Int. J. Plant Sci. 165(4 Suppl.):S000–S000. 2004.
© 2004 by The University of Chicago. All rights reserved.
1058-5893/2004/1650S4-0006$15.00

DID CRYPTERONIACEAE REALLY DISPERSE OUT OF INDIA? MOLECULAR DATING EVIDENCE FROM *rbc*L, *ndh*F, AND *rpl*16 INTRON SEQUENCES

Frank Rutschmann,[1],* Torsten Eriksson,† Jürg Schönenberger,‡ and Elena Conti*

*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland; †Bergius Foundation, Royal Swedish Academy of Sciences, SE-104 05 Stockholm, Sweden; and ‡Department of Botany, Stockholm University, Lilla Frescativägen 5, SE-106 91 Stockholm, Sweden

Biogeographical and paleontological studies indicated that some ancient Gondwanan taxa have been carried by the rafting Indian plate from Gondwana to Asia. During this journey, the Indian island experienced dramatic latitudinal and climatic changes that caused massive extinctions in its biota. However, some taxa survived these conditions and dispersed "out of India" into South and Southeast Asia, after India collided with the Asian continent in the Early Tertiary. To test this hypothesis, independent estimates for lineage ages are needed. A published *rbc*L tree supported the sister group relationship between the South and Southeast Asian Crypteroniaceae (comprising *Crypteronia*, *Axinandra*, and *Dactylocladus*) and a clade formed by the African Oliniaceae, Penaeaceae, and Rhynchocalycaceae and the Central and South American Alzateaceae. Molecular dating estimates indicated that Crypteroniaceae split from their West Gondwanan sister clade in the Early to Middle Cretaceous and reached Asia by rafting on the Indian plate. Here we present molecular evidence from additional chloroplast DNA regions and more taxa to test the validity of the out-of-India hypothesis for Crypteroniaceae. Both clock-based (Langley-Fitch) and clock-independent age estimates (nonparametric rate smoothing and penalized likelihood) based on maximum likelihood analyses of three chloroplast DNA regions (*rbc*L, *ndh*F, and *rpl*16 intron) were used to infer the age of Crypteroniaceae. Our dating results indicate an ancient Gondwanan origin of Crypteroniaceae in the Early to Middle Cretaceous, followed by diversification on the Indian plate in the Early Tertiary and subsequent dispersal to Southeast Asia. These findings are congruent with recent molecular, paleontological, and biogeographic results in vertebrates. Within the biogeographic context of this study, we explore the critical assignment of paleobotanic and geologic constraints to calibrate ultrametric trees.

Keywords: molecular dating, molecular clock, r8s, rates of substitution, penalized likelihood, NPRS, clock calibration, biogeography, Gondwana, vicariance, Crypteroniaceae, Myrtales.

Introduction

Crypteroniaceae *sensu stricto* (Myrtales; Candolle 1857) are a small group of evergreen tropical shrubs and trees comprising three genera: *Crypteronia* Bl., with seven species, is the genus with the broadest distribution in Southeast Asia, including the Malay Peninsula, Sumatra, Java, Borneo, Philippines, Thailand, Vietnam, Myanmar, and New Guinea; *Dactylocladus* Oliv. has only one species, *Dactylocladus stenostachys*, endemic to Borneo; *Axinandra* Thw. includes one species, *Axinandra zeylanica*, endemic to Sri Lanka, and three other species with restricted distribution in the Malay peninsula and the northern part of Borneo (van Beusekom-Osinga and van Beusekom 1975; Johnson and Briggs 1984; Pereira and Wong 1995; Conti et al. 2002; fig. 1).

Southeast Asia and Sri Lanka are among the taxonomically most diverse regions on earth. In addition, they harbor high proportions of endemic species. For these reasons, both areas have been included among the 25 hot spots of biological diversity identified in a recent worldwide survey (Myers et al. 2000). This remarkable species richness can be partially explained by the geologic history of Sri Lanka and Southeast Asia. The uplift of the Himalayan chain caused by the collision of the Deccan plate (comprising India, Sri Lanka, and the Seychelles) with Laurasia during the Eocene (between 55 and 40 million years ago [mya]) and the generalized Late Tertiary aridification (Partridge 1997; Willis and McElwain 2002) led to an impoverishment of the tropical biome in Asia. Pockets of this biome, however, survived in refugial areas characterized by constant, tropical conditions, for example, in Sri Lanka and Southeast Asia. Only in these refugial areas did tropical plants have a chance to survive the detrimental effects of Quaternary climate oscillations on the Indian subcontinent (Raven and Axelrod 1974). The relictual nature of the Southeast Asian flora is also reflected in the great concentration of early diverging angiosperm clades in the fossil records of the subtropical forests of Asia-Australasia (Morley 2001).

Crypteroniaceae had been proposed as being an ancient and relictual group on the basis of their distribution and morphology (van Vliet and Baas 1975; van Beusekom-Osinga 1977). They represent an interesting case study to investigate the relative contributions of Laurasian and Gondwanan elements to the South Asian flora because their members had been alternatively suggested as being of Laurasian or Gondwanan origin.

[1] Author for correspondence; e-mail frank@plant.ch.

Manuscript received April 2003; revised manuscript received November 2003.

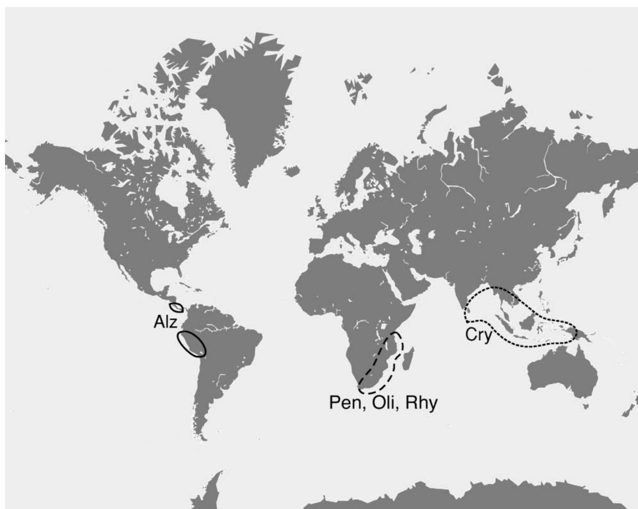


Fig. 1 Current distribution of Crypteroniaceae and related taxa. *Cry* = Crypteroniaceae, *Alz* = Alzateaceae, *Pen* = Penaeaceae, *Oli* = Oliniaceae, *Rhy* = Rhynchocalycaceae.

For example, Meijer (1972) postulated a Gondwanan origin for *Axinandra*, a genus that he interpreted as being morphologically similar to the ancestor of the entire order Myrtales. Furthermore, Ashton and Gunatilleke (1987, p. 263), referring to the biogeographic history of *Axinandra*, stated: "The disjunct distribution and generalized morphology of this lowland rain forest genus suggest considerable antiquity and possible spread into Asia by way of the Deccan Plate." The same authors suggested that *Axinandra* and other taxa were carried by the rafting Indian plate from Gondwana to Laurasia. After India collided with the Asian continent in the Early Tertiary, a few surviving Gondwanan elements dispersed "out of India" into South and Southeast Asia, which at the time lay in the same latitudinal and climatic zone (Morley 2000). The out-of-India origin of Crypteroniaceae was also supported in recent biogeographic studies based on molecular dating estimates (Conti et al. 2002; Morley and Dick 2003).

The idea that splitting plates may carry biotic elements from one continent to the other had already been proposed by Axelrod (1971) and McKenna (1973). However, Raven and Axelrod (1974) noted that it is difficult to find evidence for out-of-India dispersal because of the dramatic latitudinal and climatic changes that affected the Deccan plate during its journey from Gondwana to Laurasia and the ensuing massive extinctions in its biota. The same authors suggested a Laurasian origin for Crypteroniaceae (Raven and Axelrod 1974). Recent molecular phylogenetic analyses of *rbc*L sequences in Myrtales (Conti et al. 2002) supported that Crypteroniaceae form a monophyletic group comprising *Axinandra*, *Dactylocladus*, and *Crypteronia* and identified a sister clade comprising (1) Penaeaceae, a small group of 23 species in seven genera endemic to the Cape Province of South Africa; (2) Oliniaceae, comprising a single genus with eight species restricted to Eastern and Southern Africa; (3) Rhynchocalycaceae, with the single species *Rhynchocalyx lawsonioides*, a rare, evergreen tree endemic to the Eastern Cape and KwaZulu-Natal in South Africa (Johnson and Briggs 1984); and (4) Alzateaceae, with the single species *Alzatea verticillata*, a tree restricted to the submontane tropical forests of Bolivia, Peru, Panama, and Costa Rica (Graham 1984).

To investigate the biogeographic history of Crypteroniaceae, Conti et al. (2002) inferred the age of Crypteroniaceae by using three different molecular dating approaches applied to *rbc*L sequences. Because both phylogenetic relationships and dating estimates of relevant nodes were concordant with the geologic history of the Deccan Plate in relation to West Gondwanan continents, the authors suggested a West Gondwanan origin for Crypteroniaceae and related families, with subsequent dispersal of Crypteroniaceae to the Asian continent via India. However, these conclusions were based on evidence from only one gene (*rbc*L) and limited taxon sampling from Crypteroniaceae and related families.

In this article, we test the validity of previous conclusions on the out-of-India origin of Crypteroniaceae by expanding the taxon sampling to include four out of 12 described species of Crypteroniaceae and 13 out of 33 described species of their sister clade. Furthermore, we perform our analyses on DNA sequences of three chloroplast regions (*rbc*L, *ndh*F, *rpl*16 intron) and a combined data set, compare the results of clock-dependent (Langley-Fitch [LF], Langley and Fitch 1974) and clock-independent molecular dating methods (nonparametric rate smoothing [NPRS], Sanderson 1997; penalized likelihood [PL], Sanderson 2002), and evaluate the level of error in our divergence time estimates by implementing a bootstrap approach (Baldwin and Sanderson 1998; Sanderson and Doyle 2001). We also discuss how problems of calibration in molecular dating analyses affect different conclusions on possible biogeographic scenarios for our study system.

Material and Methods

Plant Material and DNA Extractions

For *Crypteronia paniculata*, *Crypteronia griffithii*, *Axinandra zeylanica*, *Dactylocladus stenostachys*, and *Olinia emarginata*, we extracted total genomic DNA from silica-dried leaf material. Leaf tissue was homogenized using glass beads and a MM 2000 shaker (Retsch GmbH, Haan, Germany). The DNA from these species was extracted with a method described in protocol D of Smith et al. (1991), which employs a 2% hexadecyl-trimethylammonium bromide (CTAB) extraction/lysis buffer. For all other taxa, the method of DNA extraction is given in Schönenberger and Conti (2003). Taxon names, voucher information, and GenBank accession numbers are listed in table 1.

PCR and DNA Sequencing

Amplification and sequencing primers from Zurawski et al. (1981), Olmstead and Sweere (1994), and Baum et al. (1998) were used to generate DNA sequences of *rbc*L, *ndh*F, and the *rpl*16 intron, respectively. PCR amplifications were performed in a Biometra TGradient thermocycler, applying a thermal cycling program that consisted of 34 cycles of 0.5 min at 95°C, 1 min at 49°C–52°C, and 1.7 min at 72°C, followed by a terminal extension of 10 min at 72°C. In order to successfully detect amplified DNA target regions and possible

Table 1

Species Names, Sources, and GenBank Accession Numbers of the DNA Sequences Used in the Analyses

Taxon	Voucher	GenBank accession numbers		
		*rbc*L	*ndh*F	*rpl*16 intron
Alzatea verticillata Ruiz & Pavon[a,b,c]	...	U26316[b]	AF215591[c]	AY151598[a]
Axinandra zeylanica Thwaites	Peter Ashton, s.n., Sri Lanka	AY078157[d]	AJ605094	AJ605107
Brachysiphon acutus (Thunb.) A. Juss.	J. Schönenberger 365 (Z), (BOL)	AJ605084	AJ605095	AY151605[a]
Brachysiphon fucatus (L.) Gilg	J. Schönenberger 357 (Z), (BOL)	AJ605085	AJ605096	AY151606[a]
Brachysiphon microphyllus Rourke	J. Schönenberger 386 (Z), (BOL)	AJ605086	AJ605097	AY151608[a]
Crypteronia griffithii C.B. Clarke	Shawn Lum s.n., Singapore	AJ605087	AJ605098	AJ605108
Crypteronia paniculata Blume	Peter Ashton s.n., Brunei	AY078153[d]	AJ605099	AY151597[a]
Dactylocladus stenostachys Oliver	Peter Becker, s.n., Brunei	AY078156[d]	AJ605100	AJ605109
Endonema retzioides A. Juss	J. Schönenberger 370 (Z), (BOL)	AJ605088	AJ605101	AY151611[a]
Eugenia uniflora L.[c]	...	AF294255[c]	AF215592[c]	AF215627[c]
Medinilla humbertiana Gaudich.[c]	...	AF215517[c]	AF215557[c]	AF215602[c]
Mouriri helleri Aublet[c,e]	...	AF270752[c]	AF322230[e]	AF215611[c]
Myrtus communis L.[c]	...	AF294254[c]	AF215593[c]	AF215628[c]
Olinia emarginata Davy	J. Schönenberger 579, cultivated, Kirstenbosch Botanical Garden, (Z)	AJ605089	AJ605102	AY151601[a]
Olinia ventosa (L.) Cuf.[a,c]	...	AF215546[c]	AF215594[c]	AY151604[a]
Osbeckia chinensis L.[c]	...	AF215525[c]	AF215570[c]	AF210378[c]
Penaea mucronata L.[a,d,c]	...	AJ605090	AF270756[c]	AY151620[a]
Rhexia virginica L.[b,c]	...	U26334[b]	AF215587[c]	AF215623[c]
Rhynchocalyx lawsonioides Oliver[a,b,c]	...	U26336[b]	AF270757[c]	AY151599[a]
Saltera sarcorolla (L.) Bullock	J. Schönenberger 360 (Z), (BOL)	AJ605091	AJ605103	AY151621[a]
Sonderothamnus petraeus (Barker f.) R. Dahlgren	J. Schönenberger 362 (Z), (BOL)	AY078154[d]	AJ605104	AY151622[a]
Stylapterus ericoides A. Juss. ssp.*pallidus* R. Dahlgren	J. Schönenberger 355 (Z), (BOL)	AJ605092	AJ605105	AY151625[a]
Stylapterus micranthus R. Dahlgren	M. Johns s.n. (Z)	AJ605093	AJ605106	AY151627[a]
Tibouchina urvilleana (DC.) Cogn.[b,f]	...	U26339[b]	AF272820[f]	AF322234[f]

Note. Herbaria acronyms: Z = Zurich, BOL = Bolus (University of Cape Town).
[a] Schönenberger and Conti 2003.
[b] Conti et al. 1996.
[c] Clausing and Renner 2001.
[d] Conti et al. 2002.
[e] Renner et al. 2001.
[f] Renner and Meyer 2001.

contamination, PCR products were separated on 1% agarose gels, stained with ethidium bromide, and visualized under UV light. Successfully amplified PCR products were purified with the QIAquick PCR Purification Kit (QIAGEN, Basel, Switzerland). Cycle-sequencing reactions were performed using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Applera Europe B.V., Rotkreuz, Switzerland). For a few taxa, we were unable to amplify the entire *rpl*16 intron; in these cases, two additional internal primers, MF and MR, were used (Schönenberger and Conti 2003). Cycle-sequencing reactions were performed in a GeneAmp PCR System 9700 (Applied Biosystems) by using a temperature cycle of 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C (25 cycles). The sequencing fragments were cleaned with MicroSpin G-50 columns (Amersham Pharmacia Biotech Europe GMBH, Dübendorf, Switzerland) to remove excess dye terminators before loading them on an ABI Prism 3100 genetic analyzer (Applied Biosystems).

The software Sequencher 3.1.1 (Gene Codes, Ann Arbor, Mich.) was used to edit and assemble complementary strands. Base positions were individually double-checked for agreement between the complementary strands. The *rbc*L sequences were readily aligned by eye, while *ndh*F and *rpl*16 intron sequences were first aligned using Clustal X 1.81 (Thompson et al. 1997) prior to adjusting the alignments by eye in the software MacClade 4.0 (Maddison and Maddison 2000). For the *rpl*16 intron data set, the variable region between nucleotides 810 and 1031 was deleted because we were unable to produce a reasonable alignment within that region. The data sets used for further phylogenetic analyses contained 24 taxa and 1280 (*rbc*L), 981 (*ndh*F), 1010 (*rpl*16 intron), and 3271 (all three data sets combined) aligned positions.

Phylogenetic Analyses

Maximum likelihood (ML) optimization was used to find the best tree for each of the three separate data partitions and for the combined data matrix, including all characters. A neighbor-joining tree calculated under the JC69 substitution model (Jukes and Cantor 1969) was used as the starting tree to estimate the optimal ML parameters under 56 different models of evolution in Modeltest 3.06 (Posada and Crandall 1998). The best substitution model was selected by performing hierarchical likelihood ratio tests (Felsenstein 1981; Huelsenbeck and Rannala 1997). The estimated parameters were

then used in a ML heuristic search using 100 random addition sequences, tree bisection reconnection branch swapping, and steepest descent activated, implemented in PAUP 4.0b10 (Swofford 2001). The trees were rooted on the branch leading to *Mouriri helleri* (Memecylaceae) and four representatives of Melastomataceae, which were constrained to be monophyletic. These choices were justified by the results of more inclusive phylogenetic analyses of Myrtales (Conti et al. 1996, 2002). Statistical support for each clade was estimated by generating 1000 bootstrap pseudoreplicates using the ML fast–heuristic search option in PAUP. All phylogenetic analyses were performed on a 2 Ghz Pentium IV machine under Red Hat Linux 8.0.

Molecular Dating

When performing molecular dating analyses, several crucial choices need to be made that might affect the estimated ages, including selection of molecular dating method, gene sampling, and calibration method. The first choice is to decide whether the analyses should be based on the assumption of rate constancy (molecular clock) or whether they should allow rates to vary across branches of a tree. To evaluate whether the sequences of each data partition evolved in a clocklike fashion, a likelihood ratio (LR) test was performed by comparing the scores of ML trees with and without a molecular clock enforced (Felsenstein 1981; Sanderson 1998; Nei and Kumar 2000). To gain some insight into the relative performance of different molecular dating methods, we compared the results of the clock-dependent LF (Langley and Fitch 1974) and the clock-independent NPRS (Sanderson 1997) and PL (Sanderson 2002) analyses, as implemented in r8s 1.6 (Sanderson 2003). Both latter methods relax the assumption of rate constancy by smoothing changes of substitution rates across the tree. NPRS is an entirely nonparametric method that estimates rates and times via a least squares smoothing criterion, whereas PL is a semiparametric technique that attempts to combine the statistical power of parametric methods with the robustness of nonparametric methods. Briefly, PL relies on a data-driven cross-validation procedure that sequentially prunes taxa from the tree, estimates parameters from the submatrix for a given smoothing value, predicts the data that were removed by using the estimated parameters, and calculates the χ^2 error associated with the difference between predicted and observed data of the removed submatrix. The optimal smoothing level corresponds to the lowest χ^2 error (Sanderson 2002).

The optimal ML trees estimated in PAUP 4.0b10 for each data partition and for the combined data set were saved with branch lengths (figs. 2–4, 5*A*) and then used as input trees in r8s. To establish the position of the root in the basal branch of the ML trees, *Myrtus communis* and *Eugenia uniflora* (Myrtaceae) were used as dating outgroups (see Conti et al. 1996, 1997). To evaluate the overall branch length from the root to a tip of a tree, it is necessary to know the lengths of the basal branches. In an additive tree, only the sum of these lengths is known, and the place where the root attaches to the basal branches is undefined (fig. 6*A*). Dating outgroup choice influences the position of the root attachment point, hence the lengths of the basal branches (fig. 6*B*) and the

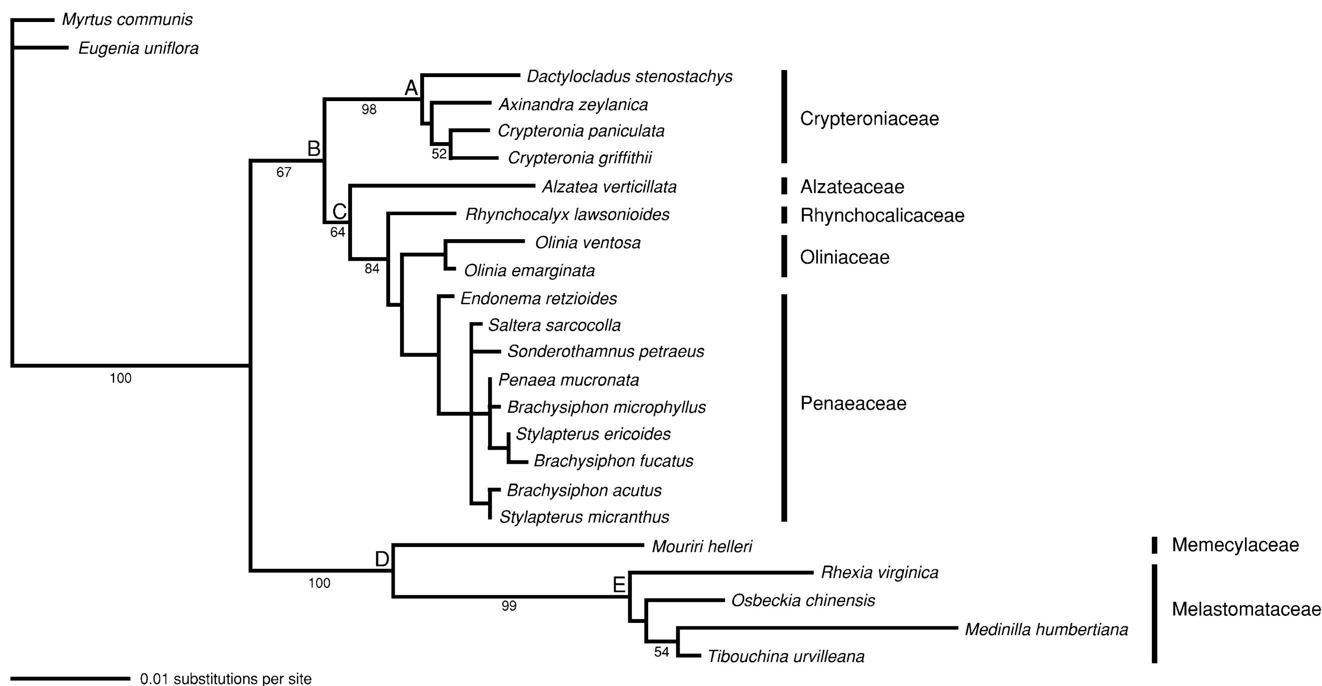


Fig. 2 Maximum likelihood tree based on the *rbc*L data set (1280 characters). Bootstrap support values are reported below the branches. Nodes of interest: *A* (diversification of Crypteroniaceae crown group), *B* (origin of Crypteroniaceae stem lineage), C (diversification of the West Gondwanan crown group), *D* (crown group of Melastomeae), and *E* (diversification of Melastomataceae crown group).

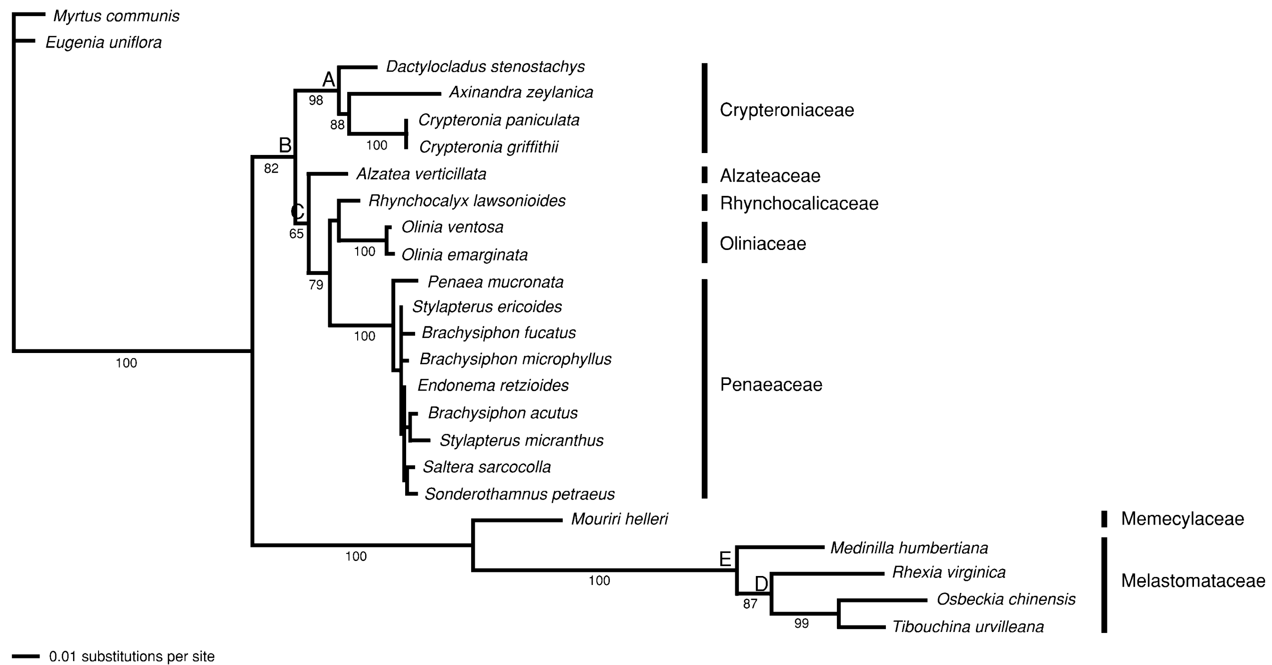


Fig. 3 Maximum likelihood tree based on the *ndh*F data set (981 characters). Bootstrap support values and description of nodes as in legend of fig. 2.

relative contribution of individual branches to the total paths from the root to the tips (Sanderson and Doyle 2001; Sanderson 2002). Therefore, root position affects the calculation of absolute substitution rates and the smoothing of differential substitution rates across the tree.

After root position was established, the dating outgroup was removed in r8s prior to molecular dating. For the PL analyses, the optimal smoothing parameter, ranging between 0.001 and 1000, was selected prior to the dating by performing a cross-validation procedure. To calculate the absolute substitution rates across the tree, optimization via the Truncated-Newton algorithm was chosen for the Langley-Fitch and PL methods and the POWELL algorithm for the NPRS dating.

All age estimations in r8s were started five times to provide different starting conditions (a random set of initial divergence times), a practice aimed at preventing the optimization algorithms from converging to a local plateau. Age estimations were performed only for nodes *A*, *B*, and *C* because these nodes are crucial to testing the out-of-India origin of Crypteroniaceae. Node *A* represents the diversification of the Crypteroniaceae crown group; node *B* represents the origin of the Crypteroniaceae stem lineage (equivalent to the time at which Crypteroniaceae split off from their West Gondwanan sister group); and node *C* represents the split between the South American *Alzatea* and its African sister clade (see figs. 2–5).

To evaluate statistical support for the estimated ages, we performed a bootstrap resampling procedure (Efron and Tibshirani 1993). For all molecular dating analyses performed on the combined data set, 100 bootstrap pseudoreplicates were generated using the program SEQBOOT from the Phylip package, version 3.6a3 (Felsenstein 2002). While the topology of the optimal ML trees was kept fixed, branch lengths for each pseudoreplicate were estimated by ML with the selected substitution model in PAUP (Sanderson 1997). With this approach, 100 bootstrap trees, with the same topology but different branch lengths, were generated and individually analyzed in 100 molecular dating procedures as described above. For the PL analysis, the optimal smoothing parameter for each bootstrap replicate was calculated prior to the dating procedures. Using the r8s bootstrap kit (Eriksson 2002), relative branch lengths from the 100 bootstrapped trees were transformed into a distribution of 100 absolute ages for each of nodes *A*, *B*, and *C*, respectively. After checking for normality, the obtained age distributions were used to calculate the mean, standard deviation, and 95% confidence interval (CI) of each age estimate (tables 2–4).

Calibration

To transform the resulting relative branch lengths into absolute ages for nodes *A*, *B*, and *C* (fig. 5*B*), it is necessary to fix or constrain a node to an absolute age. The calibration procedure represents one of the most sensitive choices in molecular dating analyses (Sanderson 1998; Wikström et al. 2001; Thorne and Kishino 2002; Yang and Yoder 2003). Calibration can be performed by reference either to the fossil record (paleobotanic dating) or to known vicariance events (geologic dating; Hillis et al. 1996; Sanderson 1998). Either approach can establish only the minimum age at the calibration point, most likely resulting in an underestimation of divergence times (Tavaré et al. 2002). In the following section, we consider the problems associated with each of the three calibration points that we selected for our analyses.

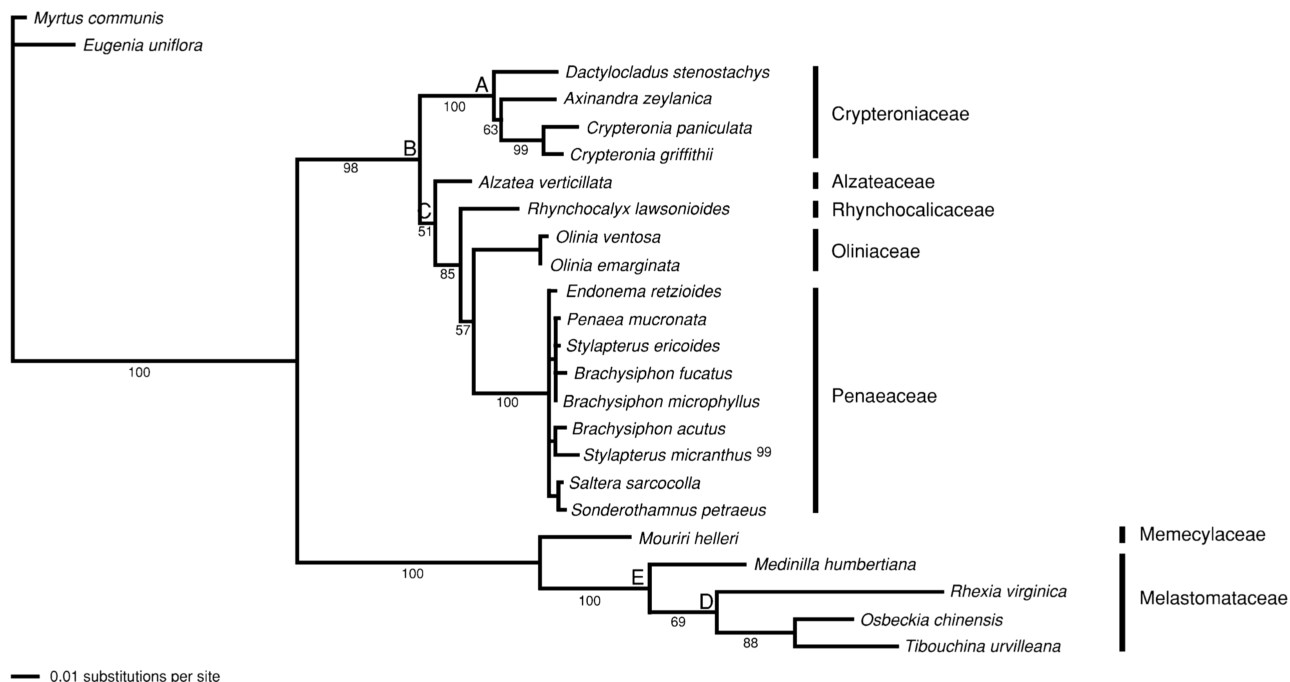


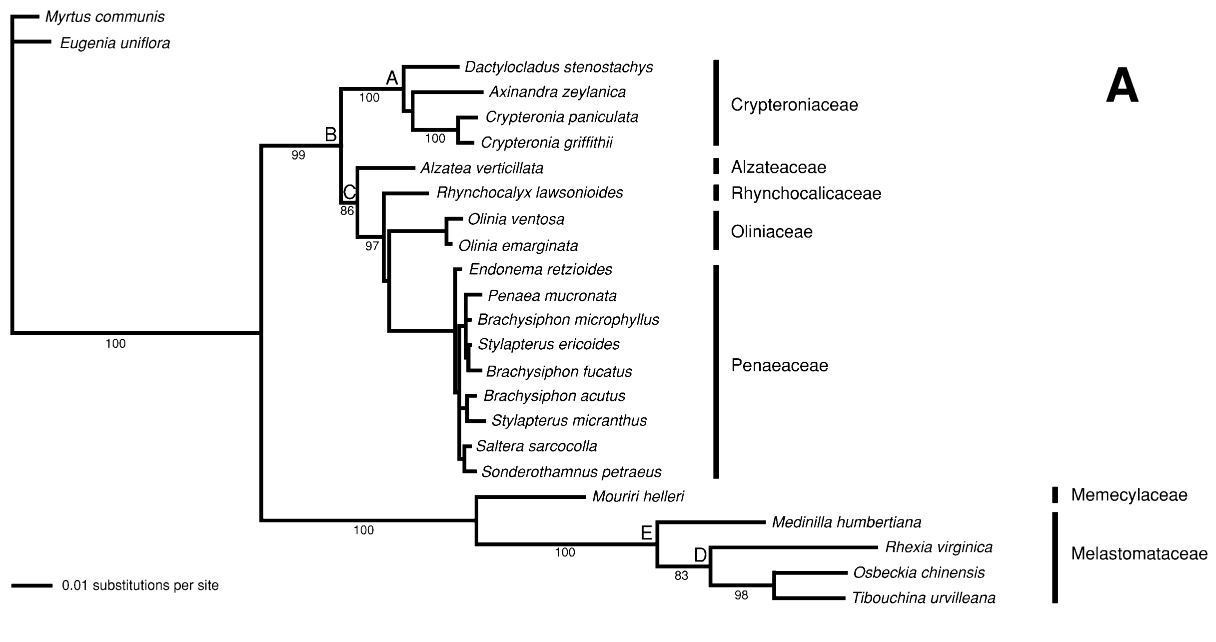
Fig. 4 Maximum likelihood tree based on the *rpl*16 intron data set (1010 characters). Bootstrap support values and description of nodes as in legend of fig. 2.

From an analytical point of view, the ideal calibration point would be as close as possible to the node to be estimated in order to reduce potential sources of error in age estimation (Wikström et al. 2001). In our tree, this was possible only with geologic calibration because the fossil record of Crypteroniaceae is too uncertain. Heterocolpate pollen tentatively assigned to Crypteroniaceae from the Middle Miocene (Muller 1975) is difficult to distinguish from heterocolpate pollen of Melastomataceae, Memecylaceae, Oliniaceae, Penaeaceae, and Rhynchocalycaceae (Morley and Dick 2003). Therefore, we were forced to look for paleobotanic calibration points outside of Crypteroniaceae.

The phylogenetically closest fossils were in Melastomataceae. Renner et al. (2001) and Renner and Meyer (2001) used fossil seeds from the Miocene of central Europe (Collinson and Pingen 1992) to constrain the origin of Melastomeae. However, the assignment of these seeds to Melastomeae is not straightforward. Collinson and Pingen (1992, p. 134) stated: "[These fossil seeds] are most similar to seeds of members of the tribes Osbeckieae [Melastomeae] and Rhexieae, but differ in several significant features, especially the presence of multicellular tubercles." Therefore, it is difficult to know whether these fossils should be assigned to the base of the Melastomeae crown group or to more recent nodes in the tribe. With these caveats in mind, we assigned a probably very conservative age of 26 mya (as suggested by Renner et al. 2001) to node *D*, representing the crown group of Melastomeae in our current taxon sampling (figs. 2–5).

Fossil leaves from the Early Eocene of North Dakota (53 mya; Hickey 1977) can also be used to constrain a node in Melastomataceae. However, the assignment of these macrofossils to a specific node is problematic. In his description of these fossil leaves, Hickey (1977, p. 144) stated that they resemble most closely the leaves of extant Miconieae and Merianieae but cautioned: "They all differ, however, in not being deeply cordate and in having tertiaries which do not form a good V pattern." Renner et al. (2001) conservatively assigned these leaves, characterized by the acrodromous leaf venation typical of extant Melastomataceae, to the node that subtends the crown group of the entire Melastomataceae, including the basal Kibessieae. However, these fossil leaves can also be assigned to the crown group that includes Miconieae and Micronieae, as the comments by Hickey (1977) might imply (see also Renner et al. 2001). Our current sampling of Melastomataceae does not include representatives of the basal Kibessieae. However, also in light of the biases in the macrofossil record discussed by Morley and Dick (2003), it does not seem unreasonable to assign an age of 53 mya to the node that comprises our current sampling of Melastomataceae (node *E*, see figs. 2–5; see also Renner 2004, in press).

Several recent studies have used geologic calibration points for molecular dating estimates, for example, in *Phylica* (Richardson et al. 2001), Laurales (Renner et al. 2000), ranid frogs (Bossuyt and Milinkovitch 2001), and ratite birds (Cooper et al. 2001). In our analyses, all ML trees from either separate or combined data sets supported the sister group relationship between the South American Alzateaceae and the African clade (see "Results"). This pattern, supported by a bootstrap value of 86% in the combined ML tree (see fig. 5*A*), represents a rather clear geologic signature and can be used as a calibration point, despite caveats of potential circularity. Therefore, we assigned an age of 90 mya, representing the final split between South America and Africa, to node *C* (see fig. 5*B*).



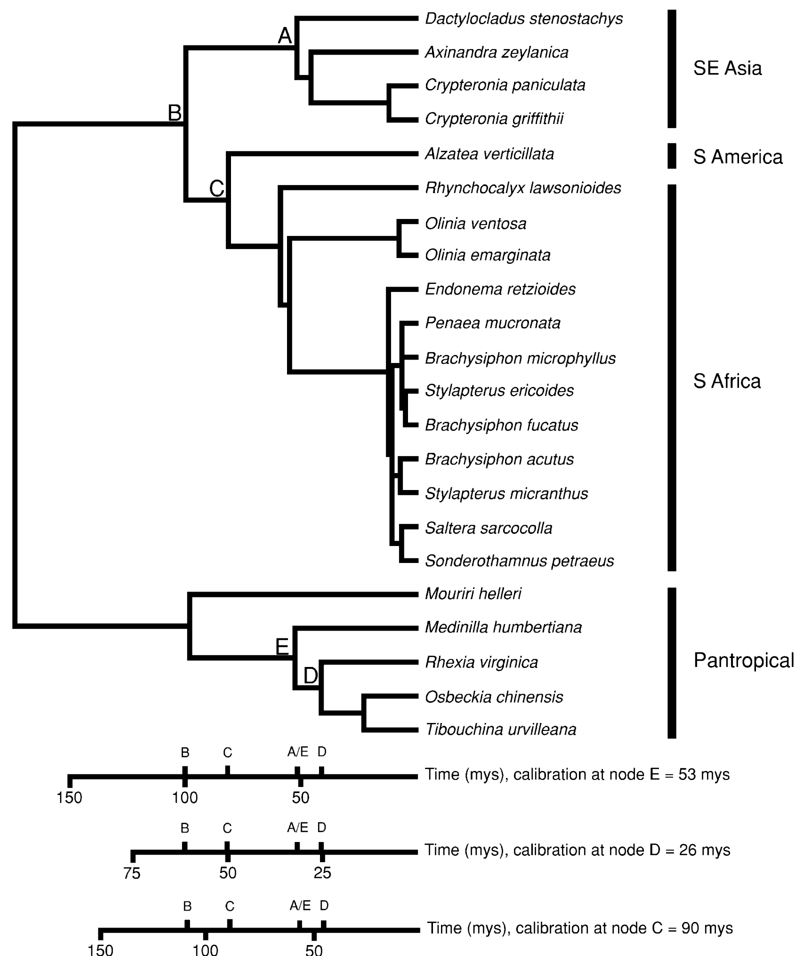


Fig. 5 *A*, Maximum likelihood tree based on the combined data set (3271 characters). Bootstrap support values and description of nodes as in legend of fig. 2. *B*, Chronogram based on a penalized likelihood molecular dating analysis of the combined data set. The independent use of three different calibration points resulted in three different time bars.

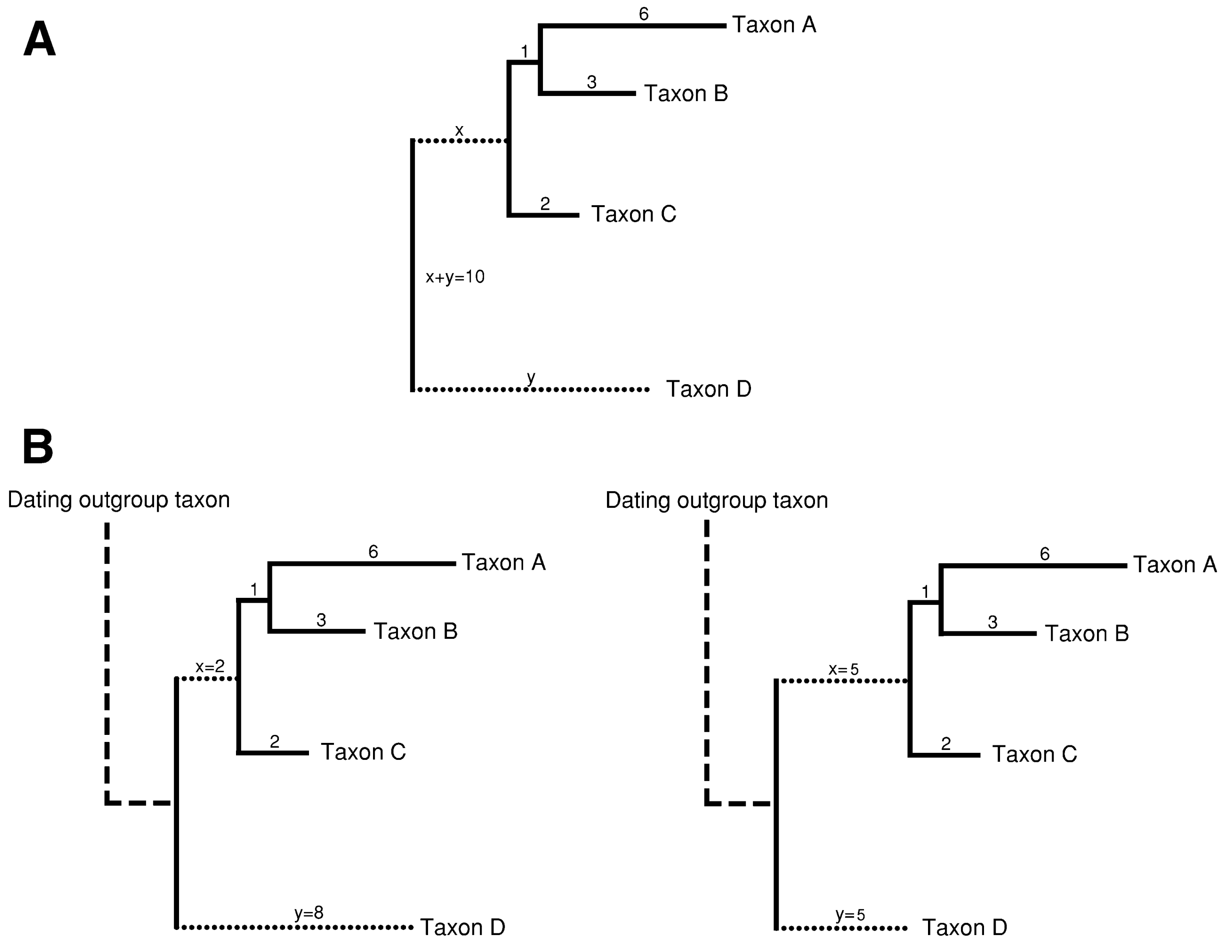


Fig. 6 Use of a dating outgroup taxon to evaluate where the root attaches to the basal branches. *A*, Additive tree to be analyzed by molecular dating. Only the sum of the lengths of the basal branches x and y is known because the point where the root attaches to the basal branches is unknown. *B*, Additional dating outgroup taxon was added to the tree to find where the root attaches to the basal branches. Two examples of possible root attachment points are shown, each determining different lengths of the basal branches x and y. Note that the two trees differ also in the overall branch lengths from the root to the tips.

Results

Phylogenetic Analyses

The selected optimal models of evolution were all submodels of the general time reversible (GTR) model (Rodríguez et al. 1990). For the *rbc*L data set, the K81uf + G + I model was selected (Kimura 1981): unequal base frequencies (A = 0.2673, C = 0.1941, G = 0.2488, T = 0.2898), one transition rate (AG/CT: 1.4245), two transversion rates (AC/GT: 1; AT/CG: 0.3281), gamma distribution of rates among sites with α shape parameter 0.7280 (Yang 1993), proportion of invariable sites 0.7098. For the *ndh*F data set, the TVM + G model was selected: unequal base frequencies (A = 0.3085, C = 0.1481, G = 0.1529, T = 0.3905), one transition rate (AG/CT: 1.5243), four transversion rates (AC: 1.2165, AT: 0.1544, CG: 1.4160, GT: 1), gamma distribution of rates among sites with α shape parameter 0.3887, no proportion of invariable sites. For the *rpl*16 intron data set, the K81uf + G + I model (Kimura 1981) was selected: unequal base frequencies (A = 0.3716, C = 0.1651, G = 0.1685, T = 0.2948), one transition rate (AG/CT: 1.3446), two transversion rates (AC/GT: 1, AT/CG: 0.4837), gamma distribution of rates among sites with α shape parameter 0.8358, no proportion of invariable sites. For the combined data set, the TVM + G + I model was selected: unequal base frequencies (A = 0.314, C = 0.1692, G = 0.1912, T = 0.3256), one transition rate (AG/CT: 1.5), four transversion rates (AC:1.1651, GT: 1, AT: 0.3469, CG: 0.7904), gamma distribution of rates among sites with α shape parameter 0.8288, proportion of invariable sites 0.4007. By using these parameters, one single optimal ML tree was found for each data set with a log-likelihood score of $-\ln L = 3408.68$ (*rbc*L), 4119.76 (*ndh*F), 4603.61 (*rpl*16 intron), and 12,460.2 (combined data set; figs. 2–4, 5*A*).

All optimal trees from the three individual and the combined data sets showed common results: (1) *Crypteronia*, *Axinandra*, and *Dactylocladus* (all Crypteroniaceae) formed a monophyletic group, with bootstrap support values (BS) between 98% and 100% (figs. 2–5*A*); (2) *Alzatea* (Alzateaceae),

Table 2

Ages in Million Years Estimated for Nodes *A, B,* and *C* Based on *rbc*L, *ndh*F, and *rpl*16 Intron Sequences and Three Different Methods Implemented in the r8s Software (Sanderson 2003; All Trees Were Calibrated at Node *E* with an Age of 53 Million Years)

	*rbc*L	*ndh*F	*rpl*16 intron	Combined data set	
LF:					
A	31.51	41.64	23.85	29.44	μ = 29.91, σ = 3.21, CI: 29.28–30.55
B	80.22	71.93	49.03	58.89	μ = 58.16, σ = 5.16, CI: 57.13–59.18
C	64.46	49.21	37.1	53.00	μ = 43.63, σ = 4.35, CI: 42.76–44.49
D	NA	45.43	42.4	42.69	μ = 42.31, σ = 2.08, CI: 41.9–42.72
NPRS:					
A	57.08	153.7	56.21	82.68	μ = 85.24, σ = 11.93, CI: 82.87–87.61
B	111.44	217.3	106	152.64	μ = 149.12, σ = 16.51, CI: 145.84–152.39
C	100.56	196.1	94.76	135.68	μ = 133.75, σ = 15.84, CI: 130.6–136.89
D	NA	42.4	40.15	38.16	μ = 38.35, n
PL:					
Optimal smoothing value:	α = 0.1	α = 0.01	α = 0.001	α = 0.00316	
A	39.37	92.75	43.06	53.00	μ = 54.26, σ = 7.58, CI: 52.76–55.76
B	96.91	154.58	91.09	100.70	μ = 103.28, σ = 12.32, CI: 100.84–105.73
C	80.26	128.08	76.19	83.03	μ = 85.91, σ = 12.45, CI: 83.44–88.38
D	NA	44.17	39.75	42.4	μ = 40.93, σ = 2.18, CI: 40.5–41.36

Note. LF = Langley-Fitch, NPRS = nonparametric rate smoothing, PL = penalized likelihood, μ = mean, σ = standard deviation, CI = 95% confidence interval, n = bootstrapped ages not normally distributed, NA = data not available because node *D* is not present in the *rbc*L data set due to different tree topology.

Rhynchocalyx (Rhynchocalycaceae), and all Penaeaceae and Oliniaceae included in this analysis formed another clade, with BS between 51% and 86%; (3) these two clades were sister to each other, with BS between 67% and 99%; and (4) *Alzatea* was sister to the clade formed by *Rhynchocalyx*, Oliniaceae, and Penaeaceae, with BS between 79% and 97%. These topological results were congruent with phylogenies published by Clausing and Renner (2001), Conti et al. (2002), and Schönenberger and Conti (2003). However, the trees differed slightly in the detailed topological resolution within the mentioned clades and in the branch lengths. In all these clades, the highest support values were obtained in the combined analysis.

Molecular Dating

By enforcing the molecular clock in PAUP 4.0b10, we obtained an optimal ML tree for each data set with log-likelihood scores of $-\ln L = 3429.36$ (*rbc*L), 4169.04 (*ndh*F), 4648.36 (*rpl*16 intron), and 12,531.27 (combined

Table 3

Ages in Million Years Estimated for Nodes *A, B,* and *C* Based on *rbc*L, *ndh*F, and *rpl*16 Intron Sequences and Three Different Methods Implemented in the r8s Software (Sanderson 2003; All Trees Were Calibrated at Node *D* with an Age of 26 Million Years)

	*ndh*F	*rpl*16 intron	Combined data set	
LF:				
A	23.83	14.63	17.93	μ = 18.44, n
B	41.17	30.06	35.86	μ = 35.82, σ = 3.59, CI: 35.11–36.53
C	28.17	22.75	26	μ = 26.87, σ = 2.92, CI: 26.29–27.44
E	30.33	32.5	32.28	μ = 32.65, σ = 1.64, CI: 32.32–32.98
NPRS:				
A	94.25	36.4	56.33	μ = 58.11, σ = 9.39, CI: 56.25–59.98
B	133.25	68.64	104	μ = 101.53, σ = 12.96, CI: 98.56–104.1
C	120.25	61.36	92.44	μ = 91.07, σ = 12.27, CI: 88.63–93.5
E	32.5	34.32	36.11	μ = 31.78, n
PL:				
Optimal smoothing value:	α = 0.01	α = 0.001	α = 0.00316	
A	54.6	28.17	32.5	μ = 34.6, σ = 5.37, CI: 33.54–35.67
B	91	59.58	61.75	μ = 65.79, σ = 8.54, CI: 64.1–67.49
C	75.4	49.83	50.92	μ = 54.73, σ = 8.35, CI: 53.07–56.38
E	31.2	34.67	32.5	μ = 33.77, σ = 1.82, CI: 33.41–34.13

Note. LF = Langley-Fitch, NPRS = nonparametric rate smoothing, PL = penalized likelihood, μ = mean, σ = standard deviation, CI = 95% confidence interval, n = bootstrapped ages not normally distributed. Ages for the *rbc*L data set are not available because node *D* is not present in this data set due to different tree topology.

Table 4

Ages in Million Years Estimated for Nodes *A, B,* and *C* Based on *rbc*L, *ndh*F, and *rpl*16 Intron Sequences and Three Different Methods Implemented in the r8s Software (Sanderson 2003; All Trees Were Calibrated at Node *C* with an Age of 90 Million Years)

	*rbc*L	*ndh*F	*rpl*16 intron	Combined data set	
LF:					
A	44	76.15	57.86	62.07	μ = 62.08, σ = 7.35, CI: 60.63–63.54
B	112	131.54	118.93	124.14	μ = 120.52, σ = 10.25, CI: 118.49
D	NA	83.08	102.86	90	μ = 88.12, n
E	74	96.92	128.57	111.72	μ = 110.42, σ = 11.09, CI: 108.22–112.63
NPRS:					
A	51.08	70.54	53.39	54.84	μ = 57.6, σ = 6.97, CI: 56.21–58.98
B	99.73	99.73	100.68	101.25	μ = 100.5, σ = 4.23, CI: 99.66–101.34
D	NA	19.46	38.14	25.31	μ = 26.15, σ = 3.45, CI: 25.46–26.83
E	47.43	24.32	50.34	35.16	μ = 36.15, σ = 4.19, CI: 35.32–36.98
PL:					
Optimal smoothing value:	α = 0.1	α = 0.01	α = 0.001	α = 0.00316	
A	44.15	65.17	50.87	57.45	μ = 57.31, σ = 7.2, CI: 55.88–58.74
B	108.68	108.62	107.61	109.15	μ = 108.89, σ = 8.52, CI: 107.2–110.58
D	NA	31.03	46.96	45.96	μ = 43.79, σ = 7.02, CI: 42.4–45.18
E	59.43	37.24	62.61	57.45	μ = 56.71, σ = 8.39, CI: 55.04–58.37

Note. LF = Langley-Fitch, NPRS = nonparametric rate smoothing, PL = penalized likelihood, μ = mean, σ = standard deviation, CI = 95% confidence interval, n = bootstrapped ages not normally distributed, NA = data not available because node *D* is not present in the *rbc*L data set due to different tree topology.

data set). Comparisons between clock and nonclock trees by applying LR tests rejected clocklike evolution for all data sets (LR = 41.35, *rbc*L; 98.56, *ndh*F; 89.52, *rpl*16 intron; 142.12, combined data set; degrees of freedom = 22, confidence interval = 95%). The results of molecular dating analyses using both clock-dependent and clock-independent approaches for the three separate and for the combined data sets are summarized in tables 2–4, and the PL chronogram for the combined data set is shown in figure 5*B*. Smoothing parameter values of 0.1 (*rbc*L), 0.01 (*ndh*F), 0.001 (*rpl*16 intron), and 0.00316 (combined data set), selected via a cross-validation procedure in r8s, were used for penalized likelihood age estimations.

Discussion

Comparisons among Dating Methods

Comparisons among the nodal ages estimated by the three dating methods showed remarkable differences (tables 2–4), depending on the methods themselves but also on the position of the calibration node within the tree. A discrepancy between clock-based and clock-independent age estimates was expected because likelihood ratio tests strongly rejected the assumption of rate constancy for all data sets. Differences in the rates of nucleotide substitution between branches in a tree are also known as lineage effects (Britten 1986; Gillespie 1991).

In general, NPRS (Sanderson 1997), which relaxes the assumption of rate constancy by smoothing changes of substitution rates across the tree, consistently produced the highest rate differences between the branches (as visualized in the ratograms produced by r8s; data not shown); thus the ages estimated using NPRS were either much younger or older than those obtained by using LF (Langley and Fitch 1974), depending on the position of the calibration node and data partition. This is because NPRS tends to overfit the data, thus causing rapid rate fluctuations in certain regions of a tree (Sanderson 2002). The semiparametric PL method (Sanderson 2002) tries to alleviate this problem by selecting the optimal smoothing parameter via a data-driven cross-validation procedure (Green and Silverman 1994). The application of PL resulted in rates of nucleotide substitution as well as age estimates, which were for most branches between those calculated with LF and NPRS.

By calibrating the trees at nodes *E* or *D* (tables 2, 3), the ages estimated for nodes *A*, *B*, and *C* using NPRS were consistently older than those produced using PL, whereas the ages obtained using LF were younger than the PL results. The likely explanation for this effect lies in the two long branches below node *E* (figs. 2–5*A*). Depending on the dating method, different rates of nucleotide substitution are assigned to these branches (ratograms not shown), producing considerably different absolute ages at nodes located on the other side of the root of the tree.

Comparisons among DNA Regions

By calibrating the trees at nodes *E* or *D* (tables 2, 3), the ages for nodes *A*, *B*, and *C* estimated using the *ndh*F data set were generally much older than those based on the other two data sets. Reciprocally, when we calibrated the trees at node *C* (table 4), the ages for nodes *E* and *D* were much younger in the *ndh*F-based analysis than by using the *rbc*L or *rpl*16-intron data sets. Nodal ages estimated from *rbc*L and *rpl*16 intron sequences were similar to each other. Comparisons of the three ML phylograms show that the branches below node *E* are significantly longer in the *ndh*F phylogram (fig. 3) than the same branches in the *rbc*L and *rpl*16-intron phylograms (figs. 2, 4).

The phenomenon of striking differences in the tempo and mode of evolution between different genes is well known, but its effects on divergence time estimation are poorly understood (Goremykin et al. 1996; Sanderson and Doyle 2001). In this article, we can only speculate on possible explanations for the anomalous results obtained from *ndh*F sequences and suggest research directions that might prove fruitful to investigate the role of gene-specific effects on molecular dating estimates.

For example, the bias of nucleotide substitutions in both coding and noncoding sequences of the plant chloroplast genome is strongly dependent on the composition of the two flanking bases (Morton 1997*a*, 1997*b*). One possible explanation for the older ages obtained from *ndh*F sequences might lie in the differential influence that the two neighboring bases could have on the substitution type of a certain nucleotide in our *ndh*F sequences as compared to *rbc*L and *rpl*16 intron sequences. It is also reasonable to ask whether the occurrence of an *ndh*F pseudogene might explain gene-specific effects on age estimates, as *ndh*F pseudogenes have been reported, for example, in orchids (Neyland and Urbatsch 1996). However, translation of *ndh*F sequences into the corresponding amino acids did not reveal the presence of any stop codons, and multiple alignment of *ndh*F sequences required only gaps in multiples of three, suggesting that our *ndh*F sequences likely represent functional gene copies. Furthermore, PCR amplifications using an *ndh*F-specific primer pair (Olmstead and Sweere 1994) did not reveal any PCR products of different lengths, and direct sequencing of double-stranded PCR products produced unequivocal electropherograms, characterized by single peaks at all positions. Finally, the topology of the optimal ML tree–based *ndh*F was congruent to the other trees based on the *rbc*L and *rpl*16 intron sequences. Nevertheless, we cannot exclude the possibility that an *ndh*F pseudogene might exist in some or all of the studied taxa, perhaps influencing the molecular evolutionary behavior of the functional *ndh*F gene copies that we likely sequenced for this study (Bromham and Penny 2003). Another potential explanation for the different dating results obtained from *ndh*F sequences might be sought in alignment effects. However, experiments using a modified *ndh*F data set from which all gapped regions were removed prior to analysis produced age estimates similar to those obtained with the gapped data set (data not shown). Additional theoretical and experimental studies of gene-specific effects on molecular dating estimates are clearly needed (Bromham and Penny 2003).

Although nested likelihood ratio tests of the three separate data sets indicated that the three chloroplast regions used in our dating analyses evolved according to different models and parameters of nucleotide substitutions, we proceeded to estimate nodal ages from the combined data matrix because we wished to compare results from the latter with those from the three separate sequence matrices. The exceedingly older or younger nodal ages estimated from *ndh*F sequences, depending on the position of the calibration point, seemed to further justify data set combination, for it has been suggested that the combination of sequences with different evolutionary patterns might compensate for unusual patterns in any single DNA region (Qiu et al. 1999; Wikström et al. 2001).

Congruence between Geologic and Biological History

One of the major goals of biogeographic studies is to elucidate the historical genesis of current plant distributions. In an evolutionary framework, it is assumed that geologic events of the past, for example the emergence and/or the elimination of major geographic barriers to range expansion, likely left a mark on the phylogenetic and biogeographic history of biotic elements (Lieberman 2000). Therefore, to support the hypothesis that geologic events shaped the current distribution of any taxa, one would need to demonstrate congruence between geologic and biological history both in terms of pattern and time (table 5). At the level of pattern, one would expect correspondence between the sequence of geologic events and the sequence of cladogenetic events. Paleogeologic reconstructions and the topology of phylogenetic trees provide the necessary evidence for pattern congruence. At the level of time, the specific timing of geologic events must be compatible with the timing of cladogenetic events (nodal ages) inferred from molecular or other dating methods. Both lines of evidence (pattern and time) are necessary, but independently not sufficient, to support a key role of geology in shaping current taxic distributions. If evidence for congruence between geology and biology can be produced at the levels of both pattern and time, there is no need to invoke other types of explanatory processes for current biotic distributions (table 5; see also Sober 1988; Hunn and Upchurch 2001).

Geologic events that influence biological distributions include plate fragmentation, as in the classic interpretation of vicariance or the expansion of a lineage due to the temporary elimination or reduction of a geographic barrier, followed by the emergence of a new barrier producing vicariant sister groups, as in the recently proposed concept of geodispersal (Lieberman 1997, 2000). In the next section, we will discuss whether the phylogenetic relationships and molecular dating estimates of Crypteroniaceae warrant a key role for geologic events in explaining the current distribution of this group and its sister clade.

Congruence between Geology and Biology for the Out-of-India Hypothesis of Crypteroniaceae

After comparing the results of different dating methods and data sets (see above), it seemed most reasonable to use the ML tree topology (fig. 5*A*) and the PL ages (fig. 5*B*) calculated from the combined data matrix to reconstruct the

Table 5

Relationships between Geology and Biology at the Levels of Pattern and Time

	Geology	Biology
Pattern	Sequence of geologic events	Sequence of cladogenetic events
Time	Timing of geologic events	Timing of cladogenetic events

Note. Correspondence at both levels is necessary to support a geodispersalist origin (*sensu* Lieberman 2000) of current biotic distributions. See "Discussion" for further explanation.

biogeographic history of Crypteroniaceae. A previous phylogenetic and molecular dating study based exclusively on *rbc*L sequences proposed an ancient Gondwanan origin for Crypteroniaceae in the Early to Middle Cretaceous, followed by dispersal to the Deccan plate (comprising Madagascar, India, Sri Lanka, and the Seychelles) as it was rafting along the African coast, and subsequent dispersal from India to Southeast Asia after collision of the Indian plate with Asia in the Middle Eocene (Conti et al. 2002). Is this biogeographic reconstruction congruent with both pattern and timing of cladogenetic events (table 5), as estimated from the phylogenetic and molecular dating analyses of the expanded data sets used in this study?

The combined ML tree (fig. 5*A*) strongly supports (BS = 99%) the sister group relationship between the Southeast Asian Crypteroniaceae and the West Gondwanan clade and the split between the South American Alzateaceae and the African clade (BS = 86%). Therefore, at the level of pattern, the sequence of cladogenetic events is congruent with the sequence of geologic events, if we consider that the Deccan Plate rafted along the African coast between the Lower and Middle Cretaceous (Scotese et al. 1988; Morley 2000), with likely island chain connections between the two plates up to the Early Maastrichtian (Morley 2000), and that separation between Africa and South America was completed by ca. 90 mya (McLoughlin 2001), although transoceanic dispersal routes between Africa and South America likely existed between 84 and 65 mya (McDougal and Douglas 1988; Hallam 1994; Morley 2000).

At the level of time, results are more controversial. The deviations in age estimates due to the use of different calibrations (tables 2–4) indicate that calibration is one of the most critical issues in molecular phylogenetic dating. The ages for the origin of Crypteroniaceae (node *B*), obtained from PL optimization on the combined ML tree, ranged from a minimum value of 62 mya, estimated by fixing node *D* to 26 mya to a higher value of 101 mya, estimated by fixing node *E* to 53 mya and a maximum value of 109 mya, estimated by fixing node *C* to 90 mya (see tables 2–4; fig. 5*B*). As explained in "Material and Methods," the assignments of fossil seeds from the Miocene of central Europe (Collinson and Pingen 1992) to node *D* (Melastomeae crown group) and fossil leaves from the Eocene of North Dakota (Hickey 1977) to node *E* (Melastomataceae crown group) most likely represent large underestimations of nodal ages. Furthermore, Morley and Dick (2003) extensively reviewed the fossil record for Melastomataceae and argued that its abrupt appearance at northern temperate latitudes during the Eocene and Miocene may simply reflect colonization from ancient Gondwanan lineages. Given these considerations, it seems more plausible to suggest an origin of the Crypteroniaceae stem lineage that is closer to the older ages (101–106 mya) obtained with our three calibration points. Which biogeographic scenario is congruent with this interpretation for the age of Crypteroniaceae?

According to paleogeographic reconstructions, East Gondwana—including India—split from West Gondwana between 165 and 150 mya (Krutzsch 1989; McLoughlin 2001; Briggs 2003). Therefore, a traditional vicariant explanation for the origin of Crypteroniaceae—with overland dispersal from West to East Gondwana, followed by tectonic split—is incompatible with our dating estimates for node *B* and indeed with molecular estimates for the age of angiosperms (190–140 mya; Sanderson and Doyle 2001; Wikström et al. 2001). It is more probable that the biogeographic history of Crypteroniaceae might reflect a temporary reduction or even elimination of the oceanic barrier between Africa and the Deccan Plate (at that time comprising Madagascar, India, Sri Lanka, and the Seychelles Plateau), as the plate drifted northward along the African coast for a rather extended period of time (over 40 mya) between the Early and Late Cretaceous (Scotese et al. 1988; Morley 2000; Briggs 2003). It has also been suggested that small islands or land bridges between West Gondwana and the Deccan Plate facilitated short- to medium-distance dispersal over the Mozambique Channel of other biotic elements, including some groups of dinosaurs, crocodiles, mammals (Krause and Maas 1990; Krause et al. 1999), frogs (Bossuyt and Milinkovitch 2001; Biju and Bossuyt 2003), lizards, snakes, turtles, and caecilians (Briggs 2003). Ashton and Gunatilleke (1987) suggested that the total distance between West Gondwana and the Deccan plate (still connected to Madagascar) remained more or less constant (ca. 420 km) until ca. 84 mya, when the plate separated from Madagascar and started to drift northward (Plummer and Belle 1995; Storey et al. 1995; McLoughlin 2001). Pollen records suggested that plant dispersal from Africa to Madagascar and the Indian plate continued on a regular basis, presumably until the middle Maastrichtian (65–71 mya ago; Morley and Dick 2003). Therefore, India's role in the biogeographic history of Crypteroniaceae most likely did not conform to a purely vicariant pattern, involving direct dispersal prior to barrier formation (Wiley 1988; Morrone and Crisci 1995), but rather to the dynamics of range expansion following barrier reduction (geodispersal; Lieberman 2000; see also Stace 1989).

Extinction played a prominent role in the history of the ancient Gondwanan elements of India's biotas, as India traveled rapidly across latitudes during the Middle to Late Cretaceous (Morley 2000; McLoughlin 2001). Its biotas were affected by massive volcanism at the Cretaceous-Tertiary boundary (ca. 65 mya ago; Officer et al. 1987), extensive aridification during the Late Tertiary (following the uplift of the Himalayan chain caused by India's collision with Southern Asia between 55 and 49 mya; Beck et al. 1995), and further cycles of aridity associated with glaciations during the Quaternary (Raven and Axelrod 1974; Bande and Prakash 1986; Ashton and Gunatilleke 1987; Morley 2000). *Axinandra zeylanica* is endemic in Sri Lanka, which was probably connected to India until 6000 years ago (McLoughlin 2001). Southwestern India together with Sri Lanka served as refugial areas, where some ancient Gondwanan taxa escaped extinction (Raven and Axelrod 1974; Guleria 1992; Morley 2000). Some of these relictual taxa dispersed to Southeast Asia, where *Crypteronia* sp., *Dactylocladus stenostachys*, and the other three species of *Axinandra* occur to this day. Southeast Asia has also long been recognized as a refugium where the equable oceanic conditions allowed tropical lineages to survive (Bande and Prakash 1986; Takhtajan 1987; Morley 2000).

To summarize, our current phylogenetic and molecular dating results from expanded taxic and genetic sampling suggest a possible congruence between biological and geologic

history that is compatible with a central role played by the Deccan Plate in transporting the stem lineage of Crypteroniaceae from West Gondwana to Asia, most likely in a time frame comprised between the Middle and Late Cretaceous. However, our results remain open to debate, especially in light of the difficult assignment of paleobotanic and geologic constraints to specific nodes in the phylogeny. It is our hope that the addition of more fossil calibration points; further taxonomic sampling from Crypteroniaceae and additional groups; and the use of dating methods that allow for multiple, contemporary constraints on the phylogeny will allow us to refine our interpretations of the biogeographic history of Crypteroniaceae and related clades.

Acknowledgments

We are very grateful to Sandro Wagen for technical assistance in the laboratory. We also thank Reto Nyffeler, Alexander Kocyan, and Niklaus Müller for fruitful discussions and general help. Financial support from the University of Zurich is gratefully acknowledged.

Literature Cited

Ashton PS, CVS Gunatilleke 1987 New light on the plant geography of Ceylon. I. Historical plant geography. J Biogeogr 14:249–285.

Axelrod DI 1971 Plate tectonics in relation to the history of angiosperm vegetation in India. Birbal Sahni Inst Paleobot Spec Publ 1:5–18.

Baldwin BG, MJ Sanderson 1998 Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc Natl Acad Sci USA 95:9402–9406.

Bande MB, U Prakash 1986 The tertiary flora of Southeast Asia with remarks on its paleoenvironment and phytogeography of the Indo-Malayan region. Rev Paleobot Palynol 49:203–233.

Baum DA, RL Small, JF Wendel 1998 Biogeography and floral evolution of Baobabs (Adansonia, Bombacaceae) as inferred from multiple datasets. Syst Biol 47:181–207.

Beck RA, DW Burbank, WJ Sercombe, GW Riley, JK Barndt, JR Berry, J Afzal, et al 1995 Stratigraphic evidence for an early collision between northwest India and Asia. Nature 373:55–58.

Biju SD, F Bossuyt 2003 New frog family from India reveals an ancient biogeographical link with the Seychelles. Nature 425: 711–714.

Bossuyt F, MC Milinkovitch 2001 Amphibians as indicators of early Tertiary "out-of-India" dispersal of vertebrates. Science 292:93–95.

Briggs JC 2003 The biogeographic and tectonic history of India. J Biogeogr 30:381–388.

Britten RJ 1986 Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1393–1398.

Bromham L, D Penny 2003 The modern molecular clock. Nat Rev Genet 4:216–224.

Candolle AD 1857 Penaeaceae. Pages 483–491 *in* Ad Candolle, ed. Prodromus systematis naturalis regni vegetabilis. Vol 14. Masson, Paris.

Clausing G, SS Renner 2001 Molecular phylogenetics of Melastomataceae and Memecylaceae: implications for character evolution. Am J Bot 88:486–498.

Collinson ME, M Pingen 1992 Seed of the Melastomataceae from the Miocene of Central Europe. Pages 129–139 *in* J Kovar-Eden, ed. Paleovegetation development in Europe. Museum of Natural History, Vienna.

Conti E, T Eriksson, J Schönenberger, KJ Sytsma, DA Baum 2002 Early Tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. Evolution 56: 1931–1942.

Conti E, A Litt, KJ Sytsma 1996 Circumscription of Myrtales and their relationships to other Rosids: evidence from *rbc*L sequence data. Am J Bot 83:221–233.

Conti E, A Litt, PG Wilson, SA Graham, BG Briggs, LAS Johnson, KJ Sytsma 1997 Interfamilial relationships in Myrtales: molecular phylogeny and patterns of morphological evolution. Syst Bot 22: 629–647.

Cooper A, C Lalueza-Fox, S Anderson, A Rambaut, J Austin, R Ward 2001 Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. Nature 409:704–707.

Efron B, RJ Tibshirani 1993 An introduction to the bootstrap. Chapman & Hall, New York.

Eriksson T 2002 The r8s-bootstrap-kit. Bergius Botanic Garden, Stockholm.

Felsenstein J 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376.

——— 2002 PHYLIP (phylogeny inference package), version 3.6a3. Department of Genetics, University of Washington, Seattle.

Gillespie JH 1991 The causes of molecular evolution. Oxford University Press, New York.

Goremykin V, V Bobrova, J Pahnke, A Troitsky, A Antonov, W Martin 1996 Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbc*L data do not support Gnetalean affinities of angiosperms. Mol Biol Evol 13:383–396.

Graham SA 1984 Alzateaceae, a new family of Myrtales from the American Tropics. Ann Mo Bot Gard 71:757–779.

Green PJ, BW Silverman 1994 Nonparametric regression and generalized linear models. Chapman & Hall, London.

Guleria JS 1992 Neogene vegetation of peninsular India. Paleobotanist 40:285–311.

Hallam A 1994 An outline of Phanerozoic biogeography. Oxford Biogeography Series, no. 10. Oxford University Press, Oxford.

Hickey LJ 1977 Stratigraphy and paleobotany of the Golden Valley Formation (Early Tertiary) of Western North Dakota. Pages 143–144 *in* Geological Society of America, Memoir 150. Geological Society of America, Boulder, Colo.

Hillis DM, BK Mable, C Moritz 1996 Applications of molecular systematics: the state of the field and a look at the future. Pages 515–543 *in* DM Hillis, C Moritz, BK Mable, eds. Molecular systematics. Sinauer, Sunderland, Mass.

Huelsenbeck JP, B Rannala 1997 Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276:227–232.

Hunn CA, P Upchurch 2001 The importance of time/space in diagnosing the causality of phylogenetic events: towards a "chronobiological" paradigm? Syst Biol 50:391–407.

Johnson LAS, BG Briggs 1984 Myrtales and Myrtaceae: a phylogenetic analysis. Ann Mo Bot Gard 71:700–756.

Jukes TH, CR Cantor 1969 Evolution of protein molecules. Pages 21–132 *in* HN Munro, ed. Mammalian protein metabolism. Academic Press, New York.

Kimura M 1981 Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA 78: 454–458.

Krause DW, MC Maas 1990 The biogeographic origins of late Paleocene–early Eocene mammalian immigrants to the Western Interior of North America. Pages 71–105 *in* TM Bown, KD Rose,

eds. Dawn of the age of mammals in the northern part of the Rocky Mountain Interior, North America. Geological Society of America Special Paper 243, Boulder, Colo.

Krause DW, RR Rogers, CA Forster, JH Hartman, GA Buckley, SD Sampson 1999 The Late Cretaceous vertebrate fauna of Madagascar: implications for Gondwanan paleobiogeography. GSA Today (Geol Soc Am) 9:1–7.

Krutzsch W 1989 Paleogeography and historical phytogeography (paleochorology) in the Neophyticum. Plant Syst Evol 162: 5–61.

Langley CH, W Fitch 1974 An estimation of the constancy of the rate of molecular evolution. J Mol Evol 3:161–177.

Lieberman BS 1997 Early Cambrian paleogeography and tectonic history: a biogeographic approach. Geology 25:1039–1042.

——— 2000 Paleobiogeography: using fossils to study global change, plate tectonics, and evolution. Kluwer, New York.

Maddison PG, DR Maddison 2000 MacClade 4: analysis of phylogeny and character evolution. Sinauer, Sunderland, Mass.

McDougal I, RA Douglas 1988 Age-progressive volcanism in the Tasmanitid seamounts. Earth Planet Sci Lett 89:207–220.

McKenna MCC 1973 Sweepstakes, filters, corridors, Noah's arks, and beached Viking funeral ships in paleogeography. Pages 291–304 *in* DH Tarling, SK Runcorn, eds. Implications of continental drift to the earth sciences. Academic Press, London.

McLoughlin S 2001 The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Austr J Bot 48: 271–300.

Meijer W 1972 The genus Axinandra-Melastomataceae: a missing link in Myrtales? Ceylon J Sci Biol Sci 10:72–76.

Morley RJ 2000 Origin and evolution of tropical rain forests. Wiley, Chichester.

——— 2001 Why are there so many primitive angiosperms in the rain forests of Asia-Australasia? Pages 185–199 *in* I Metcalfe, JMB Smith, M Morwood, I Davidson, eds. Faunal and floral migrations and evolution in Southeast Asia–Australasia. Balkeme, Lisse.

Morley RJ, CW Dick 2003 Missing fossils, molecular clocks and the origin of the Melastomataceae. Am J Bot 90:1638–1645.

Morrone JJ, JV Crisci 1995 Historical biogeography: introduction to methods. Annu Rev Ecol Syst 26:373–401.

Morton BR 1997*a* The influence of neighboring base composition on substitutions in plant chloroplast coding sequences. Mol Biol Evol 14:189–194.

——— 1997*b* The influence of specific neighboring bases on substitution bias in noncoding regions of the plant chloroplast genome. J Mol Evol 45:227–231.

Muller J 1975 Note on the pollen morphology of Crypteroniaceae s.l. Blumea 22:275–294.

Myers N, RA Mittermeier, CG Mittermeier, GAB da Fonseca, J Kent 2000 Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Nei M, S Kumar 2000 Molecular evolution and phylogenetics. Oxford University Press, Oxford.

Neyland R, LE Urbatsch 1996 Phylogeny of the subfamily Epidendroideae (Orchidaceae) inferred from *ndh*F chloroplast gene sequences. Am J Bot 83:1195–1206.

Officer CB, CL Hallam, CL Drake, JD Devine 1987 Late Cretaceous and paroxysmal Cretaceous/Tertiary extinctions. Nature 326: 143–149.

Olmstead RG, JA Sweere 1994 Combining data in phylogenetic systematics: an empirical approach using three molecular datasets in the Solanaceae. Syst Biol 43:467–481.

Partridge TC 1997 Late Neogene uplift in eastern and southern Africa and its paleoclimatic implications. Pages 63–86 *in* WF Ruddiman, ed. Tectonic uplift and climate change. Plenum, New York.

Pereira JT, KM Wong 1995 Three new species of Crypteronia (Crypteroniaceae) from Borneo. Sandakania 6:41–53.

Plummer PS, ER Belle 1995 Mesozoic tectono-stratigraphic evolution of the Seychelles microcontinent. Sediment Geol 96:73–91.

Posada D, KA Crandall 1998 Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Qiu YL, JH Lee, F Bernasconi-Quadroni, DE Soltis, PS Soltis, M Zanis, EA Zimmer, ZD Chen, V Savolainen, MW Chase 1999 The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402:404–407.

Rambaut A, M Charleston 2001 TreeEdit: an application for organising, viewing and manipulating sets of phylogenetic trees. University of Oxford, Oxford.

Raven PH, DI Axelrod 1974 Angiosperm biogeography and past continental movements. Ann Mo Bot Gard 61:539–673.

Renner SS 2004 Bayesian analysis of combined data partitions, using multiple calibrations, supports recent arrival of Melastomataceae in Africa and Madagascar. Am J Bot (in press).

Renner SS, G Clausing, K Meyer 2001 Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. Am J Bot 88:1290–1300.

Renner SS, BD Foreman, D Murray 2000 Timing trans-antarctic disjunctions in the Atherospermataceae (Laurales): evidence from coding and noncoding chloroplast sequences. Syst Biol 49:579–591.

Renner SS, K Meyer 2001 Melastomeae come full circle: biogeographic reconstruction and molecular clock dating. Evolution 55: 1315–1324.

Richardson JE, FM Weitz, MF Fay, QCB Cronk, HP Linder, G Reeves, MW Chase 2001 Rapid and ancient origin of species richness in the Cape Flora of South Africa. Nature 412:181–183.

Rodríguez F, JF Oliver, A Marín, JR Medina 1990 The general stochastic model of nucleotide substitution. J Theor Biol 142: 485–501.

Sanderson MJ 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol Biol Evol 14: 1218–1231.

——— 1998 Estimating rates and time in molecular phylogenies: beyond the molecular clock? Pages 242–264 *in* DE Soltis, PS Soltis, JA Doyle, eds. Molecular sytematics of pants. II. DNA sequencing. Kluwer, Boston.

——— 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol Biol Evol 19:101–109.

——— 2003 r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19:301–302.

Sanderson MJ, JA Doyle 2001 Sources of error and confidence intervals in estimating the age of angiosperms from *rbc*L and 18S rDNA data. Am J Bot 88:1499–1516.

Schönenberger J, E Conti 2003 Molecular phylogeny and floral evolution of Penaeaceae, Oliniaceae, Rhynchocalycaceae, and Alzateaceae (Myrtales). Am J Bot 90:293–309.

Scotese CR, LM Gahagan, RL Larson 1988 Plate tectonic reconstructions of the Cretaceous and Cenozoic ocean basins. Tectonophysics 155:27–48.

Smith JF, KS Sytsma, JS Shoemaker, RL Smith 1991 A qualitative comparison of total cellular DNA extraction protocols. Phytochem Bull 23:2–9.

Sober E 1988 The conceptual relationship of cladistic phylogenetics and vicariance biogeography. Syst Zool 37:245–253.

Stace CA 1989 Dispersal versus vicariance: no contest. J Biogeogr 16: 200–201.

Storey M, JJ Mahoney, AD Saunders, RA Duncan, SP Kelley, MF Coffin 1995 Timing of hot spot–related volcanism and the breakup of Madagascar and India. Science 267:852–855.

Swofford DL 2001 PAUP* 4.0b10: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, Mass.

Takhtajan A 1987 Flowering plant origin and dispersal: the cradle of the angiosperms revisited. Pages 26–31 *in* TC Whitmore, ed. Biogeographical evolution of the Malay Archipelago. Clarendon, Oxford.

Tavaré S, CR Marshall, O Will, C Soligo, RD Martin 2002 Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416:726–729.

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

Thorne JL, H Kishino 2002 Divergence time and evolutionary rate estimation with multilocus data. Syst Biol 51:689–702.

van Beusekom-Osinga RJ 1977 Crypteroniaceae. Flora Malesiana 82:198–204.

van Beusekom-Osinga RJ, CF van Beusekom 1975 Delimitation and subdivision of the Crypteroniaceae (Myrtales). Blumea 22:255–266.

van Vliet GJCM, P Baas 1975 Comparative anatomy of the Crypteroniaceae. Blumea 22:175–195.

Wikström N, V Savolainen, MW Chase 2001 Evolution of the angiosperms: calibrating the family tree. Proc R Soc Lond B Biol Sci 268:2211–2220.

Wiley EO 1988 Vicariance biogeography. Annu Rev Ecol Syst 19: 513–542.

Willis JK, JC McElwain 2002 The evolution of plants. Oxford University Press, New York.

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

Yang Z, AD Yoder 2003 Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst Biol 52:1–12.

Zurawski G, B Perrot, W Bottomley, PR Whitfield 1981 The structure of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Res 14:3251–3270.