Notable exceptions to this correlation between relatedness and characterized are the PsPR10-1 and PsPR10-2 sequences, which are more similar to PsNCS isoforms than to PR10 proteins from other species. Sequences included in Fig. 5 represent proteins from highly divergent plants including *Pinus monticola* (PmPR10) and *Hyacinthus orientalis* (HoPR10) PR10 homologs as non-eudicot representatives, which is apparent from their basal position in the PR10 grouping. This indicates the divergence of these three related groups might outdate the divergence of angiosperms and gymnosperms. Less phylogenetic divergence is found in the MLP group, where all MLP homologs are from eudicot species. Although *P. somniferum* MLP (PsMLP) is the most related of the MLP group to the NCS sequences (Fig. 5), this arrangement does not have high posterior probability

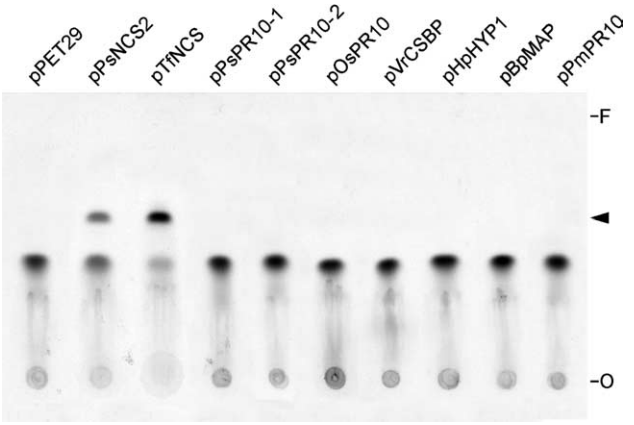

Fig. 3. Norcoclaurine synthase (NCS) activity of recombinant proteins produced by heterologous expression of coding sequences for NCS, pathogenesis-related (PR)10 and major allergen (MAP) proteins in *Escherichia coli*. Total soluble protein extracts from induced *E. coli* cells harboring expression constructs encoding the following proteins: pET29, empty vector; pPsNCS2, *Papaver somniferum* NCS2; pTfNCS, *Thalictrum flavum* NCS; pPsPR10-1, *Papaver somniferum* PR10-1; pPsPR10-2, *Papaver somniferum* PR10-2; pOsPR10, *Oryza sativa* PR10; pVrCSBP, *Vigna radiata* cytokinin-specific binding protein; pHpHYP1, *Hypericum perforatum* HYP1; pBpMAP, *Betula pendula* MAP (Bet v 1a); pPmPR10, *Pinus monticola* PR10-1.1. The arrowhead shows the migration distance of authentic (*S*)-norcoclaurine. Abbreviations: O, origin; F, front.

support. The major alternate position for PsMLP is between AtMLP and VvPRP, meaning it is not necessarily more related to NCS sequences than other MLP-homologs. The two sequences with the nearest phylogenetic relationship to the NCS group are a functionally uncharacterized *Oryza sativa* PR10 homolog (OsPR10) and a *Vigna radiata* cytokinin-specific binding protein sequence (VrCSBP). These sequences swap positions in different MCMC runs, accounting for the relatively low posterior probability support for their inclusion in the clades shown (Fig. 5).

Many alternative topologies are available for the depicted NCS phylogenetic tree as indicated by the relatively low posterior probability support for several of the clades, however, most of these are due to swapped locations between closely related sequences, as for OsPR10 and VrCSBP. The low support for monophyletic origins of functional NCS genes (Fig. 5; 0.54) is due to alternate topologies, which include PsPR10-1 and/or PsPR10-2 as a clade with PsNCS1 (Fig. 5). In these alternative trees, TfNCS is an outgroup to PsNCS1 and one or both of the PsPR10 sequences. However, if the PsPR10 sequences are considered as nonfunctional evolutionary products of PsNCS1 the NCS sequences are still a monophyletic clade.

For the phylogenetic analysis of BBE related sequences, ORFs of four known BBE genes, 22 predicted FAD-dependant genes from *Arabidopsis thaliana*, four predicted FAD-dependant carbohydrate oxidases from other species and a *Vigna unguiculata* drought-inducible gene were included. Fig. 6 shows a 50% majority rule consensus tree from a MCMC run of 1×10^6 cycles in MrBayes using the GTR + Γ + I 4by4 nucleotide model with a burn in value of 2×10^5 cycles. The tree topology from other MCMC runs was similar. Topology around the functionally characterized BBE genes is supported by a posterior probability of 1.00 (Fig. 6), and remained unchanged in all alternative 50% majority rule consensus trees. The nearest related sequence, a functionally uncharacterized *O. sativa* genomic clone with a FAD-binding motif, is quite distant. No genes currently available on public databases show a more significant phylogenetic relationship to functionally characterized BBE genes.

The lowest support for clades on the tree shown in Fig. 6 are for the inclusion of AtFAD17 through OsFAD1 as a clade (0.59) and the inclusion of AtFAD17 through AtFAD20 as a clade (0.60). Both are due to a major alternate topology with an alteration in the position of the AtFAD17-AtFAD16 group. The inclusion of LcFAD with AtFAD21 and AtFAD22 represents the only other major topology variation seen in 50% majority rule consensus trees from additional MCMC runs. Interestingly, while the BBE function appears to be monophyletic among the genes sampled, the carbohydrate oxidases do not. *Helianthus annuus* (HaCOX) and *Latuca sativa* (LsCOX) carbohydrate oxidase sequences cluster with a number of uncharacterized *A. thaliana* FAD-dependant proteins, whereas a *Nicotiana langsdorfii* × *Nicotiana sanderae* glucose oxidase (NlCOX) appears to be closely related to a different clade in which the only other functionally characterized gene is the *Cannabis sativa* tetrahydrocannabinolic acid synthase (CsTAS).

For analysis of selected OMT sequences, an initial alignment of 50 sequences was performed and then trimmed to sufficiently cover OMT sequence space and eliminate highly homologous sequences. Fig. 7 shows a 50% majority rule consensus tree from one of the Bayesian MCMC runs of 1×10^6 cycles with a burn in value of 2×10^5 cycles. Other runs produced a similar topology. High posterior probability support was indicated for a clade including norcoclaurine 6-*O*-methyltransferase (Cj6OMT, Tf6OMT, Ps6OMT) and 3′-hydroxy-*N*-methylnorcoclaurine 4′-*O*-methyltransferase (Cj4′OMT, Tf4′OMT, Ps4′OMT), along with *C. japonica* columbamine *O*-methyltransferase (CjCoOMT). Strong support was seen for a separate monophyletic clade including the *C. japonica* and *T. flavum* scoulerine *O*-methyltransferases (CjCOMT and TfCOMT). The *P. somniferum* (*R,S*)-reticuline 7-*O*-methyltransferase (Ps7OMT) appears to have a unique origin. Interestingly, representative *P. somniferum* and *T. tuberosum* catechol *O*-methyltransferases cluster with *O. basilicum* and *C. breweri* *O*-methyltransferases of phenylpropanoid biosynthesis, again with strong support.

The *O*-methyltransferase tree displays the most differential rates of nucleotide substitution, compared

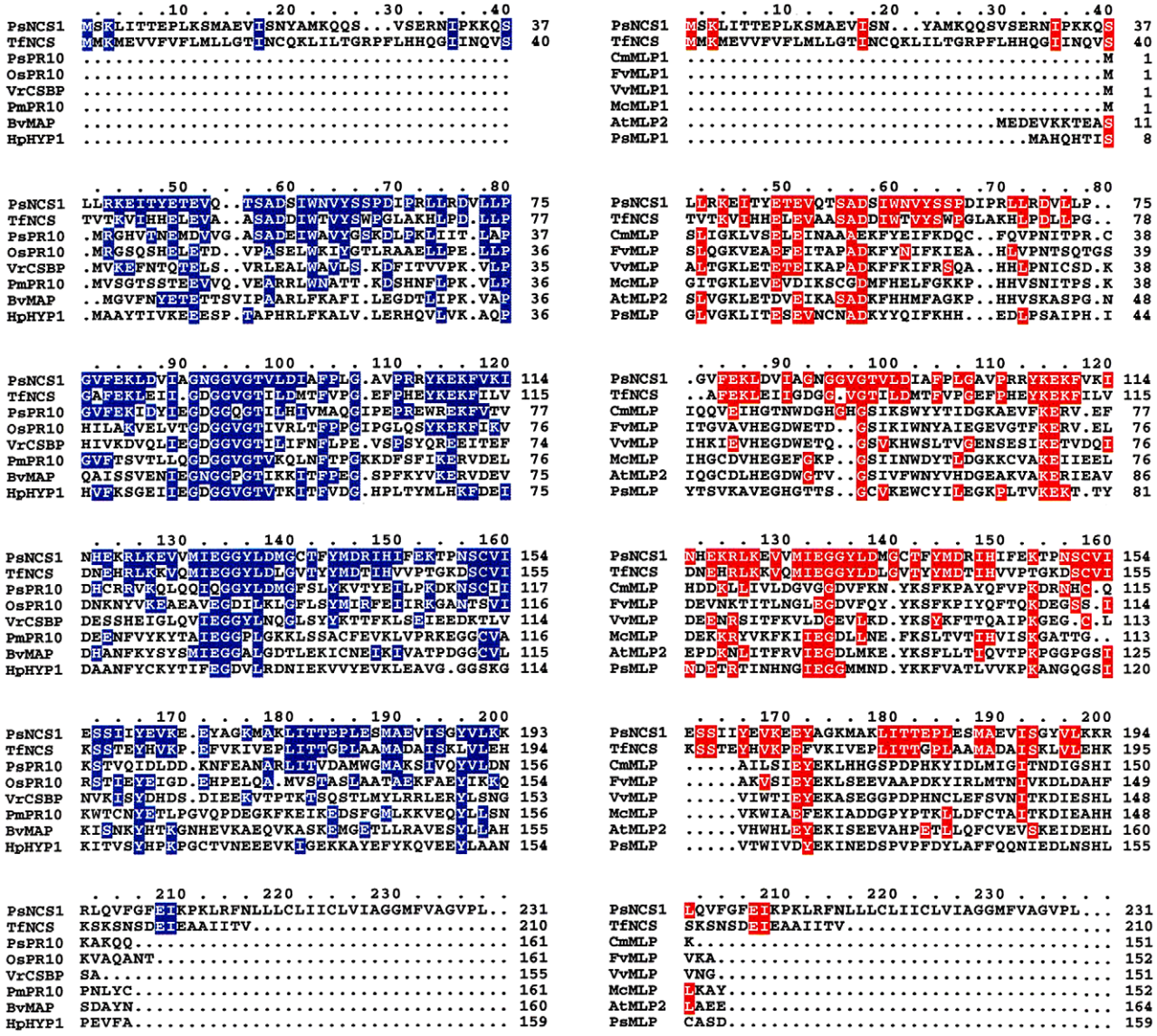

Fig. 4. Alignment (shown in blue) of the deduced amino acid sequence of *Papaver somniferum* norcoclaurine synthase 1 (PsNCS1) and *Thalictrum flavum* NCS (TfNCS) with Bet v 1 major allergen (MAP) and pathogenesis-related (PR)10 proteins from various plant species. Alignment (shown in red) of the deduced amino acid sequence of PsNCS1 and TfNCS with ripening related (RRP) and major latex (MLP) protein homologs from various plant species. Coloured boxes indicate residues identical to those in PsNCS1. Abbreviations and accessions are listed in Section 5.

with the NCS and BBE trees, despite the high support for clades representing functionally similar genes. *P. somniferum* Ps6OMT and Ps4′OMT have higher nucleotide substitution rates compared with functionally identical genes in *T. flavum* and *C. japonica*. For example, the distance between Ps6OMT and Cj6OMT is over 1.2 substitutions per site, while the difference between Cj6OMT and Tf6OMT is less than 0.2 substitutions per site. This differential rate of evolution could be explained by the closer relationship between *C. japonica* and *T. flavum* (Ranunculaceae) compared with *P. somniferum* (Papaveraceae), although the same pattern does not occur with BBE genes (Fig. 6).

3. Discussion

A combination of biochemical and molecular phylogenetic approaches were used to conduct an empirical investigation into the evolution of BIA biosynthesis in plants. NCS activity was assayed in 90 diverse plant species, including a large number of basal angiosperms, as a key biochemical marker due to the central role of this enzyme as the first committed step in BIA metabolism. The occurrence of NCS activity was compared to the distribution of BIA accumulation superimposed on a molecular phylogeny that included the genera represented in the NCS screen. We also identify and functionally characterize

Table 2
Percent amino acid sequence identity among selected norcoclaurine synthases (NCS), pathogenesis relation (PR)10 proteins, major allergen proteins (MAP), cytokinin-specific binding proteins (CSBP), major latex proteins (MLP), and ripening-related proteins (RRP)

Protein	TfNCS	PsPR10	OsPR10	VrCSBP	HpHYP1	BvMAP	PmPR10	CmMLP	FvMLP	VvMLP	McMLP	AtMLP2	PsMLP
PsNCS1	39	30	22	16	17	16	17	13	11	13	11	12	15
TfNCS	–	31	28	19	19	20	20	14	14	18	16	17	16
PsPR10	–	–	30	24	18	19	23	16	18	18	18	17	15
OsPR10	–	–	–	22	23	22	26	18	17	23	19	17	15
VrCSBP	–	–	–	–	20	26	29	15	19	20	16	19	18
HpHYP1	–	–	–	–	–	36	34	13	21	18	14	16	21
BvMAP	–	–	–	–	–	–	37	15	17	23	15	19	20
PmPR10	–	–	–	–	–	–	–	21	20	26	20	20	18
CmRRP	–	–	–	–	–	–	–	–	47	40	31	31	27
FvRRP	–	–	–	–	–	–	–	–	–	43	32	41	30
VvRRP	–	–	–	–	–	–	–	–	–	–	39	37	30
McMLP	–	–	–	–	–	–	–	–	–	–	–	42	27
AtMLP2	–	–	–	–	–	–	–	–	–	–	–	–	32

Abbreviations and accessions are listed in Section 5.

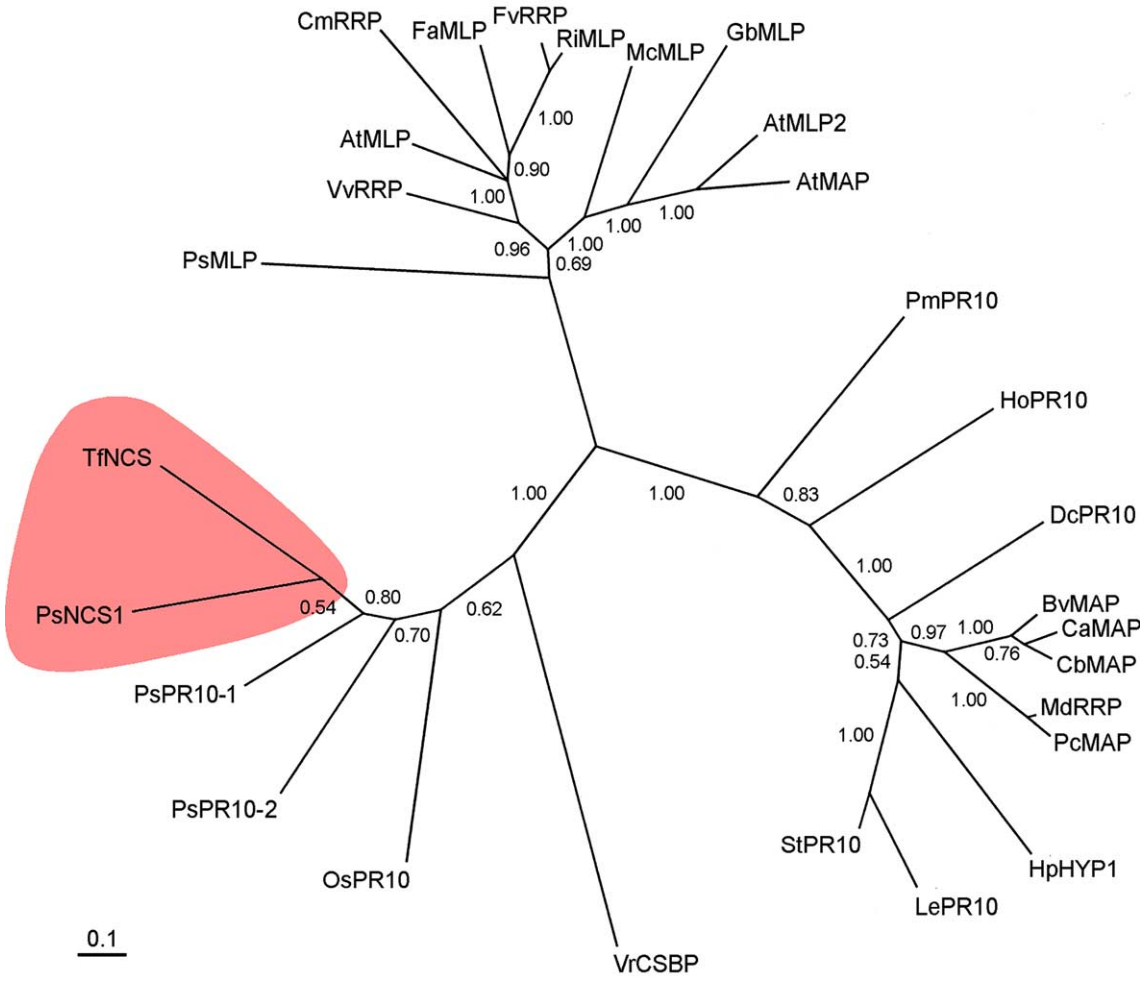

Fig. 5. Consensus Bayesian tree derived from proteins homologous to nococlaurine synthase (NCS). The tree was constructed as an unrooted 50% majority rule consensus tree from 1×10^6 Markov chain Monte Carlo runs using a GTR + Γ + I model in MrBayes on a codon alignment of NCS-related ORFs. Gene products functionally characterized as NCS are shaded in red. The scale represents 0.1 nucleotide substitutions per site and internal labels give posterior probabilities for each clade. Abbreviations and accessions are listed in Section 5.

the gene encoding NCS from opium poppy. Six homologous, recombinant proteins were assayed for NCS activity to further investigate the origin of this unique activity. Molecular phylogenetic analysis of BIA biosynthetic enzymes was performed to gain additional insight into the evolution of the pathway. Our results support a

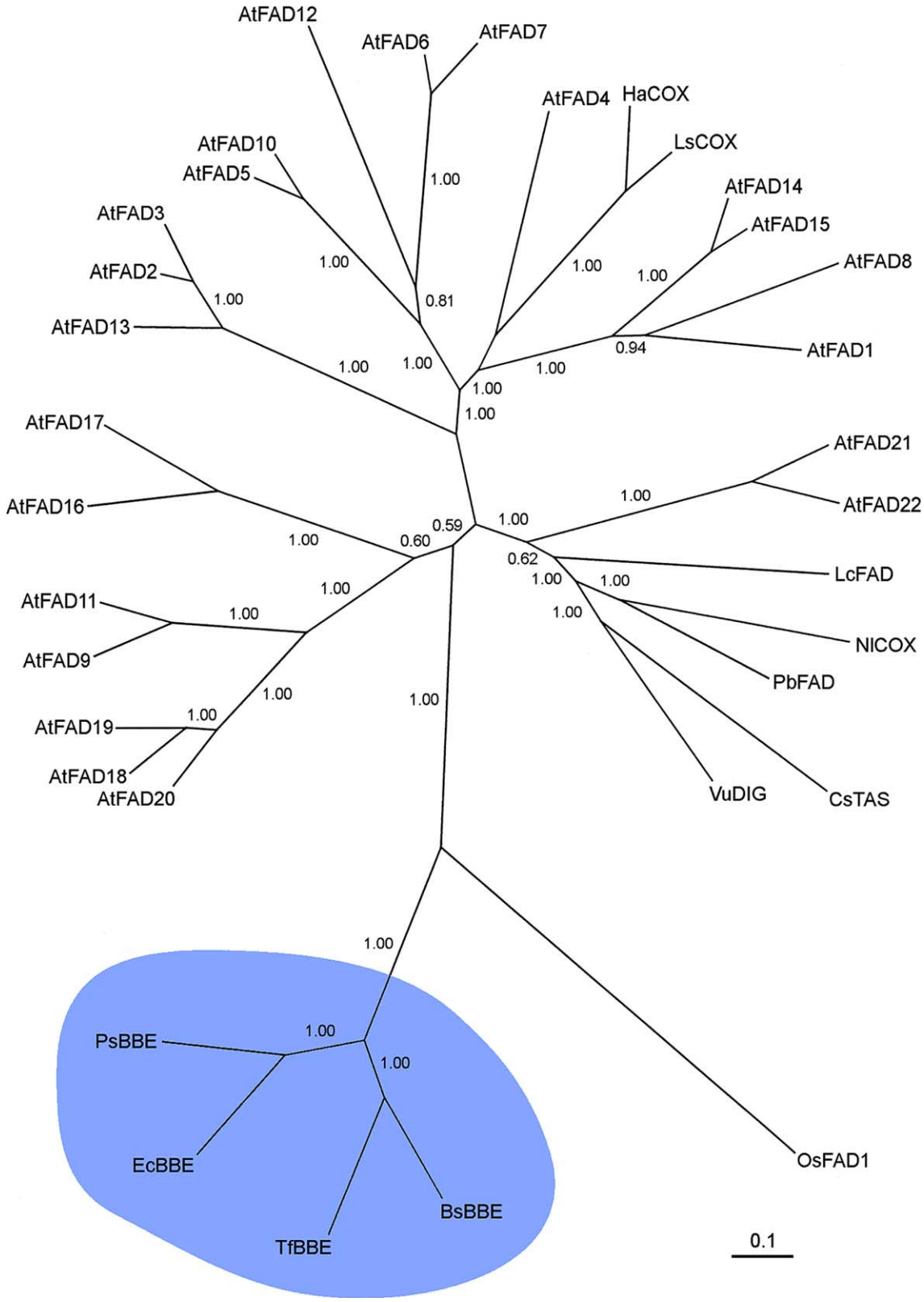

Fig. 6. Consensus Bayesian tree derived from putative or functionally characterized FAD-dependent proteins homologous to berberine bridge enzyme (BBE). The tree was constructed as an unrooted 50% majority rule consensus tree from 1×10^6 Markov chain Monte Carlo runs using a GTR + Γ + I model in MrBayes on a codon alignment of BBE-related ORFs. Gene products functionally characterized as BBE are shaded in blue. The scale represents 0.1 nucleotide substitutions per site and internal labels give posterior probabilities for each clade. Abbreviations and accessions are listed in Section 5.

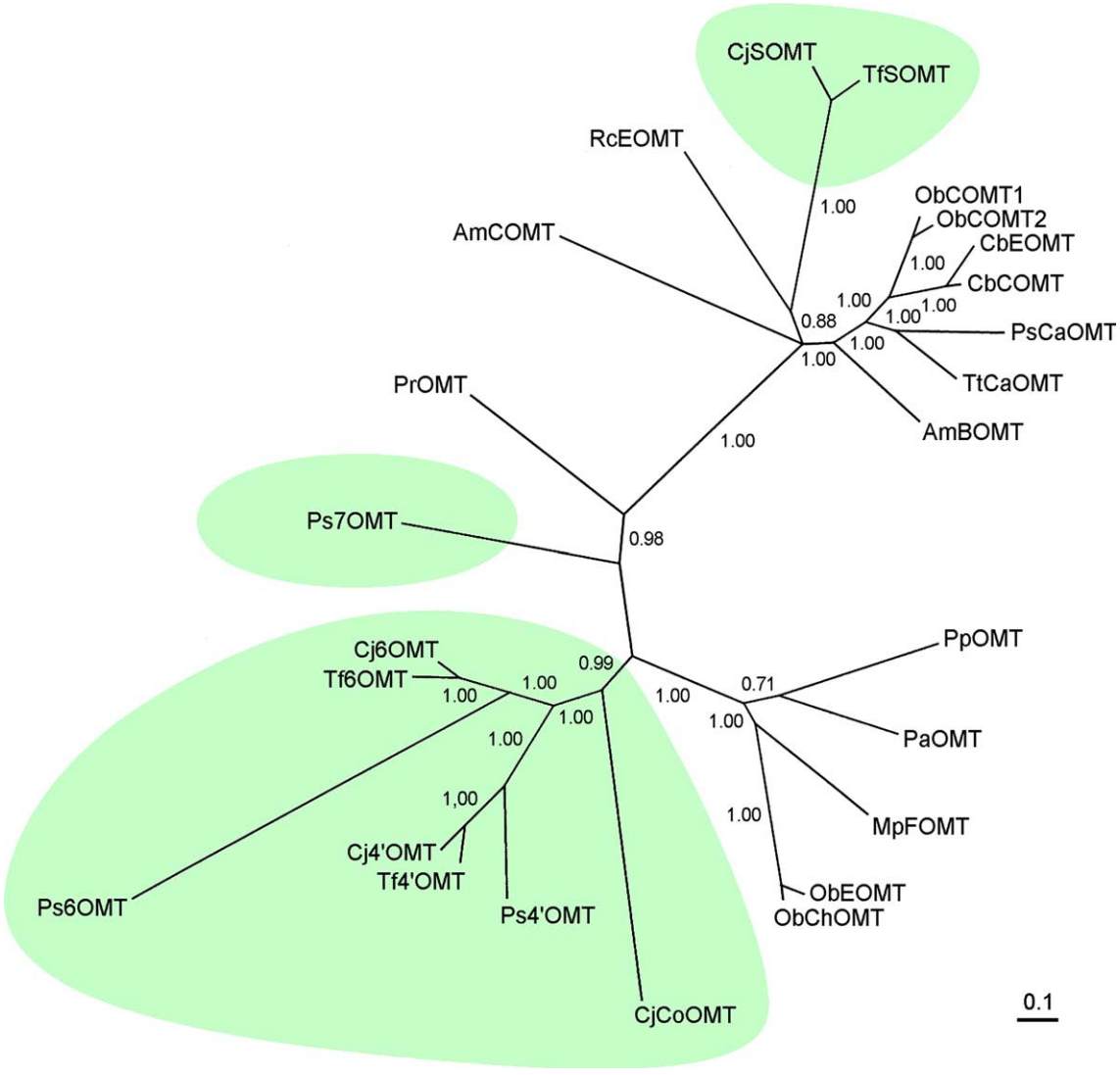

Fig. 7. Consensus Bayesian tree derived from selected *O*-methyltransferases (OMTs). The tree was constructed as an unrooted 50% majority rule consensus tree from 1×10^6 Markov chain Monte Carlo runs using a GTR + Γ + I model in MrBayes on a codon alignment of OMT-related ORFs. Gene products functionally characterized as OMT are shaded in green. The scale represents 0.1 nucleotide substitutions per site and internal labels give posterior probabilities for each clade. Abbreviations and accessions are listed in Section 5.

model for the monophyletic evolution of BIA biosynthesis in angiosperms.

3.1. Identification and functional characterization of opium poppy NCS

With the exception of NCS, enzymes involved in BIA biosynthesis generally exhibit strong sequence similarity across taxa (Samanani et al., 2005). Despite the relatively low homology with *T. flavum* TfNCS (Table 2), the most similar predicted proteins detected in *P. somniferum*, PsNCS1 and PsNCS2, each catalyze the efficient formation of (*S*)-norcoclaurine from dopamine and 4-HPAA (Fig. 3). The limited amino acid sequence identity is particularly evident in the first 20 N-terminal residues, which were predicted to represent a signal peptide in TfNCS. However, a hydrophobic N-terminal signal peptide is not predicted in either PsNCS isoform, although the presence or absence of a signal peptide must be tested empirically.

Significant overall homology is shared between NCS and members of the MAP (Bet v 1) and PR10 protein families, and with the RRP and MLP families (Fig. 4), suggesting an ancient recruitment of NCS from a common ancestor as the primary evolutionary event that allowed certain plant taxa to produce BIAs. The occurrence of such an event is supported by molecular phylogenetic analysis (Fig. 5). We assayed recombinant forms of the six MAP (Bet v 1) and PR10 proteins shown in Fig. 4, and demonstrated their complete inability to catalyze the formation of norcoclaurine (Fig. 3). This suggests that NCS is a unique enzyme that is solely

responsible for the biosynthesis of norcoclaurine in angiosperms. However, few sequence motifs unique to NCS, compared with MAP and PR10 protein homologs, that could confer this enzymatic activity can be readily discerned (Fig. 4). The availability of somewhat divergent NCS sequences from two different plant families (i.e. Ranunculaceae and Papaveraceae), combined with the phylogenetic analysis presented in this study, create the potential to identify the structural basis for the unique function of this enzyme.

NCS is not the only member of the MAP (Bet v 1)/PR10 protein family that functions in a secondary metabolic pathway. HYP1 (Figs. 4 and 5) catalyzes the formation of the bioactive compound hypericin in St-John's wort (*Hypericum perforatum*), which involves the condensation of emodine and emodine anthrone followed by dehydration and successive phenolic oxidations to yield naphthodianthrone (Bais et al., 2003). Catalytic activity has also been reported for Bet v 1 proteins, which were shown to exhibit a ribonuclease function in vitro (Swoboda et al., 1996). Another NCS homolog, VrCSBP from *V. radiata*, was identified based on an ability to selectively bind cytokinins (Fujimoto et al., 1998). Overall, it is possible that the apparent propensity for proteins of this class to interact specifically with secondary metabolites and other small molecules might have eventually led to the evolution of several unique catalytic functions, such as NCS activity. The phylogenetic relationship between NCS and various MLP/RRP homologs (Figs. 4 and 5) is intriguing since BIAs accumulate in the cytoplasm of laticifers, which are specialized cells in opium poppy that contain MLPs as the most abundant cytosolic proteins (Facchini, 2001; Nessler et al., 1990). The divergence of the MLP/RRP and MAP (Bet v 1)/PR10 protein families represents an ancient evolutionary event that preceded that emergence of NCS (Fig. 5).

3.2. Evolution of the BIA biosynthetic pathway

The reliability of NCS activity as a biochemical marker for BIA biosynthesis was evident from the general correlation of NCS activity with the reported occurrence of BIAs in most genera tested (Fig. 2; Table 1). The few taxa in which NCS activity was not detected despite a documented accumulation of BIAs might be expected due to the spatio-temporal regulation of alkaloid biosynthesis (Facchini, 2001), which could lead to a restricted level of NCS activity the samples tested. *Cephalotaxus fortunei* is an interesting example in which NCS activity was not expected. The *Cephalotaxus* genus is well-known for the production of homoerythrina alkaloids arising from a 1-phenethyl-1,2,3,4-tetrahydroisoquinoline derivative via oxidative phenolic coupling of tyrosine and phenylalanine (Parry and Schwab, 1975; Parry et al., 1980). The possibility that dopamine could be the ultimate precursor to these alkaloids has never been investigated. However, dopamine incorporation would conserve the ^{14}C-labelling pattern reported by Parry et al. (1980) and explain the origin of the phenolic oxygens, which could account for the NCS-like activity detected in *C. fortunei*.

The widespread and somewhat sporadic distribution of BIAs throughout the eudicots and the concomitant occurrence of NCS activity support a model for the monophyletic evolution of the BIA biosynthetic pathway. The ability of angiosperms to produce BIAs appears to have originated just before the emergence of the eudicots. This is supported by the apparent lack of BIAs in taxa that diverged before the eudicots (i.e. the basal angiosperms, the monocots and the commelinids), by the high levels of NCS activity detected in several members of the Chloranthaceae, and by the accumulation of aporphine and protoberberine alkaloids in the magnoliids. However, there are some indications that NCS may have arisen as early as the emergence of the basal angiosperms since low levels of NCS activity were detected in members of the Austrobaileyales. Moreover, the simple BIA coclaurine has been reported from *Nymphaea stellata* (Mukherjee et al., 1986), and *Lysichiton camtschatcense* reportedly produces aporphinoids (Katsui and Sato, 1966). The plants are members of the Nymphaceae and Araceae, both of which represent basal angiosperm families. However, BIA production by these or other closely related species has not been demonstrated.

The diverse distribution of BIAs is comparable to patterns observed for other secondary metabolites, such as the tropane and quinolizidine alkaloids. Wink (2003) argues that these phylogenetic patterns result from differential gene expression, and that the biosynthetic capacity to produce a certain secondary metabolite is switched off, but not lost. A similar explanation might apply to BIA biosynthesis, where plant families devoid of BIAs occur phylogenetically between taxa that accumulate these products. The evolution of the genetic blueprint for BIA biosynthesis appear to have been monophyletic, but mutations in biosynthetic genes, regulatory genes, or genes encoding transport proteins led to the inactivation of the pathway in several taxa. Such functional changes might have arisen as the result of altered cell type-specific gene expression patterns or the disruption of protein–protein interactions (de Meaux et al., 2005; Bird et al., 2003; Samanani et al., 2005; Moll et al., 2002).

The ecophysiological functions of specific BIAs, generated by selective environmental pressures, and/or the emergence of specialized physiological mechanisms might have contributed to the maintenance or reactivation of the pathway in certain angiosperm taxa. For example, the evolution of laticifers as an internal sequestration site for BIAs in the Papaveraceae might be responsible for accommodating (indirectly) the biosynthesis and ultimate

storage of copious cytotoxic alkaloids (Facchini, 2001). The absence of laticifers in the Ranunculaceae and Berberidaceae would limit storage options and might contribute a sporadic loss of the BIA pathway in certain genera (Fig. 2).

The detection of abundant NCS activity in *Vancouveria hexandra* suggests that a disruption in the regulation of the pathway, or another biosynthetic enzyme, has led to the inactivation of BIA biosynthesis in this genus. The ancestral pathway might survive in plants that do not produce BIAs as a latent molecular fingerprint that once defined the biosynthetic machinery of a complex metabolic pathway. The *Arabidopsis thaliana* genome sequence has revealed several predicted proteins with strong homology to BIA biosynthetic enzymes, including BBE (Fig. 6), NCS (Figs. 4 and 5) and (*S*)-coclaurine-*N*-methyltransferase (Facchini et al., 2004).

In contrast to the evolution of BIA biosynthesis, strong evidence supports the polyphyletic origin of pyrrolizidine alkaloid pathways, which also occur in diverse taxa. Molecular phylogenetic and biochemical analyses have demonstrated that the independent recruitment of HSS from DHS, conferring the ability to produce pyrrolizidine alkaloids, has occurred repeatedly in separate angiosperm lineages (Reimann et al., 2004). A polyphyletic origin for BIA biosynthesis is unlikely since the evolution of NCS and BBE, for example, would require the establishment of unique biochemical mechanisms in many distant taxa in which (*S*)-scoulerine-derived compounds are found (e.g. the families Aristolochiaceae, Rutaceae, Cornaceae, Magnoliaceae, and the order Ranunculales). Isolated polyphyletic events cannot be ruled out, but the lack of molecular data makes the interpretation of these cases quite difficult. For example, the protoberberine alkaloid bharatamine in *Alangium* spp. is thought to derive from monoterpenoid rather than benzylisoquinoline precursors (Pakrashi et al., 1983) suggesting an independent origin for BIA metabolism in this genus.

The availability of sequence information for a growing number of enzymes involved in BIA biosynthesis has allowed us to perform basic phylogenetic analyses to investigate their evolutionary origin in the Ranunculales. BBE sequences are clearly of monophyletic origin (Fig. 7), and the topology of this clade correlates with the phylogenetic relationships between the source species. BBE genes appear to share ancestral origin with several genes encoding FAD-dependant proteins. Unfortunately, most of these genes have not been characterized and the distant origin of BBE remains uncertain. It is interesting to note that another enzyme involved in secondary metabolism, tetrahydrocannabinolic acid synthase from *Cannabis sativa*, shares an ancestral origin with BBE.

The scoulerine OMTs of *C. japonica* and *T. flavum* have a monophyletic origin relatively near the catechol and caffeic acid OMT-type sequences, while the *P. somniferum* 7OMT appears to be of unique origin. The remaining cloned OMTs of BIA biosynthesis form a monophyletic clade. The length of the *P. somniferum* 6OMT and 4′OMT branch indicates a particularly high rate of sequence change relative to other 6OMT and 4′OMT sequences, which may be due to positive selection as reported for CbEOMT (Barkman, 2003). Positive selection for *P. somniferum* 6OMT might be explained by the extensive breeding of this plant for morphinan content. Such a hypothesis fits with the lack of change in rate of evolution between taxa in BBE (Fig. 7), as this enzyme is not involved in morphinan biosynthesis.

4. Conclusion

More than 2500 diverse BIAs are largely restricted to the order Ranunculales and eumagnoliids, but also occur in the Rutaceae, Lauraceae, Cornaceae and Nelumbonaceae, and sporadically throughout the order Piperales. Our results support a model for the monophyletic evolution of BIA biosynthesis in angiosperms prior to the divergence of the eudicots. Biochemical and phylogenetic analyses of key biosynthetic enzymes suggest a latent molecular fingerprint for BIA biosynthesis in angiosperms not known to accumulate such alkaloids suggesting the requirement for a highly specialized, yet evolutionarily unstable cellular platform to accommodate, or potentially reactivate the pathway in divergent taxa.

5. Experimental

5.1. Plant material

Young leaves, roots and stems lacking extensive secondary growth were harvested from the Zürich Botanical Garden (Zürich, Switzerland). Samples were immediately frozen until used for norcoclaurine synthase assays.

5.2. Chemicals

[8-^{14}C]Dopamine hydrochloride (2035 MBq mol^{-1}) was purchased from Sigma–Aldrich (St. Louis, MO, USA). 4-HPAA was synthesized according to the method of Hirose et al. (2000). All reagents were passed through a 0.45 μm filter.

5.3. cDNA library preparation and expressed sequence tag database assembly

An opium poppy (*Papaver somniferum* cv Marianne) cDNA library was constructed using mRNA isolated

from cell cultures treated for 10 h with a fungal elicitor (Facchini and Park, 2003). Mass excision of the cDNA library was performed using the ExAssist helper phage (Stratagene, La Jolla, CA, USA) and 13,000 independent clones were randomly selected and sequenced in pBluescript SK- (Stratagene, La Jolla, CA, USA) using the T7 promoter as a primer. The resulting FASTA sequences were organized into a searchable EST database using DNATools 6.0 (http://www.dnatools.dk).

5.4. Nucleic acid isolation

A partial NCS cDNA was identified by screening the opium poppy EST database in silico using the tBLASTn algorithm (Altschul et al., 1990) in DNATools 6.0 and *Thalictrum flavum* NCS (Samanani et al., 2004) as the query. The partial NCS cDNA was used to screen the opium poppy cDNA library by membrane hybridization. Full-length cDNAs representing two NCS isoforms (NCS1 and NCS2) were isolated. Several full-length cDNAs encoding PR10 protein homologs were also identified in the EST collection using the tBLASTn algorithm (Altschul et al., 1990), based on sequence homology to known PR10 proteins (Liu and Ekramoddoullah, 2003), and two (PR10-1 and PR10-2) were selected for further characterization.

5.5. Construction of recombinant protein expression vectors

PCR was used to amplify ORFs encoding selected proteins from cDNA templates using sense primers containing a *Bam*HI or *Hin*dIII restriction site and antisense primers containing a *Xho*I restriction site. The OsJNBa9P12 BAC clone containing the *OsPR10* gene was obtained from CUGI (Clemson University, USA). Amplicons were ligated into pET29b (Novagen, Madison, WI, USA) using engineered restriction sites. Construction of the *T. flavum* NCSΔ19, HpHYP1 and BvMAP (Bet v 1) expression vectors are described in Samanani et al. (2004), Bais et al. (2003), and Hoffmann-Sommergruber et al. (1997), respectively.

5.6. Heterologous expression of recombinant proteins in *E. coli*

Escherichia coli ER2566 pLysS cells (New England Biolabs, Beverly, MA, USA) were transformed with expression vectors, grown at 37 °C in Lauria–Bertani medium supplemented with kanamycin (50 μg/ml) and chloramphenicol (34 μg/ml) to an A_{600} of 0.6. Cultures were subsequently induced with 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4.5 h at 28, 30, or 37 °C and harvested by centrifugation at 5000*g* for 5 min. Recombinant proteins were detected in total bacterial protein extracts with Coomassie brilliant blue following separation by SDS–PAGE.

5.7. Total protein extraction

Plant tissues (500 mg) were ground to a fine powder with a mortar and pestle under liquid nitrogen, and extracted in 1 ml buffer A (100 mM Tris–HCl, pH 7.0; 12 mM 2-mercaptoethanol). The slurries were centrifuged at 20,000*g* for 20 min, and the supernatants were desalted on a Sephadex G-25 column. Bacterial cells were resuspended in buffer A supplemented with phenylmethylsulphonyl fluoride (0.6 mM), and lysed using a French press at 15,000 psi. The lysate was centrifuged at 5000*g* for 15 min and the supernatant concentrated by ultrafiltration. Protein concentration was determined according to the method of Bradford (1976).

5.8. NCS assays

NCS activity was measured by incubating total soluble protein extracts (10 μl for plant tissues, 30 μg for recombinant proteins) in buffer A, containing 1 nmol [8-^{14}C]dopamine and 10 nmol 4-HPAA in a total volume of 30 μl, for 1.5 h at 37 °C. Assays using plant tissue extracts were supplemented with diethyldithiocarbamic acid (10 mM) as an inhibitor of amine oxidase activity (Cogoni et al., 1989). The reactions were applied to a silica gel 60 F_{254} TLC plate (EM Science, Gibbstown, NJ), which was developed in *n*-BuOH:HOAc:H_2O (4:1:5; v/v/v). For assays using plant tissue extracts, the [^{14}C](*S*)-norcoclaurine produced was analyzed using a Bio-Imaging Analyzer (FUJIX BAS 1000, Fuji Film, Tokyo, Japan) and quantified using MacBAS software (Fuji Film, Tokyo, Japan). TLC plates from recombinant protein assays were autoradiographed on Kodak OMAT AR film.

5.9. Phylogenetic analysis

Reconstruction of a phylogenetic tree using genera that included the species assayed for NCS activity was performed using molecular marker sequences obtained from GenBank. A combination of *5.8S*, *18S*, *26S*, *atpB*, *ndhF*, *phyA*, *phyC* and *rbcL* genes was selected to minimize the number of markers required to represent all species tested. Accessions are grouped by sequence: ***5.8S*** – *Chelidonium* (AJ001959) *Papaver* (AF098920) *Stylophorum* (AJ001962) *Bocconia* (AJ001952) *Sarcandra* (AF280408) *Acorus* (AF209786) *Houttuynia* (AF275211) *Nymphaea* (AJ012308) *Neolitsea* (AY265401) *Laurus* (AY265392) *Schisandra* (AF163720) *Ceratophyllum* (AY335974) *Sedum* (AY545717) *Holboellia* (AY029795) *Akebia* (AY029788) *Illicium* (AF163726) *Nelumbo* (AY615194) *Paeonia* (U27687) *Adonis* (AY148280) *Aquilegia* (U75659) *Helleborus* (AY148281) *Nigella* (AB020376)

Knowltonia (AB120213) *Ranunculus* (AY680167) *Actaea* (Z98278) *Caltha* (AJ496614) *Cimicifuga* (Z98299) *Trautvetteria* (AY680202) *Trollius* (AY148266) *Consolida* (AY150244) *Pulsatilla* (AB120215) *Thalictrum* (U75661) *Podophyllum* (AF328964) *Chimonanthus* (AY786094) *Stephania* (AY017404) *Macropiper* (AF275193) *Arabidopsis* (X52320) Pinellia (AF4690040); ***18S*** – *Dicentra* (L37908) *Corydalis* (AF094544) *Ceratophyllum* (L54064) *Nandina* (L37911) *Chloranthus* (AY032650) *Sarcandra* (AF207012) *Chimonanthus* (AF503352) *Calycanthus* (U38318) *Stephania* (AJ604528) *Cocculus* (AF197581) *Eupomatia* (AF469771) *Acorus* (AF197584) *Anemopsis* (AF197576) *Houttuynia* (AY032645) *Saururus* (AY032646) *Nuphar* (AF206972) *Nymphaea* (AF206973) *Buxus* (X16599) *Peperomia* (AF206985) *Laurus* (AF197580) *Schisandra* (AF094561) *Liquidambar* (U42553) *Liriodendron* (AJ235981) *Aristolochia* (AF206855) *Alangium* (AF206843) *Asarum* (D29774) *Sedum* (U42528) *Euptelea* (L75831) *Triglochin* (AF197586) *Holboellia* (L37909) *Akebia* (L37905) *Illicium* (L75832) *Nelumbo* (AF094556) *Platanus* (U42794) *Paeonia* (AF274605) *Grevillea* (AF197577) *Ranunculus* (D29780) *Cimicifuga* (AB029382) *Arabidopsis* (X16077) *Austrobaileya* (AF206858); **26S** – *Adonis* (U52625) *Anemone* (U52618) *Aquilegia* (U52607) *Clematis* (U52623) *Helleborus* (U52634) *Nigella* (U52635) *Pulsatilla* (U52620) *Ranunculus* (AF389269) *Actaea* (U52628) *Caltha* (U52632) *Cimicifuga* (U52629) *Trautvetteria* (U52630) *Trollius* (U52624) *Consolida* (U52626) *Dicentra* (AF389262) *Caulophyllum* (AF389240) *Jeffersonia* (U52604) *Vancouveria* (U52602) *Nandina* (AF389241) *Chloranthus* (AY095457) *Hedyosmum* (AY095461) *Saururus* (AY095468) *Nuphar* (AY292901) *Nymphaea* (AY095465) *Pachysandra* (AF389244) *Peperomia* (AY292895) *Liquidambar* (AF479217) *Hamamelis* (AF036495) *Liriodendron* (AY095464) *Alangium* (AY260009) *Ceratophyllum* (AY292904) *Sedum* (AF274667) *Euptelea* (AF389249) *Triglochin* (AY292915) *Akebia* (AF389253) *Nelumbo* (AF389259) *Platanus* (AF274662) *Paeonia* (AF274660) *Thalictrum* (U52610) *Podophyllum* (U52603) *Arabidopsis* (X52320) *Austrobaileya* (AY292886); ***atpB*** – *Papaver* (U86394) *Dicentra* (AJ235454) *Stylophorum* (U86388) *Corydalis* (AF093372) *Podophyllum* (AF092109) *Caulophyllum* (AF092108) *Chloranthus* (AF092113) *Hedyosmum* (AJ235491) *Sarcandra* (AJ235593) *Calycanthus* (AJ235422) *Eupomatia* (AJ235473) *Acorus* (AJ235381) *Anemopsis* (AF197608) *Houttuynia* (AF528851) *Saururus* (AJ235596) *Dictamnus* (AF066830) *Ruta* (AF035913) *Zanthoxylum* (AF066843) *Melicope* (AF066826) *Nuphar* (AF209640) *Nymphaea* (AJ235544) *Pachysandra* (AF528854) *Buxus* (AF092110) *Peperomia* (AJ235556) *Laurus* (AJ235518) *Schisandra* (AJ235599) *Hamamelis* (AF093380) *Liquidambar* (AF092104) *Liriodendron* (AJ235522) *Aristolochia* (AF092106) *Asarum* (U86383) *Alangium* (AJ235386) *Austrobaileya* (AJ235403) *Butomus* (AY147593) *Ceratophyllum* (AJ235430) *Sedum* (AJ235600) *Tamus* (AF308017) *Euptelea* (AF528850) *Triglochin* (AF197601) *Holboellia* (L37928) *Akebia* (AF209523) *Illicium* (U86385) *Nelumbo* (AF528853) *Platanus* (U86386) *Paeonia* (AF209643) *Grevillea* (AF060434); ***ndhF*** – *Aquilegia* (AF130233) *Clematis* (AY145147) *Epimedium* (AY145163) *Podophyllum* (AY145155) *Caulophyllum* (AY145149) *Jeffersonia* (AY145152) *Vancouveria* (AY145159) *Nandina* (AY145148) *Sarcandra* (AY394745) *Cocculus* (AY145144) *Eupomatia* (AY218175) *Acorus* (AF546992) *Arisaema* (AF546995) *Pachysandra* (AF241594) *Buxus* (AF241604) *Schisandra* (AF238062) *Liriodendron* (AF130230) *Alangium* (AF130221) *Butomus* (AF546997) *Ceratophyllum* (AF130232) *Triglochin* (AF546998) *Akebia* (AY145143) *Platanus* (AY818909) *Paeonia* (AF130223) *Arabidopsis* (NC_000932); ***phyA*** – *Sarcandra* (AF276741) *Calycanthus* (AF190072) *Eupomatia* (AF190082) *Acorus* (AF190060) *Houttuynia* (AF276726) *Saururus* (AF190106) *Pachysandra* (AF276734) *Liquidambar* (AY674458) *Aristolochia* (AF276712) *Akebia* (AF276710) *Nelumbo* (AF190096) *Aquilegia* (AF190066) *Hedyosmum* (AF276722) *Chloranthus* (AF190076) *Nymphaea* (AF190098) *Arabidopsis* (NM_100828); ***phyC*** – *Sarcandra* (AF276742) *Calycanthus* (AF190073) *Eupomatia* (AF190083) *Acorus* (AF190061) *Houttuynia* (AF190088) *Saururus* (AF190107) *Nymphaea* (AF190099) *Pachysandra* (AF276735) *Aristolochia* (AF276713) *Akebia* (AF276711) *Illicium* (AF276729) *Nelumbo* (AF190097) *Aquilegia* (AF190067) *Ranunculus* (AY674461) *Hedyosmum* (AF276723) *Chloranthus* (AF190077) *Arabidopsis* (X17343); ***rbcL*** – *Papaver* (L08764) *Dicentra* (L37917) *Stylophorum* (U86633) *Glaucium* (U86626) *Corydalis* (U86622) *Epimedium* (L75869) *Podophyllum* (AF203488) *Caulophyllum* (L08760) *Jeffersonia* (L75867) *Nandina* (L75843) *Chloranthus* (L12640) *Hedyosmum* (L12649) *Calycanthus* (L12635) *Eupomatia* (L12644) *Cocculus* (L12642) *Peltandra* (AJ005628) *Arisaema* (AJ005629) *Anemopsis* (AF197597) *Houttuynia* (L08762) *Saururus* (L14294) *Dictamnus* (AF066801) *Ruta* (AY128251) *Zanthoxylum* (U39282) *Nuphar* (M77029) *Nymphaea* (M77034) *Buxus* (AF543712) *Pachysandra* (AF203486) *Peperomia* (L12661) *Laurus* (AF193970) *Schisandra* (L12665) *Liquidambar* (AF061997) *Lidriodendron* (L12654) *Aristolochia* (AF543711) *Alangium* (L11209) *Butomus* (AY149345) *Tamus* (AF307474) *Euptelea* (L12645) *Triglochin* (AB088811) *Holboellia* (AF398182) *Akebia* (AF335305) *Illicium* (L12652) *Nelumbo* (AF543715) *Platanus* (L01943) *Grevillea* (AF193973) *Ranunculus* (AY395557) *Caltha* (L02431) *Eschscholzia* (U86625) *Arabidopsis* (U91966).

Alignments were performed using ClustalX (Thompson et al., 1997) and edited manually and with the aid of TuneClustalX (Hall, in press). Optimized alignments for each gene were first run through ModelTest (Posada and Crandall, 1998) and MrModelTest (http://www.ebc.uu.se/systzoo/staff/nylander.html) using PAUP* 4.0 beta10 (Swofford, 2003) to determine the model best suited for analysis in MrBayes (Huelsenbeck and Ronquist, 2001). The aligned genes were then combined into a single matrix of 78 taxa, each with 9412 characters for analysis in MrBayes using a model based on ideas presented for analysis of combined data (Nylander et al., 2004; Huelsenbeck and Ronquist, 2004). Ambiguous alignments and polynucleotide stretches were excluded from the analysis. A GTR + Γ + I 4by4 nucleotide model (Lanave et al., 1984; Tavare, 1986) was used for all genes except *5.8S*, which was analyzed using a GTR + Γ + I 4by4 nucleotide model for loop regions and a GTR + Γ + I doublet nucleotide model (Schoniger and von Haeseler, 1994) for paired stem nucleotides. The matrix was partitioned by gene sequence and substitution rates were allowed to vary using a gamma distribution (Yang, 1993) with a proportion of invariable sites (Reeves, 1992; Steel et al., 2000) across partitions. Rates of substitution were also allowed to vary by sequence position. Analysis was performed in duplicate for 1×10^6 and 2×10^6 MCMC cycles to determine any major alternate topologies in clades with weak support. From these analyses, 50% majority rule consensus trees were viewed using TreeView (Page, 1996). Clades and individual taxa, which repeatedly altered position after separate MCMC runs, were manually adjusted to the position most congruent with known phylogenies (Hoot et al., 1997, 1999; Ro et al., 1997; Davies et al., 2004; Kim et al., 2004). The final tree was rooted with *Nuphar* and *Nymphaea* and displayed unrooted, without lengths.

For phylogenetic analysis of alkaloid biosynthetic enzymes, sequences homologous to available BIA biosynthetic enzymes were selected based on initial neighbor-joining trees produced using ClustalX to avoid repetition of similar sequences. Protein abbreviations and accessions are grouped by sequence: **PR10/MLP** – AtMAP, *Arabidopsis thaliana* Bet v 1 major allergen protein (MAP) homolog (NM_122684); AtMLP, *A. thaliana* major latex protein (MLP) homolog (BT015555); AtMLP2, *A. thaliana* MLP homolog (BT014717); BvMAP, *Betula verrucosa* MAP (Bet v 1; AJ006912); GbMLP, *Gossypium barbadense* MLP homolog (AY271667); McMLP, *Mesembryanthemum crystallinum* MLP homolog (AF054451); FvRRP, *Fragaria vesca* ripening-related protein (RRP) (AJ001449); RiMLP, *Rubus idaeus* MLP homolog (AJ001449); FaMLP, *Ficus awkeotsand* MLP homolog (AF497749); CmRRP, *Cucumis melo* RRP (Z70522); VvRRP, *Vitis vinifera* RRP (AJ237994); PsMLP, *Papaver somniferum* MLP (X54305); CaMAP, *Corylus avellana* MAP homolog (X70997); CbMAP, *Carpinus betulus* MAP homolog (Z80169); MdRRP, *Malus domestica* RRP (L42952); PcMAP, *Pyrus communis* MAP homolog (Pyrc1; AF057030); DcPr10, *Daucus carota* pathogenesis related (PR)10 protein (AB082377); HpHYP1, *Hypericum perforatum* phenolic oxidative coupling protein (HYP1, AY148090); LePR10, *Lycopersicon esculentum* PR10 protein (Y15846); StPR10, *Solanum tuberosum* PR10 protein (M25156); HoPR10, *Hyacinthus orientalis* PR10 protein (AY389712); PsNCS1, *P. somniferum* NCS1 (AY860500); TfNCS, *Thalictrum flavum* NCS (AY376412); OsPR10, *Oryza sativa* PR10 protein homolog (AK110687); VrCSBP, *Vigna radiata* cytokinin-specific binding protein (AB012218); PsPR10-1, *P. somniferum* PR10 protein homolog 1 (AY861682); PsPR10-2, *P. somniferum* PR10 protein homolog 2 (AY861683); PmPR10, *Pinus monticola* PR10 protein homolog 1.1 (AY064193); **FAD-dependent** – AtFAD1, *Arabidopsis thaliana* FAD protein (NM_100078); AtFAD2, *A. thaliana* FAD protein (NM_102403); AtFAD3, *A. thaliana* FAD protein (NM_102405.2); AtFAD4, *A. thaliana* FAD protein (NM_102806.2); AtFAD5, *A. thaliana* FAD protein (NM_102807.2); AtFAD6, *A. thaliana* FAD protein (NM_102808.3); AtFAD7, *A. thaliana* FAD protein (NM_102809); AtFAD8, *A. thaliana* FAD protein (NM_102810); AtFAD9, *A. thaliana* FAD protein (NM_102813); AtFAD10, *A. thaliana* FAD protein (NM_103181); AtFAD11, *A. thaliana* FAD protein (NM_129032.3); AtFAD12, *A. thaliana* FAD protein (NM_129034.2); AtFAD13, *A. thaliana* FAD protein (NM_118198.2); AtFAD14, *A. thaliana* FAD protein (NM_202851); AtFAD15, *A. thaliana* FAD protein (NM_118202); AtFAD16, *A. thaliana* FAD protein (NM_118204); AtFAD17, *A. thaliana* FAD protein (NM_123803); AtFAD18, *A. thaliana* FAD protein (NM_123805); AtFAD19, *A. thaliana* FAD protein (NM_123806); AtFAD20, *A. thaliana* FAD protein (NM_123807); AtFAD21, *A. thaliana* FAD protein (NM_123808); AtFAD22, *A. thaliana* FAD protein (NM_123811); BsBBE, *Berberis stolonifera* BBE (AF049347); CsTAS, *Cannabis sativa* tetrahydrocannabinolic acid synthase (AB057805); EcBBE, *Eschscholzia californica* BBE (S65550); HaCOX, *Helianthus annuus* carbohydrate oxidase (COX) (AF472609); LsCOX, *Lactuca sativa* COX (AF472608); LcFAD, *Lotus corniculatus* FAD protein (AP004983); NlCOX, *Nicotiana langsdorffii* x *Nicotiana sanderae* COX (AF503442; AF503441); OsFAD1, *Oryza sativa* FAD protein (AP004698); PsBBE, *Papaver somniferum* BBE (AF025430); PbFAD, *Populus balsamifera subsp. trichocarpa* FAD protein (AC149424); TfBBE, *Thalictrum flavum* BBE (AAU20769); VuDIG, *Vigna unguiculata* drought induceable gene (AB056448); **OMT** – AmBOMT, *Ammi majus* bergaptol OMT (AY443006); AmCOMT, *A. majus* caffeic acid OMT (AY443008); CbCOMT, *Clarkia breweri* caffeic acid

OMT (AF006009); CbEOMT, *C. breweri* (iso)eugenol OMT (U86760); Cj4′OMT, *Coptis japonica* 3′-hydroxy-*N*-methylcoclaurine 4′-OMT (D29812); Cj6OMT, *C. japonica* norcoclaurine 6-OMT (D29811); CjCoOMT, *C. japonica* columbamine OMT (AB073908); CjSOMT, *C. japonica* scoulerine 9-OMT (D29809); MpFOMT, *Mentha piperita* flavonoid OMT (AY337462); ObChOMT, *Ocimum basilicum* chavicol OMT (AF435007); ObCOMT1, *O. basilicum* caffeic acid OMT1 (AF154918); ObCOMT2, *O. basilicum* caffeic acid OMT2 (AF154917); ObEOMT, *O. basilicum* eugenol OMT (AF435008); Ps4′OMT, *Papaver somniferum* 3′hydroxy-*N*-methylcoclaurine 4′-OMT (AY217334); Ps6OMT, *P. somniferum* norcoclaurine 6-OMT (AY268894); Ps7OMT, *P. somniferum* reticuline 7-OMT (AY268893); PsCaOMT, *P. somniferum* catechol OMT (AY268895); PrOMT, *Pinus radiata* OMT homolog (U70873); PaOMT, *Prunus armeniaca* OMT homolog (U82011); PpOMT, *Pyrus pyrifolia* OMT homolog (AB014456); RcEOMT, *Rosa chinensis* eugenol OMT (AB086103); Tf4′OMT, *Thalictrum flavum* 3′-hydroxy-*N*-methylcoclaurine 4′-OMT (AAU20768); Tf6OMT, *T. flavum* norcoclaurine 6-OMT (AAU20765); TfSOMT, *T. flavum* scoulerine 9-OMT (AAU20770); TtCaOMT, *T. tuberosum* catechol OMT (AF064693).

Caffeic acid, chavicol and (iso)eugenol *O*-methyltransferase sequences from basil (*Ocimum basilicum*) and *Clarkia breweri* were also included in the *O*-methyltransferase phylogenetic analysis for comparison to prior phylogenetic studies. Codon alignments for collections of NCS, BBE and OMT sequence homologs were performed with CodonAlign (www.sinauer.com/hall) analyzed using ModelTest, MrModelTest and PAUP* 4.0 beta10 to determine the model best suited for analysis in MrBayes. Codon alignments were then analyzed in MrBayes in a partitioned GTR + Γ + I with a 4by4 nucleotide model. Sequences were partitioned by codon position and parameters were allowed to vary across these partitions using a gamma distribution with a proportion of invariable sites. Rates of substitution were also allowed to vary by sequence position. The analysis for each sequence set was performed in duplicate using MCMC chains of 1×10^6, 2×10^6 and 3×10^6 cycles to identify any alternate topologies. Finally, 50% majority rule consensus trees were displayed using TreeView. All sequence identity and similarity values were calculated using Ident and Sim (Stothard, 2000).

Acknowledgements

We are grateful to Peter Enz and his staff at the Zürich Botanical Garden for their cooperation and assistance with the collection of plant materials. We also thank Heimo Breiteneder for the Bet v 1a expression construct, Jorge Vivanco for the HYP1 expression construct, Abul Ekramoddoullah for the PmPR10-1 cDNA, and Yasuyuki Fujimoto for the VrCSBP cDNA. P.J.F. holds the Canada Research Chair in Plant Biotechnology. D.K.L. is the recipient of a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship. This work was funded by a grant from the Natural Sciences and Engineering Research Council of Canada to P.J.F.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Anke, S., Niemuller, D., Moll, S., Hansch, R., Ober, D., 2004. Polyphyletic origin of pyrrolizidine alkaloids within the Asteraceae. Evidence from differential tissue expression of homospermine synthase. Plant Physiol. 136, 4037–4047.

Bais, H.P., Vepachedu, R., Lawrence, C.B., Stermitz, F.R., Vivanco, J.M., 2003. Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's Wort (*Hypericum perforatum* L.). J. Biol. Chem. 278, 32413–32422.

Barkman, T.J., 2003. Evidence for positive selection on the floral scent gene isoeugenol-*O*-methyltransferase. Mol. Biol. Evol. 20, 168–172.

Bird, D.A., Franceschi, V.R., Facchini, P.J., 2003. A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell 15, 2626–2635.

Bohlmann, J., Meyer-Gauen, G., Croteau, R., 1998. Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA 95, 4126–4133.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Castedo, L., Tojo, G., 1990. Phenanthrene alkaloids. In: Brossi, A. (Ed.), The Alkaloids, vol. 39. Academic Press, New York, pp. 99–136.

Cogoni, A., Farci, R., Medda, R., Rinaldi, A., Floris, G., 1989. Amine oxidase from *Lathyrus cicera* and *Phaseolus vulgaris*: purification and properties. Prep. Biochem. 19, 95–112.

Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E., Savolainen, V., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. Proc. Natl. Acad. Sci. USA 101, 1904–1909.

de Meaux, J., Goebel, U., Pop, A., Mitchell-Olds, T., 2005. Allele-specific assay reveals functional variation in the *Chalcone Synthase* promoter of *Arabidopsis thaliana* that is compatible with neutral evolution. Plant Cell 17, 676–690.

Facchini, P.J., 2001. Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. Ann. Rev. Plant Physiol. Plant Mol. Biol. 52, 29–66.

Facchini, P.J., Park, S.-U., 2003. Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy. Phytochemistry 64, 177–186.

Facchini, P.J., Bird, D.A., St-Pierre, B., 2004. Can Arabidopsis make complex alkaloids?. Trends Plant Sci. 9, 116–122.

Firn, R.D., Jones, C.G., 2003. Natural products – a simple model to explain chemical diversity. Nat. Prod. Rep. 20, 382–391.

Fujimoto, Y., Nagata, R., Fukasawa, H., Yano, K., Azuma, M., Iida, A., Sugimoto, S., Shudo, K., Hashimoto, Y., 1998. Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*). Eur. J. Biochem. 258, 794–802.

Gang, D.R., Lavid, N., Zubieta, C., Chen, F., Beuerle, T., Lewinsohn, E., Noel, J.P., Pichersky, E., 2002. Characterization of phenylpropene *O*-methyltransferases from sweet basil: facile change of

substrate specificity and convergent evolution within a plant *O*-methyltransferase family. Plant Cell 14, 505–519.
Gozler, B., Lantz, M.S., Shamma, M., 1983. The pavine and isopavine alkaloids. J. Nat. Prod. 46, 293–309.
Hall, B.G., in press. Comparison of the accuracies of several phylogenetic methods using protein and DNA sequences. Mol. Biol. Evol 22, 792–802.
Hirose, T., Sunazuka, T., Zhi-Ming, T., Handa, M., Uchida, R., Shiomi, K., Harigaya, Y., Omura, S., 2000. Total syntheses of kurasoins a and b, novel protein farnesyltransferase inhibitors, and absolute structures of kurasoins a and b. Heterocycles 53, 777–784.
Hocquemiller, R., Oztekin, A., Roblot, F., Hutin, M., Cave, A., 1984. Alkaloids of *Papaver pilosum*. J. Nat. Prod. 47, 342–346.
Hoffmann-Sommergruber, K., Susani, M., Ferreira, F., Jertschin, P., Ahorn, H., Steiner, R., Kraft, D., Scheiner, O., Breiteneder, H., 1997. High-level expression and purification of the major birch pollen allergen, Bet v 1. Protein Exp. Purif. 9, 33–39.
Hoot, S.B., Kadereot, J.W., Blattner, F.R., Jork, K.B., Schwarzbach, Crane, P.R., 1997. Data congruence and phylogeny of the Papaveraceae based on four data sets: *atpB* and *rbcL* sequences, *trnK* restriction sites, and morphological characters. Syst. Bot. 22, 575–590.
Hoot, S., Magallón, S.B., Crane, P.R., 1999. Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcL*, and *18S* nuclear ribosomal DNA sequences. Ann. Miss. Bot. Gar. 86, 1–32.
Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.
Huelsenbeck, J.P., Ronquist, F., 2004. Bayesian analysis of molecular evolution using MrBayes. In: Neilson, R. (Ed.), Statistical Methods in Molecular Evolution. Springer, New York, pp. 183–232.
Katsui, N., Sato, K., 1966. Alkaloids of *Lysichiton camtschatcense* Schott var. Japonicum Makino. Tet. Lett. 50, 6257–6261.
Kim, S.K., Ryu, S.Y., No, J., Choi, S.U., Kim, Y.S., 2001. Cytotoxic alkaloids from *Houttuynia cordata*. Arch. Pharm. Res. 24, 518–521.
Kim, Y.-D., Kim, S.-H., Kim, C.H., Jansen, R.K., 2004. Phylogeny of Berberidaceae based on sequences of the chloroplast gene *ndhF*. Biochem. Syst. Ecol. 32, 291–301.
Kumar, V., Poonam, Prasad, A.K., Parmar, V.S., 2003. Naturally occurring aristolactams, aristolochic acids and dioxoaporphines and their biological activities. Nat. Prod. Rep. 20, 565–583.
Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.
Liu, J.J., Ekramoddoullah, A.K., 2003. Root-specific expression of a western white pine PR10 gene is mediated by different promoter regions in transgenic tobacco. Plant Mol. Biol. 52, 103–120.
Lopes, L.M.X., Humpfer, E., 1997. 8-Benzylberbine and *N*-oxide alkaloids from *Aristolochia gigantea*. Phytochemistry 45, 431–435.
Lorence, A., Nessler, C.L., 2004. Camptothecin, over four decades of surprising findings. Phytochemistry 65, 2735–2749.
Moll, S., Anke, S., Kahmann, U., Hansch, R., Hartmann, T., Ober, D., 2002. Cell-specific expression of homospermidine synthase, the entry enzyme of the pyrrolizidine alkaloid pathway in *Senecio vernalis*, in comparison with its ancestor, deoxyhypusine synthase. Plant Physiol. 130, 45–57.
Mukherjee, K.S., Bhattacharya, P., Mukherjee, R.K., 1986. Chemical examination of *Nymphaea stellata* Willd. J. Ind. Chem. Soc. 63, 538.
Nakajima, K., Oshita, Y., Yamada, Y., Hashimoto, T., 1999. Insight into the molecular evolution of two tropinone reductases. Biosci. Biotechnol. Biochem. 63, 1819–1822.
Neilsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.
Nessler, C.L., Kurz, W.G.W., Pelcher, L.E., 1990. Isolation and analysis of the major latex protein genes of opium poppy. Plant Mol. Biol. 15, 951–953.
Nylander, J.A.A., Ronquist, F., Huesenbeck, J.P., Nieves-Aldrey, J.L., 2004. Beyesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.
Ober, D., Hartmann, T., 2000. Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids. Plant Mol. Biol. 44, 445–450.
Page, R.D.M., 1996. TreeView: An application to display phylogenetic trees on personal computers. Comp. Appl. Biosci. 12, 357–358.
Pakrashi, S.C., Mukhopadhyay, R., Dastidar, P., Bhattacharjya, A., Ali, E., 1983. Bharatamine – a unique protoberberine alkaloid from *Alangium lamarckii* Thw., biogenetically derived from monoterpenoid precursor. Tet. Lett. 24, 291–294.
Parry, R.J., Schwab, J.M., 1975. Biosynthesis of the cephalotaxus alkaloids. I. Novel mode of tyrosine incorporation into cephalotaxine. J. Am. Chem. Soc. 97, 2555–2557.
Parry, R.J., Chang, M.N.T., Schwab, J.M., Foxman, B.M., 1980. Biosynthesis of the cephalotaxus alkaloids. Investigations of the early and late stages of cephalotaxine biosynthesis. J. Am. Chem. Soc. 103, 1099–1111.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Rastrelli, L., Capasso, A., Pizza, C., De Tommasi, N., 1997. New protopine and benzyltetrahydroprotoberberine alkaloids from *Aristolochia constricta* and their activity in isolated guinea pig ileum. J. Nat. Prod. 60, 1065–1069.
Reeves, J.H., 1992. Heterogenetity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA. J. Mol. Evol. 35, 17–31.
Reimann, A., Nurhayati, N., Backenkohler, A., Ober, D., 2004. Repeated evolution of the pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages. Plant Cell 16, 2772–2784.
Ro, K.-E., Keener, C.S., McPheron, B.A., 1997. Molecular phylogenetic study of the Ranunculaceae: utility of the nuclear 26S ribosomal DNA in inferring intrafamilial relationships. Mol. Phylogenet. Evol. 8, 117–127.
Rozwadowska, M.D., 1988. Secoisoquinoline alkaloids. In: Brossi, A. (Ed.), The Alkaloids, vol. 33. Academic Press, New York, pp. 231–306.
Rueffer, M., El-Shagi, H., Nagakura, N., Zenk, M.H., 1981. (*S*)-Norlaudanosoline synthase: The first enzyme in the benzylisoquinoline biosynthetic pathway. FEBS Lett. 129, 5–9.
Samanani, N., Facchini, P.J., 2002. Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy. Planta 213, 898–906.
Samanani, N., Liscombe, D.K., Facchini, P.J., 2004. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. Plant J. 40, 302–313.
Samanani, N., Park, S.U., Facchini, P.J., 2005. Cell type-specific localization of transcripts encoding nine consecutive enzymes involved in protoberberine alkaloid biosynthesis. Plant Cell. 17, 915–926.
Sari, A., Sariyar, G., 1997. Isolation of triniifoline, miltanthaline and some medicinally important alkaloids from *Papaver triniifolium*. Planta Med. 63, 575–576.
Sariyar, G., 2002. Biodiversity in the alkaloids of Turkish *Papaver* species. Pure Appl. Chem. 74, 557–574.
Schoniger, M., von Haeseler, A., 1994. A stochastic model and the evolution of autocorrelated DNA sequences. Mol. Phyl. Evol. 3, 240–247.
Schmeller, T., Latz-Brüning, B., Wink, M., 1997. Biochemical activities of berberine, palmatine and sanguinarine mediating chemical defence against microorganisms and herbivores. Phytochemistry 44, 257–266.

Shamma, M., Rothenberg, A.S., Jayatilake, G.S., Hussain, S.F., 1978. A new group of isoquinoline alkaloids: the secoberbines. Tetrahedron 34, 635–640.

Shulgin, A.T., Perry, W.E., 2002. The Simple Plant Isoquinolines. Transform Press, Berkeley.

Stadler, R., Kutchan, T.M., Löffler, S., Nagakura, N., Cassels, B.K., Zenk, M.H., 1987. Revision of the early steps of reticuline biosynthesis. Tetrahedron Lett. 28, 1251–1254.

Stadler, R., Kutchan, T.M., Zenk, M.H., 1989. Norcoclaurine is the central intermediate in benzylisoquinoline alkaloid biosynthesis. Phytochemistry 28, 1083–1086.

Steel, M., Huson, D., Lockhart, P.J., 2000. Invariable sites models and their use in phylogeny reconstruction. Syst. Biol. 49, 225–232.

Steffens, P., Nagakura, N., Zenk, M.H., 1985. Purification and characterization of the berberine bridge enzyme from *Berberis beaniana* cell cultures. Phytochemistry 24, 2577–2583.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques 28, 1102–1104.

Swoboda, I., Hoffman-Sommergruber, K., Scheiner, O., Heberle-Bors, H., Vincente, O., 1996. Bet v 1 proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity in vitro. Physiol. Plant 93, 433–438.

Swofford, D.L., 2003. PAUP*, Phylogenetic Analysis Using Parsimony, and Other Methods, Version 4. Sinauer Associates, Sunderland, MA.

Tavare, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lect. Math. Life Sci. 17, 57–86.

Theis, N., Lerdau, M., 2003. The evolution of function in secondary metabolites. Int. J. Plant Sci. 164, S93–S102.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Wang, J., Pichersky, E., 1999. Identification of specific residues involved in substrate discrimination in two plant *O*-methyltransferases. Arch. Biochem. Biophys. 368, 172–180.

Wink, M., 2003. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. Phytochemistry 64, 3–19.

Xu, Q., Lin, M., 1999. Benzylisoquinoline alkaloids from *Gnetum parvifolium*. J. Nat. Prod. 62, 1025–1027.

Yang, Z., 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol. Biol. Evol. 10, 1396–1401.