Org. Divers. Evol. 3, 227–237 (2003)

http://www.urbanfischer.de/journals/ode

ORGANISMS DIVERSITY & EVOLUTION

Phylogeny and biogeography of Morinaceae (Dipsacales) based on nuclear and chloroplast DNA sequences

Charles D. Bell*, Michael J. Donoghue

Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, USA

Received 7 January 2003 · Accepted 6 May 2003

Abstract

The Morinaceae (Dipsacales) contains 13 species placed in *Acanthocalyx*, *Cryptothladia* or *Morina*, and is distributed from the mountains of southeastern Europe through the Himalayas to the Qinghai-Tibetan Plateau, mainly in alpine habitats. Sequence data from two chloroplast regions (the *trn*K intron and the *trn*L-F region) and the internal transcribed spacer (ITS) region of nuclear ribosomal DNA were used to infer phylogenetic relationships of Morinaceae and related Dipsacales. Both the nuclear and chloroplast datasets, as well as the combined data, provide strong support for relationships within the Valerina clade, placing Morinaceae as the sister group of a clade containing Valerianaceae and Dipsacaceae plus *Triplostegia*. The Morinaceae, *Acanthocalyx*, *Cryptothladia*, and a clade containing *Morina* and *Cryptothladia*, are all supported as monophyletic. However, *Morina* was found to be paraphyletic in several of our analyses, with *Morina longifolia* more closely related to *Cryptothladia* than to other *Morina* species. There is some evidence that *Morina longifolia* produces cleistogamous flowers, as do *Cryptothladia* species. Dispersal-vicariance analyses support the view that Valerina radiated initially within Asia, with subsequent movement into Europe in Morinaceae, Dipsacaceae, and Valerianaceae, and into the New World in Valerianaceae. For Morinaceae, as for a number of plant groups, the Brahmaputra river drainage marks a significant biogeographic divide, although this has been spanned within *Acanthocalyx* and the *Morina-Cryptothladia* lineage.

Key words: Bayesian methods, character evolution, Dipsacales, Morinaceae, Valerina

Introduction

An excellent taxonomic revision of Morinaceae was published by Cannon & Cannon in 1984. They recognized 13 species extending from southeastern Europe and Israel to the Himalayas of Nepal, Bhutan, and China, and further east and north in the Qinghai-Tibetan Plateau. These plants most often grow at high elevations, sometimes in excess of 4,200 meters, but they occupy a range of habitats, including rock ledges, alpine meadows, dry slopes, the margins of pine forests, and even swamps.

Within Morinaceae, Cannon & Cannon (1984) recognized three genera: *Morina*, *Acanthocalyx*, and *Cryptothladia*. *Acanthocalyx* has basal, opposite leaves, corolla lobes that are more or less equal in size, four fertile stamens, and pollen grains that lack prominent protrusions. In contrast, *Morina* and *Cryptothladia* share whorled leaves, distinctly zygomorphyic, 2-lipped corollas with unequal lobes, two fertile stamens and two reduced sterile stamens, and a unique pollen form with three prominent equatorial protrusions (Blackmore & Cannon 1983, Cannon & Cannon 1984). *Cryptothladia* differs from *Morina* in having 2–4 corolla lobes as opposed to five, and in that the corolla is equal to or, in most cases, shorter than the calyx. *Cryptothladia* flowers appear to be predominantly cleistogamous.

Cannon & Cannon (1984) compared Morinaceae to Valerianaceae and Dipsacaceae, but did not specifically hypothesize relationships among these groups. Mori-

*** Corresponding author:** Charles D. Bell, Department of Ecology and Evolutionary Biology, P.O. Box 208106, Yale University, New Haven, Connecticut 06520, USA; e-mail: charles.bell@yale.edu

naceae have sometimes been united with Dipsacaceae, and in some treatments are even considered to be a tribe within Dipsacaceae (e.g., Cronquist 1988). Such treatments highlight the presence of a distinctive epicalyx surrounding the ovary in these groups plus *Triplostegia* (which has a double epicalyx). Other authors have highlighted similarities shared by Morinaceae, Valerianaceae, and *Triplostegia*, such as the presence of three distinct carpels (two of which abort), and of clearly zygomorphic corollas. Backlund & Bremer (1998) summarized the varying taxonomic treatments of the group.

In recent years a variety of phylogenetic analyses have supported the monophyly of the Morinaceae and clarified its relationships within Dipsacales (Caputo & Cozzolino 1994, Peng et al. 1995, Backlund & Donoghue 1996, Backlund & Bremer 1997, Bell et al. 2001, Donoghue et al. 2001, Pyck 2001, Pyck et al. 2002, Zhang et al. 2003 Donoghue et al. in press). The more recent and most comprehensive of these analyses have agreed that Morinaceae, Valerianaceae, *Triplosetgia*, and Dipsacaceae form a clade (the Valerina clade of Donoghue et al. 2001), and that this is in turn united with the Linnaeeae of the traditional Caprifoliaceae (together forming the Linnina clade of Donoghue et al. 2001). Within Valerina there have been some differences among phylogenetic analyses, although these have been weakly supported. Several analyses combining different data sources have supported the hypothesis that Morinaceae form the sister group of a clade containing Valerianaceae plus a clade containing Dipsacaceae plus *Triplostegia*.

Representatives of the three genera of Morinaceae have been included in several previous phylogenetic studies (Caputo & Cozzolino 1994, Peng et al. 1995, Backlund & Donoghue 1996, Bell et al. 2001, Donoghue et al. 2001, Pyck 2001, Donoghue et al. in press). In all of these cases, *Acanthocalyx* has been found to be the sister group of *Morina* plus *Cryptothladia*. However, only one species has been included from each of these groups, and it has therefore not been possible to test their monophyly.

In the present paper we use DNA sequences from both chloroplast and nuclear genomes to infer the phylogeny of Morinaceae using parsimony, maximum likelihood, and Bayesian methods. The main difference between this study and those conducted previously is that we have been able to include 11 of the 13 recognized species of Morinaceae (collections of *Cryptothladia ludlowii* and *Cryptothladia parviflora* were unavailable). With this greatly improved sample we reconsider the monophyly of Morinaceae and, for the first time, test the monophyly of the three previously recognized genera. Our phylogenetic results clarify the evolution of a variety of morphological characters, as well as the historical biogeography of the group.

Materials and methods

Plant material

Sequence data from 24 taxa were collected from herbarium specimens, silica-preserved and fresh plant material (Table 1). Eleven of the thirteen species of Morinaceae were sampled, along with 10 species of Valerianaceae and Dipsacaceae. Based on previous analyses (Judd et al. 1994, Pyck et al. 1999, Pyck & Smets 2000, Bell et al. 2001, Donoghue et al. 2001, Donoghue et al. in press), three species of Linnaeeae were included for rooting purposes.

DNA extraction, amplification and sequencing

Total DNA was extracted using the CTAB methods of Doyle & Doyle (1987) or Qiagen DNeasy plant mini extraction kits (Qiagen). Double stranded copies of all regions were amplified using standard PCR reactions. Reactions were initially denatured at 94 °C for 3 min, then 35 cycles were carried out of 94 °C for 1.5 min, 46–56 °C for 2 min, and 72 °C for 3 min. Amplification products were purified using QIAquick PCR purification kits (Qiagen).

The *trn*K intron and a portion of the 5′ end of the *mat*K gene were amplified with primers trnk-11 and matK510R (Young et al. 1999). Two newly designed internal primers – matK-seq1 (5′-ATTWGAATGATTKCCCGATCTA-3′) and matK-seq2 (5′-TAGATCGGGMAATCATTCWAAT-3′) – were used along with the amplification primers for sequencing. All amplification and sequencing of the *trn*L-F IGS region and *trn*L intron and of the 3′ *trn*L exon utilized the universal primers c, d, e, and f of Taberlet et al. (1991). The ITS region was amplified and sequenced using the primers ITS2, ITS3B, ITS4 (White et al. 1990), and ITS-I (Urbatsch et al. 2000).

Sequencing was conducted using dye-terminator cycle sequencing (protocols specified by the manufacturer), and then visualized on an ABI 377 automated DNA sequencer or an MJ Research BaseStation51 DNA fragment analyzer.

Sequence alignment and phylogenetic analyses

Sequence fragments were edited using the computer package Sequencher 3.0 (Gene Codes Corp., Ann Arbor) to build contig sequences. All sequences were then aligned visually with the aid of MacClade version 4.0 (Maddison and Maddison 2000). In the primary, 'full' dataset containing all of the taxa, several regions of the ITS sequences could not be confidently aligned and were therefore omitted (see details below). However, in a 'reduced' ITS dataset containing only species of Morinaceae, regions of ambiguity in the large dataset could readily be aligned. Our alignments, along with the published tree, are available in TreeBASE (www.treebase. org/treebase) and from the first author. Although we only present one tree retrieved from the maximum-likelihood and Bayesian analyses of the combined data, trees from separate ITS and cpDNA analyses are available in TreeBASE.

All parsimony and likelihood analyses were performed using the computer software PAUP* vers. 4.0b10 for UNIX (Swofford 2002). For the full dataset, maximum parsimony

Table 1. Taxa included in phylogenetic analyses, voucher information for new sequences, and GenBank accession numbers.

Species	Voucher	ITS	*trn*L-F	*trn*K-intron/*mat*K
Morinaceae				
Acanthocalyx alba (Hand.-Mazz.) M. Cannon	Boufford et al., 28401 (A)	AY236183	AF446973	AY290027
Acanthocalyx delavayi (Franchet) M. Cannon	Boufford et al., 28374 (A)	AY290014	AY290003	AY290028
Acanthocalyx nepalensis (D. Don) M. Cannon	Boufford et al., 28955 (A)	AY290015	AY290004	AY290029
Cryptothladia chinesis (Pai) M. Cannon	Boufford et al., 27846 (A)	AY236184	AF446974	AY290025
Cryptothladia chlorantha (Diels) M. Cannon	Ho et al., 1095 (A)	AY290012	AY290001	AY290024
Cryptothladia kokonorica (Hao) M. Cannon	Boufford et al., 26762 (A)	AY290011	AY290000	AY290023
Cryptothladia polyphylla (Wallich ex DC) M. Cannon	Takayama et al., 923915 (A)	AY290013	AY290002	AY290026
Morina coulteriana Royle	Grady et al., 6553 (A)	AY290009	AY289998	AY290021
Morina kokanica Regel	Bobrov, 4167 (A)	AY290019	n/a	n/a
Morina longifolia Wallich ex DC.	Cult. Bergius Bot. Gard., Sweden; Eriksson, s.n., 2 Nov. 1999 (SBT)	AY236185	AF446975	AY290020
Morina persica L.	Coode & Jones, 2323 (A)	AY290010	AY289999	AY290022
Other Dipsacales				
Dipsacus pilosa L.	Cult. National Botanic Garden of Belgium, F. Billiet/H1304	AY290016	AY290005	AY290030
Fedia cornucopiae (L.) Gaertner	Cult. Bergius Bot. Gard., Sweden; Eriksson, s.n., 2 Nov. 1999 (SBT)	AY236193	AF446983	AY290037
Nardostachys jatamansi (D. Don) DC.	Boufford et al., 28099 (A)	AY236190	AF446980	AY290038
Patrinia triloba Miq.	Cult. Bergius Bot. Gard., Sweden; Eriksson, 807 (SBT)	AY236191	AF446981	AY290039
Plectritis macroceras Torr. & Gray	Patterson, 2000 (SFSU)	AY236195	AF446985	AY290036
Scabiosa atropurpurea L.	Cult. National Botanic Garden of Belgium, F. Billiet/H1446	AY290017	AY290006	AY290033
Scabiosa columbaria L	Cult. National Botanic Garden of Belgium, F. Billiet/H434	AY236188	AF446078	AY290032
Succisa pratenis Moench	Cult. National Botanic Garden of Belgium, F. Billiet/V263	AY290018	AY290007	AY290033
Triplostegia glandulifera Wallich ex DC.	Boufford et al., 28440 (A)	AY236189	AF446979	AY290034
Valeriana urticifolia HBK	C. Bell EC024 (Yale)	AY290019	AY290008	AY290035
Linnaeeae (outgroups)				
Abelia x grandiflora Rehder	Cult. Wake Forest Univ., NC., 1999; Donoghue, voucher lacking	AY236179	AF446969	AY290041
Dipelta yunnanensis Franchet	Boufford et al., 29340 (A)	AY236180	AF446970	AY290042
Linnaea borealis L.	Door County, WI; Donoghue, 1990, voucher lacking	AY236181	AF446971	AY290040

searches were conducted using heuristic search methods with tree-bisection-reconnection (TBR) branch swapping, zero-length branches collapsed, and all characters weighted equally. Analyses were repeated 100 times with the RANDOM ADDITION option. Equally parsimonious trees were summarized using strict consensus. Confidence in clades was assessed using the bootstrap (Felsenstein 1985); 300 replicates were performed with heuristic search settings identical to those of the original search. For the reduced datasets (11 species for ITS; 10 species for the chloroplast and combined analyses, where we were unable to amplify the chloroplast regions for *Morina kokanica*, possibly due to degraded DNA sampled from an herbarium specimen) we performed exhaustive searches; branch-and-bound searches were used in bootstrap tests.

A series of likelihood ratio tests were performed – on a variety of tree topologies, e.g. neighbor-joining (NJ) and maximum parsimony (MP) – using the computer program PORN* (Bell 2001) to determine which model of sequence evolution best fit the data. Maximum likelihood tree searches were carried out using the appropriate model, with parameters for each search simultaneously estimated via maximum likelihood. Heuristic search methods were used with TBR branch swapping and collapse of zero-length branches. Again, analyses were repeated 100 times with RANDOM ADDITION. Bootstrap tests were performed using 1,000 replicates with nearest-

neighbor-interchange (NNI) branch swapping. Parameters for bootstrap tests were fixed to values estimate on the maximum likelihood tree.

Bayesian analyses were carried out using Metropolis-coupled Markov chain Monte Carlo (MC^3) methods as implemented in the computer software MrBayes ver. 2.01 (Huelsenbeck & Ronquist 2001). The chain was run for 40–50 million generations, initiated with a random starting tree. The chain was sampled every 1,000 generations, for a total of 40,000–50,000 trees sampled from the posterior distribution of trees, which were used to calculate the posterior probabilities of clades. Sample points collected prior to stationary were eliminated from the sample (i.e. 'burn-in'). To detect the number of generations that constituted the size of the burn-in, we plotted –*ln*L values versus generation and checked for when the chain was stationary. The incremental heating scheme for each MC^3 search used the default settings in MrBayes (three 'heated' chains, and one 'cold' chain). We also used the defaults for the priors on the rate matrix (0–100), branch lengths (0–10), gamma shape parameter (0–10), and the proportion of invariable sites (0–1). A Dirichlet distribution was used for the base frequency parameters, and an uninformative prior was used for the tree topology.

A parametric bootstrap approach was used to assess potential conflict between the nuclear and chloroplast datasets (Goldman 1993, Huelsenbeck & Rannala 1997, Swofford et al. 1996). The maximum likelihood tree for the combined dataset (T_A) was used as a null hypothesis (assuming that this topology underlies both data partitions). We then simulated 100 datasets based on this null hypothesis and the maximum likelihood estimates of the parameters of the best-fitting model derived for both of the original datasets (nuclear and chloroplast), using the computer software Seq-Gen ver. 1.2.5 (Rambaut & Grassly 1997). The –*ln*L scores were then calculated with PAUP*, using T_A and re-estimating the free parameters of the model for each of the 100 datasets. Next, tree searches were conducted, maximizing model parameters for each of the simulated datasets. We then used the difference of the –*ln*L scores as an estimate of the distribution under the null hypothesis and compared it to the difference in –*ln*L values from the original data. To further investigate potential conflict between the data partitions, we also performed the homogeneity partition test (i.e. ILD test) as implemented in PAUP*, using 1,000 replicate searches.

A parametric bootstrap approach was also used to test for the monophyly of specific clades (e.g. *Morina*). For each test we used MacClade to build constraint trees, enforcing the monophyly of each clade of interest. We then used maximum likelihood to search for trees under the given constraint. The maximum likelihood tree from the constraint search (T_c) was used as a null hypothesis. We then simulated 100 datasets based on this null hypothesis and the maximum likelihood estimates of the parameters of the best-fitting model estimated from the original dataset using Seq-Gen. The –*ln*L scores were then calculated with PAUP*, using T_c and re-estimating the free parameters of the model for each of the 100 datasets. Next, tree searches were conducted, maximizing model parameters for each of the simulated datasets. We then used the difference of the –*ln*L scores as an estimate of the distribution under the null hypothesis and compared it to the difference in –*ln*L values from the original data.

Dispersal-vicariance analysis

We used the computer software DIVA ver. 1.1 (Ronquist 1996, 1997) to estimate ancestral areas within Valeriana and Morinaceae. DIVA reconstructs ancestral areas by minimizing the number of dispersal and extinction events needed to explain a given distribution pattern. In analyses of the entire Valerina clade we coded areas as "Asia", "Europe", "North America" and "South America". In analyses focused on Morinaceae, "Asia" was subdivided into "western Himalayas" and "eastern Himalayas", marked by the Brahmaputra river drainage.

Results

ITS sequences

Several ITS regions were excluded from the large dataset because they could not be confidently aligned. These corresponded to positions 199–261 and 459–485, for a total of 88 excluded characters. The full matrix therefore consisted of 628 aligned nucleotides (ranging in length from 472 to 576), of which 299 were variable and 182 parsimony informative. Parsimony analyses resulted in 6 trees of 658 steps (CI = 0.73, 0.62 excluding invariant characters; RI = 0.72). Our maximum likelihood search resulted in a single tree with a –*ln*L value of 3,393.64. This was identical to one of the topologies found under parsimony. Bayesian analysis recovered a tree topology with a –*ln*L value of 3,393.64, again identical to the ML and parsimony trees.

The reduced matrix consisted of 653 aligned nucleotides (ranging in length from 532 to 643), of which 132 were variable and 95 parsimony informative. Parsimony analyses resulted in 3 trees of 162 steps (CI = 0.90, 0.87 excluding invariant characters; RI = 0.91). Our maximum likelihood search resulted in a single tree with a –*ln*L value of 1,747.60. This was one step longer than the topologies found under parsimony (163 steps). Bayesian analysis recovered a tree topology with a –*ln*L value of 1,747.60, again identical to the ML tree.

Chloroplast sequences

The large chloroplast data matrix consisted of 2,326 aligned base pairs (ranging in length from 1,887 to 2,177) from the *trn*K-intron, a portion of the *mat*K coding region, the *trn*L intron, and the intergenic spacer (IGS) regions of *trn*L-F. Of these, 900 were variable and 444 parsimony informative. The *trn*K intron varied in length from 798 bp (in *Fedia cornucopiae*) to 852 bp (in *Acanthocalyx nepalensis*), the 5′ end of the *mat*K coding region ranged from 878 to 964 bp, and the *trn*L-F region from 326 to 366 bp. A number of potentially informative gaps were present in the chloroplast regions (both coding and non-coding), but were not included in the phylo-

genetic analyses. Maximum parsimony analyses resulted in a single tree of 1,449 steps (CI = 0.75, 0.62 excluding invariant characters; RI = 0.73). The maximum likelihood and Bayesian searches both resulted in a single tree with a –*ln*L value of 11,285.82 (1,449 steps under parsimony).

The reduced chloroplast data matrix consisted of 2,181 aligned base pairs from the *trn*K-intron, a portion of the *mat*K coding region, the *trn*L intron and the intergenic spacer (IGS) regions of *trn*L-F. Of these, 271 were variable and 100 parsimony informative. Maximum parsimony analyses resulted in 3 most parsimonious trees of 342 steps (CI = 0.85, 0.68 excluding invariant characters; RI = 0.76). The maximum likelihood and Bayesian searches both resulted in a single tree with a –*ln*L value of 5,027.91 (342 steps under parsimony).

Parametric bootstrapping and ILD tests

Parametric bootstrap analyses suggested that there was not significant incongruence (at the 5% level) between the chloroplast dataset and the result obtained from the combined analysis (i.e. the null hypothesis T_A), at a level greater than expected from sampling error for both the full and reduced datasets (P = 0.39 and P = 0.32, respectively). The results for the ITS dataset were also not significant (P = 0.12 and P = 0.09, respectively). Results from the ILD test also suggested no significant (P = 0.32) conflict between the nuclear and chloroplast datasets.

Combined analyses

The combined nuclear and chloroplast dataset consisted of 2,934 aligned base pairs (ranging in length from 2,463 to 2,728), of which 1,199 were variable and 626 parsimony informative. Maximum parsimony analyses resulted in 2 trees of 2,128 steps (CI = 0.72, 0.61 excluding invariant characters; RI = 0.72). Our maximum likelihood and Bayesian searches both resulted in a single tree with a –*ln*L value of 15,621.13 (2,128 steps under parsimony). Of the two parsimony trees, ML favored the one in which Valerianaceae is sister to Dipsacaceae plus *Triplostegia* over the one in which Morinaceae is sister to Dipsacaeae. This is shown in Figure 1.

The reduced nuclear and chloroplast dataset consisted of 2,835 aligned base pairs, of which 403 were variable and 195 parsimony informative. Maximum parsimony analyses resulted in 2 trees of 498 steps (CI = 0.86, 0.76 excluding invariant characters; RI = 0.82). Our maximum likelihood and Bayesian searches both resulted in a single tree with a –*ln*L value of 6,932.45 (500 steps under parsimony).

Summary statistics

Information on the full and reduced ITS, chloroplast, and combined datasets is summarized in Table 2. Support values for the major clades in the resulting trees are presented in Table 3. All analyses recovered trees in which *Acanthocalyx*, *Cryptothladia*, *Morina* plus *Cryptothladia*, Morinaceae, Dipsacaceae, and Valerianaceae were monophyletic, with moderate to very strong support (Table 3). ITS, chloroplast, and combined analyses also supported grouping *Triplostegia* with Dipsacaceae rather than with Valerianaceae. Likewise, all analyses (except the full ITS and combined parsimony searches) supported a sister group relationship between Valerianaceae and Dipsacaceae plus *Triplostegia*. These parsimony searches also find optimal trees that place Morinaceae sister to Dipsacaceae plus *Triplostegia*. In all cases clade support increased in the combined analyses as compared to the separate analyses. *Morina* was not well supported as monophyletic in any of the full analyses nor in the separate reduced analyses; it is very weakly supported only in the ML and Bayesian analyses of the reduced dataset (54% and 0.64, respectively). Parametric bootstrap tests, however, could not reject the monophyly of *Morina* for the nuclear, chloroplast, or combined datasets (P = 0.39, P = 0.32, P = 0.16).

Dispersal-vicariance analysis

DIVA reconstructions support an initial diversification of the Valerina clade in Asia, with subsequent movement into Europe within each of the three major clades – Morinaceae, Dipsacaceae, and Valerianaceae. Within Valerianaceae there are additional shifts into the New World on as many as four separate occasions (Bell & Donoghue, unpublished data). For Morinaceae, DIVA reconstructions inferred a widespread ancestral area at the base of the clade, followed by an initial splitting event marked by the Brahmaputra river. Within the *Morina* plus *Cryptothladia* clade there were two inferred dispersal events and two vicariance events. An additional dispersal and vicariance event was inferred for *Acanthocalyx* (Fig. 2).

Discussion

Results from the separate ITS and cpDNA analyses were basically the same as those found in the combined analysis. Differences seen in topologies across analyses and datasets were never supported by bootstrap values greater than 50% or posterior probabilities greater than 0.80 (see trees in TreeBASE).

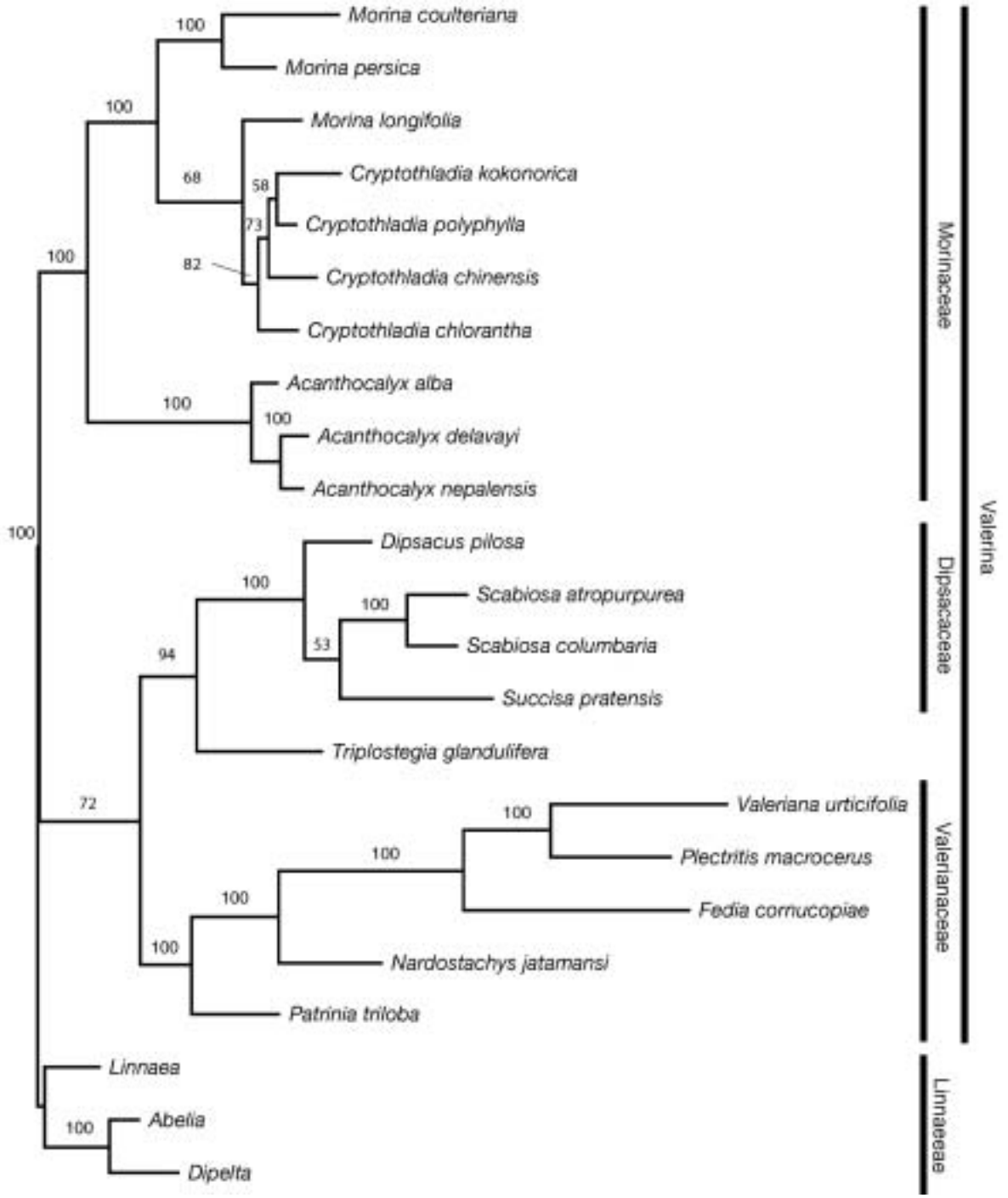

Fig. 1. Maximum likelihood tree (–*ln*L = 15621.1364) of the Linnina clade from the full combined dataset. Numbers above branches represent nonparametric bootstrap proportions (greater than 50%) from maximum likelihood searches.

Phylogeny and character evolution

Maximum likelihood analysis of the combined data provides strong support for relationships within the Valerina clade. Dipsacaceae plus *Triplostegia* form a clade that is sister to Valerianaceae, and Morinaceae is sister to this entire clade. These results are entirely consistent with the analysis of Bell et al. (2001) based on five chloroplast markers. Of special interest is the well-supported connection between *Triplostegia* and Dipsacaceae, in

Table 2. Summary statistics. Summary of maximum parsimony, maximum likelihood, and Bayesian analyses. CI = consistency index, Γ = among-site rate variation modeled to fit a discrete gamma distribution, GTR = general time reversible, I = invariable sites, ML = maximum likelihood, MP = most parsimonious, RI = retention index.

	Chloroplast	ITS	Combined
Full dataset			
Aligned sequence length	2326	608	2934
No. of variable characters	900	299	1199
No. of parsimony-informative characters	444	182	626
No. of MP trees	1	6	2
Length of MP trees	1449	658	2128
CI (all characters)	0.7522	0.6854	0.7270
CI (excluding invariant characters)	0.6173	0.5957	0.6050
RI	0.7274	0.7210	0.7180
–*ln*L of ML tree	11286.0918	3393.6442	15621.1367
Model of sequence evolution	GTR + Γ	GTR + Γ	GTR + Γ
No. of steps on MP tree	1449	659	2128
–*ln*L of Bayesian tree	11286.0918	3393.6442	15621.1367
No. of steps on MP tree	1449	659	2128
Reduced dataset			
Aligned sequence length	2181	653	2835
No. of variable characters	271	132	397
No. of parsimony-informative characters	100	95	196
No. of MP trees	3	3	2
Length of MP trees	342	162	498
CI (all characters)	0.8480	0.9010	0.8635
CI (excluding invariant characters)	0.6790	0.8710	0.7631
RI	0.7559	0.9083	0.8191
–*ln*L of ML tree	5027.9138	1747.6010	6932.4525
Model of sequence evolution	GTR + Γ	GTR + Γ	GTR + Γ
No. of steps on MP tree	342	163	500
–*ln*L of Bayesian tree	5027.9138	1747.6010	6932.4525
No. of steps on MP tree	342	163	500

Table 3. Support values. Bootstrap support greater than 50% for maximum parsimony and maximum likelihood, and posterior probabilities for Bayesian analyses for major clades (MP/ML/BA). * = bootstrap value less than 50%.

	ITS	Chloroplast	Combined
Full dataset			
Acanthocalyx	100/100/1.0	100/100/1.0	100/100/1.0
Cryptothladia	52/*/0.72	63/63/0.78	93/82/0.95
Morina	*/*/*	*/*/*	*/*/*
Morina + Cryptothladia	100/100/1.0	90/93/1.0	96/100/1.0
Morinaceae	98/97/1.0	100/100/1.0	100/100/1.0
Dipsacaceae	100/100/1.0	100/100/1.0	100/100/1.0
Dipsacaceae + *Triplostegia*	73/75/0.80	79/71/0.92	100/94/0.98
Valerianaceae	77/60/0.78	99/100/1.0	100/100/1.0
Valerianaceae, Dipsacaceae + *Triplostegia*	96/98/1.0	98/99/1.0	100/100/1.0
Reduced dataset			
Acanthocalyx	100/100/1.0	100/100/1.0	100/100/1.0
Cryptothladia	75/*/0.82	71/67/0.90	99/98/1.0
Morina	*/*/*	*/*/*	*/54/0.64
Morina + Cryptothladia	100/100/1.0	100/100/1.0	100/100/1.0

contrast to several previous analyses in which *Triplostegia* was instead united with Valerianaceae (Backlund & Donoghue 1996, Backlund & Bremer 1997, Pyck 2001). These relationships are consistent with the interpretation of the epicalyx as having been derived from the fusion of the "supernumerary" inflorescence bracts present in the related Linnaeeae (Manchester & Donoghue 1995, Roels & Smets 1996, Donoghue et al. in press). Under this interpretation supernumerary bracts are a synapomorphy for the Linnina clade. However, as emphasized by Donoghue et al. (in press), this pattern of relationships necessitates at least two originations of the epicalyx – in Morinaceae with more than 8 ribs (often 12 ribs), and again in Dipsacaceae plus *Triplostegia* with 8 ribs – or a single origin at the base of Valerina followed by loss in Valerianaceae. The later interpretation would be favored if the wing-like structures subtending the ovary in fruits of *Patrinia* are interpreted as evidence of an epicalyx at the base of Valerianaceae (Hofmann & Gottmann 1990).

We find strong support for the monophyly of Morinaceae in all analyses, regardless of analytical method. This is consistent with previous analyses based on morphological characters (Caputo & Cozzolino 1994, Peng et al. 1995, Backlund & Donoghue 1996) as well as a variety of molecular datasets (Backlund & Donoghue 1996, Backlund & Bremer 1997, Pyck et al. 1999, Pyck & Smets 2000, Donoghue et al. 2001, Bell et al. 2001, Donoghue et al. in press). Diagnostic morphological characters for Morinaceae include the production of verticillate inflorescences, an epicalyx with 12 nerves, and a distinctly zygomorphic calyx.

For the first time we are able to evaluate the monophyly of the

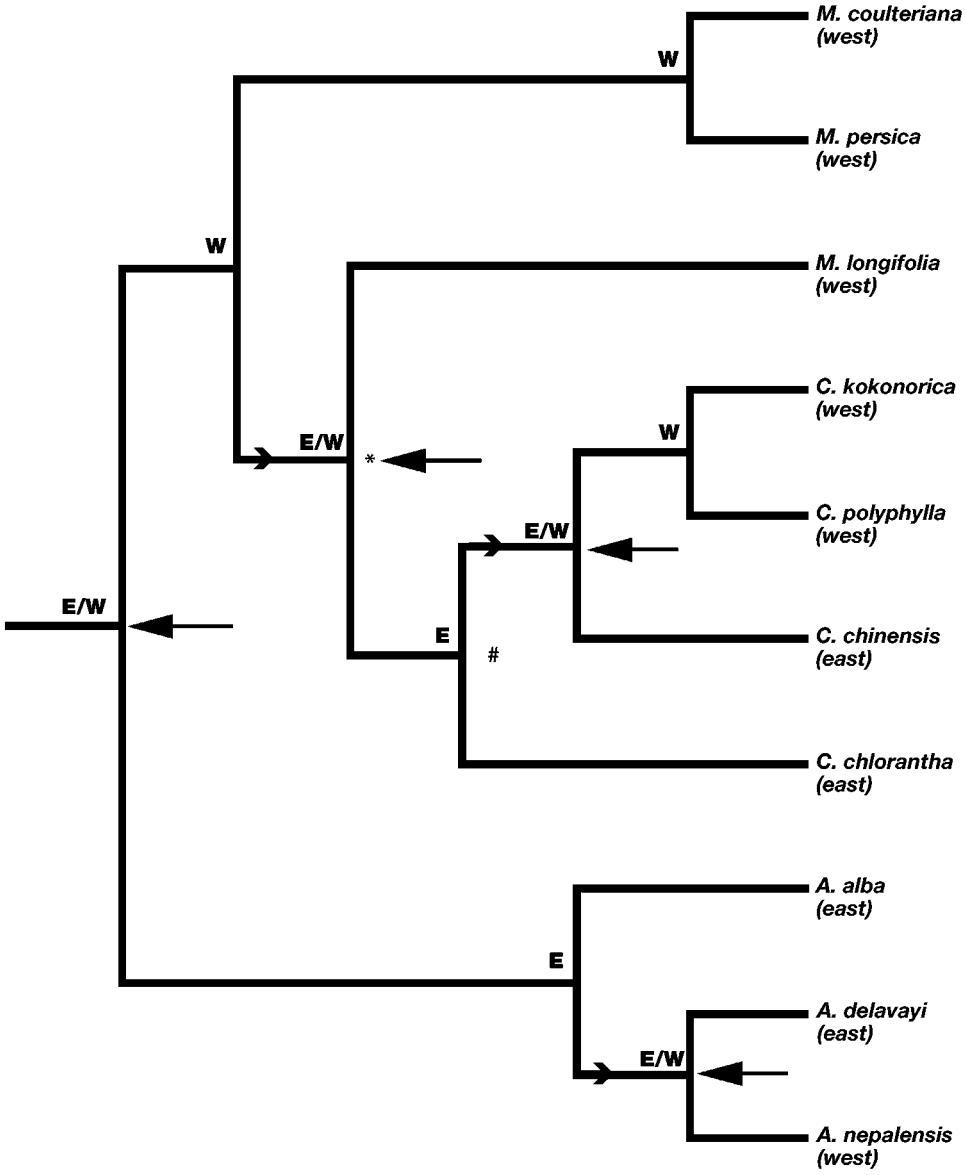

Fig. 2. One of two possible ancestral area reconstructions for Morinaceae obtained from DIVA (Ronquist 1996). E = east of Brahmaputra; W = west of Brahmaputra: E/W = east and west. Arrowheads along branches indicate inferred dispersal events; arrows pointing at nodes indicate inferred vicariance events. * = alternatively reconstructed as W; # = alternatively reconstructed as E/W.

three previously recognized genera and their presumed relationships to one another. We found strong support in all analyses for a basal split separating *Acanthocalyx* from a clade containing *Morina* plus *Cryptothladia* (bootstrap value ≥ 90% and clade probability values = 1.0 for all analyses, see Table 3). *Acanthocalyx* appears to have retained ancestral characteristics, while *Morina* and *Cryptothladia* are linked by a variety of clearly derived morphological features. These include whorled leaves, spike-like inflorescences (with well-separated verticels), two fertile stamens and two staminodia, lobed or separate nectaries, and pollen grains with prominent equatorial protrusions (Cannon & Cannon 1984).

Within *Acanthocalyx* all of our datasets strongly unite *A. deleveya* and *A. nepalensis* as the sister group of *A. alba*. This is consistent with the distribution of corolla color; *A. deleveya* and *A. nepalensis* have pink corollas, whereas *A. alba* has white or cream-colored corollas and sometimes puce. *Acanthocalyx alba* also has a somewhat shorter calyx and a narrower, more curved corolla tube than the other two species (Cannon & Cannon 1984).

Within the *Morina-Cryptothladia* clade we found moderate to strong support for the monophyly of *Cryptothladia* (depending on data partition, see Table 3). This is consistent with a number of apparently derived mor-

phological traits (Cannon & Cannon 1984). Specifically, *Cryptothladia* species have shortened corolla tubes (as long as or shorter than the calyx) and a reduction from five to four (or sometimes two) corolla lobes, apparently through fusion of the two upper lobes. Reduction in the length of the corolla tube correlates with a difference in the position of the sterile stamens. In *Morina* the staminodia are attached in the upper portion of the corolla tube slightly below the two fertile stamens. In contrast, in *Cryptothladia* the two sterile stamens are located at the very base of the tube (or 1/3 of the way up the corolla tube in *C. chinensis*), each stamen associated with a lobed nectary (Cannon & Cannon 1984). This suggests the possibility that corolla tube reduction occurred primarily through retardation of growth in the abaxial "stamen-corolla" portion of the tube (Leins 2000). As noted by Cannon & Cannon (1984), reduction in flower size in *Cryptothladia* may be related to cleistogamy. Relationships among the species of *Cryptothladia* are poorly supported in our analyses. However, combined analyses (Fig. 1) found moderate support (73%) for placing *C. chlorantha* as sister to a clade containing *C. kokonorica*, *C. polyphylla*, and *C. chinensis*.

Our most surprising result concerns *Morina*, which in some of our analyses appears to be paraphyletic, with *Cryptothladia* nested within it. In particular, our data provide some evidence that *Morina longifolia* is more closely related to *Cryptothladia* than it is to the well-supported clade containing *M. coulteriana* and *M. persica*. In the ITS dataset, where we were able to include *M. kokanica*, this species appeared as sister to the *M. coulteriana-M. persica* clade. One possible apomorphy of this clade concerns the pollen grains: the presence of parallel bands of verrucae between the protrusions. There are no obvious characters uniting *M. longifolia* with *Cryptothladia*, but corolla size is of interest in this regard. *Morina kokanica* and *M. persica* have exceptionally long corolla tubes (35–45 mm), followed by *M. coulteriana* (25–30 mm). In general, it appears that the corolla tube (20–30 mm) as well as the corolla lobes are smallest in *M. longifolia*, and in this respect this species more closely approaches the drastically reduced corollas seen in *Cryptothladia* (generally included within the calyx, which varies from 6 to 8 mm in length) (Cannon & Cannon 1984). Of great interest is the observation by Cannon & Cannon (1984: p. 31) that unopened flowers of *M. longifolia* "showed pollen already adhering to the stigma." They interpreted this as possibly indicating cleistogamy in this species, "paralleling the probable cleistogamy in *Cryptothladia*." Our result, uniting *M. longifolia* with *Cryptothladia*, suggests the possibility that *M. longifolia* represents a step in the direction of the well-developed cleistogamy found in *Cryptothladia*. However, it should be emphasized that support for the *M. longifolia-Cryptothladia* connection is weak, and we cannot yet reject the monophyly of *Morina* with confidence.

Biogeography

DIVA analyses imply that the Valerina clade originated in Asia. All basal lineages of the three major clades – Morinaceae, Valerianaceae, Dipsacaceae – have their present distributions restricted to Asia. Valerianaceae and Dipsacaceae have since achieved a widespread distribution across temperate Asia, Europe, Mediterranean Africa, and, in the case of Valerianaceae, North and South America. Morinaceae, in contrast, has remained restricted to the higher elevations of the Himalayas and adjacent areas. The one exception is the widespread *Morina persica*, which apparently spread to the west, reaching Israel and the mountains of Turkey and Greece.

Within Morinaceae it is noteworthy that individual species are primarily restricted to one side or the other of the Brahmaputra river drainage (Cannon & Cannon 1984: figs 3, 5, 8) – either to the east (in eastern Tibet, Yunnan, Sichuan, and Qinghai, China) or to the west (in Bhutan, Nepal, and westward to the Middle East). Specifically, *Acanthocalyx alba*, *A. delevayi*, *Cryptothladia chlorantha* and *C. chinensis* occur primarily to the east of the Brahmaputra, while *A. nepalensis*, *C. kokonorica*, *C. polyphylla* and all *Morina* species basically occupy the western region. This particular east-west split is seen in other Himalayan taxa. In *Roscoea* (Zingiberaceae), for example, one clade occurs to the east and another to the west of the so-called "Brahmaputra gap" (Ngamriabsakul et al. 2000).

In DIVA analyses of the Morinaceae several results are obtained, depending, for example, on whether *M. longifolia* is considered to be the sister group of *Cryptothladia* (Fig. 2). However, in all cases at least three episodes of dispersion are inferred across the Brahmaputra gap: one within *Acanthocalyx*, perhaps from east to west; one within *Cryptothladia*, again perhaps from east to west; and one nearer the base of Morinaceae, perhaps from west to east. However, in the absence of information on the timing of the relevant speciation events it is unclear which, if any, of them correspond to the split seen in *Roscoea*. The distributional gap across the Brahmaputra River may be a more general pattern, warranting further attention in other groups.

Conclusions

It appears from our analyses that Morinaceae is monophyletic and sister to a clade containing Valerianaceae, Dipsacaceae and *Triplostegia*. This result is consistent with those based on five chloroplast markers (Bell et al. 2001) and on a subset of these data analyzed by Zhang

et al. (2003). These relationships suggest that the epicalyx either evolved at the base of Valerina (via the fusion of supernumerary bract as found in the Linnaeeae) and was lost in Valerianaceae (with the exception of *Patrinia*), or that it was gained twice, in Morinaceae and in Dipsacaceae plus *Triplostegia*. Additional data, especially from other nuclear markers, would be helpful in clarifying these relationships.

Within Morinaceae, *Acanthocalyx* is strongly supported as sister to a clade containing species of *Morina* and *Cryptothladia*. This supports a reduction of two of the four stamens to staminodia, and the acquisition of a unique pollen morphology. Relationships within *Acanthocalyx* are strongly supported in all analyses. However, relationships within the *Morina-Cryptothladia* clade are not yet entirely well resolved. Although we find a paraphyletic *Morina* in several of our analyses, parametric bootstrap tests could not reject its monophyly. More data are needed to resolve this issue with confidence. In addition, including the two remaining species of *Cryptothladia* (*C. ludlowii* and *C. parviflora*) in future studies would help in understanding the evolution and biogeography of Morinaceae.

Acknowledgments

We thank Torsten Eriksson and the National Botanic Garden of Belgium for providing plant materials, and Nancy Pyck for sharing unpublished information concerning the phylogeny of Dipsacales and Morinaceae. We also benefited from discussions of this work with members of the Donoghue Lab group at Yale. This study was supported by a NSF Biotic Surveys and Inventories grant to collect plants in the Hengduan Mountain region of China (DEB-9705795), and a Doctoral Dissertation Improvement award (DEB-0104968).

References

Backlund, A. A. & Bremer, K. (1997): Phylogeny of the Asteridae s. str. based on *rbc*L sequences, with particular reference to the Dipsacales. Pl. Syst. Evol. 207: 225–254.

Backlund, A. A. & Bremer, K. (1998): To be or not to be – principles of classification and monotypic plant families. Taxon 47: 391–400.

Backlund, A. A. & Donoghue, M. J. (1996): Morphology and phylogeny of the order Dipsacales. In: Backlund, A. A. Phylogeny of the Dipsacales. Doct. diss., Dept Syst. Bot., Uppsala Univ., Sweden.

Bell, C. D. (2001): PORN*: a hierarchical likelihood ratio calculator for LINUX. http://www.phylodiversity.net/pornstar. Yale University, New Haven, Connecticut.

Bell, C. D., Edwards, E. J., Kim, S.-T. & Donoghue, M. J. (2001): Dipsacales phylogeny based on chloroplast DNA sequences. Harvard Pap. Bot. 6: 481–499.

Blackmore, S. & Cannon, M. J. (1983): Palynology and systematics of Morinaceae. Rev. Paleobot. Palynol. 40: 207–226.

Cannon, M. J. & Cannon, J. F. M. (1984): A revision of the Morinaceae (Magnoliophyta-Dipsacales). Bull. Brit. Mus. (Nat. Hist.) Bot. 12: 1–35.

Caputo, P. & Cozzolino, S. (1994): A cladistic analysis of Dipsacaceae (Dipsacales). Pl. Syst. Evol. 189: 41–61.

Cronquist, A. (1988): The Evolution and Classification of Flowering Plants. New York Botanical Garden, Bronx, New York.

Donoghue, M. J., Bell, C. D. & Winkworth, R. C. (in press): The evolution of reproductive characters in Dipsacales. Int. J. Plant Sci.

Donoghue, M. J., Eriksson, T., Reeves, P. A. & Olmstead, R. G. (2001): Phylogeny and phylogenetic taxonomy of Dipsacales, with special reference to *Sinadoxa* and *Tetradoxa* (Adoxaceae). Harvard Pap. Bot. 6: 459–479.

Doyle, J. J. & Doyle, J. L. (1987): Isolation of plant DNA from fresh tissue. Focus 12: 13–15.

Felsenstein, J. (1985): Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Goldman, N. (1993): Statistical tests of models of DNA substitution. J. Mol. Evol. 36: 182–198.

Hofmann, U. & Gottmann, J. (1990): *Morina* L. und *Triplostegia* Wall. Ex DC im Vergleich mit Valerianaceae und Dipsacaceae. Bot. Jahrb. Syst. 111: 499–553.

Huelsenbeck, J. P. & Rannala, B. (1997): Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276: 227–232.

Huelsenbeck, J. P & Ronquist, F. (2001): MrBayes: Bayesian inference of phylogeny. Bioinformatics 17: 754–755.

Judd, W. S., Sanders, R. W. & Donoghue, M. J. (1994): Angiosperm family pairs – preliminary phylogenetic analyses. Harvard Pap. Bot. 5: 1–51.

Leins, P. (2000): Blüte und Frucht. 390 pp. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.

Maddison, W. P. & Maddison, D. R. (2000): MacClade 4: Interactive Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.

Manchester, S. R. & Donoghue, M. J. (1995): Winged fruits of Linnaeeae (Caprifoliaceae) in the Tertiary of western North America: *Diplodipelta* gen. nov. Int. J. Plant Sci. 156: 709–722.

Ngamriabsakul, C., Newman, M. F. & Cronk, Q. C. B. (2000): Phylogeny and disjunction in *Roscoea* (Zingiberaceae). Edinburgh J. Bot. 57: 39–61.

Peng, C.-I., Tobe, H. & Takahashi, M. (1995): Reproductive morphology and relationships of *Triplostegia* (Dipsacales). Bot. Jahrb. Syst. 116: 505–516.

Pyck, N. (2001): Phylogenetic relations within Dipsacales. PhD thesis, Katholieke Universiteit Leuven, Belgium.

Pyck, N. & Smets, E. (2000): A search for the position of the seven-son flower (*Heptacodium*, Dipsacales): combining molecular and morphological evidence. Plant Syst. Evol. 225: 185–199.

Pyck, N., Roels, P. & Smets, E. (1999): Tribal relationships in Caprifoliaceae: evidence from a cladistic analysis using *ndh*F sequences. Syst. Geogr. Plants 69: 145–159.

Pyck, N., Van Lysebetten, A., Stessens, J. & Smets, E. (2002): The phylogeny of Patrinieae sensu Grabner (Valerianaceae)

revisited: additional evidence from *ndh*F sequence data. Plant Syst. Evol. 233: 29–46.

Rambaut, A. & Grassly, N. C. (1997): Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comp. Appl. Biosci. 13: 235–238.

Roels, P. & Smets, E. (1996): A floral ontogenetic study in Dipsacales. Int. J. Plant Sci. 157: 203–218.

Ronquist, F. (1996): DIVA, ver. 1.1. Computer program and manual available by anonymous FTP (ftp.uu.se or ftp.systbot.uu.se). Uppsala University, Sweden.

Ronquist, F. (1997): Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46: 195–203.

Swofford, D. L. (2002): PAUP* 4.0: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996): Phylogenetic inference. Pp. 407–514 in: Hillis, D. M., Moritz, C. & Mable, B. K. (eds.) Molecular Systematics. Second Edition. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. (1991): Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105–1109.

Urbatsch, L. E., Baldwin, B. G. & Donoghue, M. J. (2000): Phylogeny of the coneflowers and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. Syst. Bot. 25: 539–565.

White, T. J., Bruns, T., Lee, S. & Taylor, J. (1990): Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in: Innis, M., Gelfand, D. Sninsky, J. & White, T. J. (eds) PCR Protocols: a Guide to Methods and Applications. Academic Press, San Diego.

Young, N. D., Steiner, K. E. & dePamphilis, C. W. (1999): The evolution of parasitism in Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series. Ann. Missouri Bot. Gard. 86: 876–893.

Zhang, W.-H., Chen, Z.-D., Li, J.-H., Chen, H.-B. & Tang, Y.-C. (2003): Phylogeny of the Dipsacales s.l. based on chloroplast *trn*L-F and *ndh*F sequences. Mol. Phylog. Evol. 26: 176–189.