Phylogenetic patterns and diversification in the caesalpinioid legumes[1]

Anne Bruneau, Marjorie Mercure, Gwilym P. Lewis, and Patrick S. Herendeen

Abstract: Subfamily Caesalpinioideae is a paraphyletic grade of 171 genera that comprises the first branches of the Leguminosae and from which are derived the monophyletic subfamilies Mimosoideae and Papilionoideae. We have sequenced the chloroplast *matK* gene, and the *trnL* and 3′-*trnK* introns for 153 genera of caesalpinioid legumes. Parsimony and Bayesian phylogenetic analyses of these data support the monophyly of several major groups within the caesalpinioid legumes: the Cercideae, Detarieae, Detarieae s. str., *Prioria*, Amherstieae, Dialiinae, *Cassia*, *Caesalpinia*, *Peltophorum*, and *Tachigali* clades. Relationships among the first branching lineages of the legumes are not well supported, with Cercideae, Detarieae, and the genus *Duparquetia* alternatively resolved as sister group to all of the legumes. The division of certain large genera (e.g., *Caesalpinia* s. l., *Bauhinia* s. l.) into segregate genera generally is supported by our molecular data. Using 18 well-documented fossils as calibration points, fixing the stem node of the legumes at 65 Ma, and using the Penalised Likelihood method, we estimate the crown node of the Leguminosae at 64 Ma and the crown age of each of the major caesalpinioid lineages varying from 34 to 56 Ma. Fossil cross-validation suggests that none of the 18 fossil calibrations is internally inconsistent. Analyses done without fossil calibrations yield much younger divergence times. The age estimates for the Detarieae clade are more sensitive to the presence of calibration points than other caesalpinioid clades, a situation which we attribute to the slow rate of chloroplast DNA evolution in this group.

Key words: Leguminosae, *trnL*, *matK/trnK*, phylogenetic analyses, fossil calibration, divergence time estimates.

Résumé : La sous-famille des Caesalpinioideae est un groupe paraphylétique de 171 genres qui comprend les premières lignées des Légumineuses, desquelles sont dérivées les sous-familles monophylétiques des Mimosoideae et des Papilionoideae. Nous avons séquencé les régions chloroplastiques du gène *matK* et des introns *trnL* et 3′-*trnK* pour 153 genres des Caesalpinioideae. Des analyses phylogénétiques de parcimonie et bayesiennes appuient le monophylétisme de plusieurs groupes, dont les tribus des Cercideae et des Detarieae, ainsi que les clades Detarieae s. str., *Prioria*, Amherstieae, Dialiinae, *Cassia*, *Caesalpinia*, *Peltophorum* et *Tachigali*. Les relations demeurent floues entre les premières lignées des Légumineuses, notamment concernant la position des Cercideae, des Detarieae et du genre *Duparquetia*, alternativement résolus comme groupe frère de toute la famille. La ségrégation récente de certains grands genres (ex., *Caesalpinia* s. l., *Bauhinia* s. l.) en de plus petits genres monophylétiques est généralement supportée par les données moléculaires. Des analyses de datations effectuées par la méthode de vraisemblance pénalisée, en utilisant 18 fossiles bien documentés comme points de calibrations et en fixant le nœud basal des Légumineuses à 65 Ma, nous permettent d'estimer un âge de 64 Ma pour la famille et de 34 à 56 Ma pour chacune des principales lignées des Caesalpinioideae. Une validation croisée des fossiles suggère qu'aucun des 18 fossiles n'est incohérent avec les autres. Des analyses effectuées sans ces 18 points de calibrations produisent des âges de divergence beaucoup plus jeunes. Les âges estimés dans le clade des Detarieae sont particulièrement sensibles à la présence ou l'absence des points de calibration, une situation que nous attribuons au faible taux d'évolution du génome chloroplastique au sein de ce groupe.

Mots-clés : Leguminosae, *trnL*, *matK/trnK*, analyses phylogénétiques, calibration par fossiles, temps de divergence.

Introduction

Subfamily Caesalpinioideae is a paraphyletic grade that comprises the first branches of the large and economically important family Leguminosae. The subfamily includes approximately 2250 species in 171 genera, currently divided into four tribes: Cercideae, Detarieae, Cassieae, and Caesalpinieae (Lewis et al. 2005). Of these, only the former two are supported as monophyletic in recent phylogenetic analyses (Bruneau et al. 2001; Herendeen et al. 2003*a*). Lewis et al. (2005), in a synthesis of studies by Polhill and Vidal (1981), Polhill (1994), Lewis and Schrire (1995), Bruneau

Received 11 March 2008. Published on the NRC Research Press Web site at botany.nrc.ca on 18 June 2008.

A. Bruneau[2] and M. Mercure. Institut de Recherche en Biologie Végétale, Université de Montréal, 4101 Sherbrooke est, Montréal, QC H1X 2B2, Canada.
G.P. Lewis. Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB, UK.
P.S. Herendeen.[3] Department of Biological Sciences, George Washington University, Washington, DC 20052, USA.

[1]This paper is one of a selection of papers published in the Special Issue on Systematics Research.
[2]Corresponding author (e-mail: anne.bruneau@umontreal.ca).
[3]Present address: Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, IL 60022, USA.

 doi:10.1139/B08-058

et al. (2000, 2001), Kajita et al. (2001), Gervais and Bruneau (2002), Fougère-Danezan et al. (2003), Herendeen et al. (2003*a*, 2003*b*), Simpson and Lewis (2003), Simpson et al. (2003), Haston et al. (2003, 2005), and Wieringa and Gervais (2003), have reevaluated tribal limits, proposed informal generic groupings, and redefined the generic limits of certain large genera (e.g., *Bauhinia* s.l., *Caesalpinia* s.l.). Although some of these groupings are well supported and consistently resolved in a number of phylogenetic studies, others are tentative arrangements that merit further study with thorough taxon sampling.

Despite numerous recent phylogenetic studies in the monophyletic Leguminosae, relationships among the early branching lineages of the family remain ambiguous and the position of a number of key taxa has yet to be definitively determined. Most analyses (e.g., Bruneau et al. 2001; Kajita et al. 2001; Wojciechowski 2003) suggest that Cercideae is sister to all the Leguminosae, followed by the Detarieae, the Dialiinae (a subtribe of the Cassieae), the Papilionoideae, and a large, poorly resolved clade that includes the *Umtiza* clade (sensu Herendeen et al. 2003*b*), some other Cassieae genera, various Caesalpinieae lineages, and the Mimosoideae. However, support for the resolution among these first branches is consistently low, and other studies variously place the Detarieae clade as the first lineage (e.g., Forest et al. 2002), as sister to the Cercideae (Wojciechowski et al. 2004), or with Cercideae and Detarieae as a trichotomy with the remaining Leguminosae (Käss and Wink 1996). In addition, relationships in the clade that includes the Mimosoideae and most Cassieae and Caesalpinieae subgroups (subtribes or informal generic groups) are poorly resolved in all molecular phylogenetic analyses published to date. Here we provide a phylogenetic analysis based on chloroplast DNA sequences for nearly all genera of the Caesalpinioideae. This allows us to substantiate the tribal assignments and generic groupings proposed by Lewis et al. (2005), and to provide a strongly supported scheme of phylogenetic relationships for future biogeographic and evolutionary studies in these early offshoots of the Leguminosae.

Using fossil evidence and molecular dating methods, Lavin et al. (2005) fixed the stem node of the legumes at 60 Ma and estimated the age of the Leguminosae crown node at 59 Ma. They also noted that all of the oldest clades in the Leguminosae were present from about 39 to 59 Ma. However, their sampling of caesalpinioid taxa was limited, and thus, a critical assessment of the age of the lineages in this paraphyletic grade, from which the other subfamilies are derived, is warranted. Based on such dates, Lavin et al. (2004, 2005) and Schrire et al. (2005*a*, 2005*b*) have suggested that dispersal mechanisms are necessary to account for most continental disjunctions in the Leguminosae. Although this may indeed be the case for most clades in the Leguminosae, it remains to be demonstrated, for example, whether some of the continental disjunctions observed in the caesalpinioid legumes could have occurred earlier, and thus have vicariance explanations. It is also of interest, given the high diversity in floral morphology among lineages of caesalpinioid legumes (e.g., Tucker 2003), to evaluate whether this morphological diversification represents a rapid process or whether these lineages have slowly accumulated changes during the course of their evolution. To properly assess both biogeographical and morphological diversification in the caesalpinioid legumes, the object of future studies, well supported phylogenies are required. Using a number of well-documented fossils, some of which have not previously been used as calibration points in divergence time analyses, we critically evaluate the timing of divergence for major clades in these early branching lineages of the Leguminosae.

Materials and methods

Taxon sampling

Genera from all but one of the subtribes and generic groups in the Caesalpinioideae recognised by Polhill (1994) are included in this study (see supplementary data[4], Table S1). We could not include the monospecific, taxonomically and geographically isolated South American *Orphanodendron* Barneby & J.W. Grimes (tribe Caesalpinieae), recognised by Polhill (1994) as a distinct generic group, because of lack of material. Representative samples from 18 other caesalpinioid genera (eight Detarieae, six Caesalpinieae, three Cassieae, one Cercideae) are not present because of lack of material. In addition, for five genera we only were able to generate a *trnL* sequence because of poor quality DNA (Table S1[4]). A total of 259 sequences are included in the *matK/trnK* analysis, of which 225 are caesalpinioid samples representing 153 genera. For the *trnL* analysis, a total of 376 sequences are included, of which 368 are caesalpinioid sequences representing 157 genera (Table S1[4]). Almost all of the *matK/trnK* sequences were generated in the Bruneau laboratory (33 are from published sequences; Table S1[4]) and several *trnL* sequences are presented here for the first time. For the *trnL* analysis, most genera are represented by more than a single sequence and often several species per genus are included when material was available.

In Mimosoideae, we include representatives of three of the four tribes recognised in Lewis et al. (2005). The monogeneric Mimozygantheae is not sampled, but Fortunato (2005) remarked that, based on recent analyses, this tribe should be disbanded. Our representation of subfamily Papilionoideae is limited to members of tribes Sophoreae and Swartzieae, considered to be sister to the rest of the subfamily in all recent, as well as older treatments (e.g., Polhill 1981; Doyle et al. 1997; Ireland et al. 2000; Pennington et al. 2000, 2001; Kajita et al. 2001; Wojciechowski et al. 2004). As outgroup taxa for rooting our phylogenetic trees, we include species from the Polygalaceae, Surianaceae, and Quillajaceae. All recent molecular studies place these families as part of a Fabales clade in which relationships are poorly resolved among the monophyletic families (e.g., Chase et al. 1993; Käss and Wink 1996; Doyle et al. 1997, 2000; Soltis et al. 1997; Kajita et al. 2001; Wojciechowski et al. 2004; Lavin et al. 2005; Bello et al. 2008).

[4] Supplementary data for this article are available on the journal Web site (http://botany.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Building M-55, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada. DUD 3673. For more information on obtaining material refer to http://cisti-icist.nrc-cnrc.gc.ca/cms/unpub_e.html.

Molecular methods

For most taxa studied, DNA was extracted using the Doyle and Doyle (1987) CTAB protocol, a modification of this protocol described by Joly et al. (2006) or the QIAGEN DNeasy Plant Mini Kit (Mississauga, Ont.), following instructions from the manufacturer.

PCR amplifications in reaction volumes of 50 μL included 4 units of *Taq* polymerase, 1× *Taq* DNA polymerase buffer (Roche Diagnostics, Indianapolis, Ia. final $MgCl_2$ concentration of 1.5 mmol/L), 3.5 mmol/L of additional $MgCl_2$ (Roche Diagnostics, Laval, Que.), 200 μmol/L of each dNTP, 3 mmol/L of each primer, and 50–300 ng of genomic DNA. Bovine serum albumin (2.5–5 μg, New England BioLabs, Ipswich, Mass.) and tween-20 (0.03%–0.05%) were added to samples that were difficult to amplify. Amplifications were conducted using an ABI Gene Amp 9700 thermocycler (Applied Biosystems, Foster City, Calif.) programmed for 35 cycles of 30 s at 94 °C, 1 min at 48 °C and 1 min at 72 °C. The 35 cycles were preceded by a 2 min heating phase (94 °C) before adding the *Taq* polymerase, followed by an additional 1 min at 94 °C, and the cycles ended with 7 min at 72 °C. Product purification involved incubation in one volume of 20% polyethylene glycol with 2.5 mol/L NaCl for 15 min at 37 °C, followed by two washes in 95% and 70% ethanol. Cycle sequencing was performed using the BigDye Terminator chemistry (Applied Biosystems) and following instructions from the manufacturer. Between 15 and 60 ng of PCR products were cycle sequenced. Sequenced products were purified with an ethanol – sodium acetate precipitation and two washes with 70% ethanol. Sequenced products then were run on an ABI 3100-*avant* automated sequencer (Applied Biosystems). Both DNA strands of the loci were sequenced. Sequences were verified, edited, and the two strands were combined using Sequencher (version 3.1–4.7, Gene Codes Corporation, Ann Arbor, Mich.).

The *matK* gene and the flanking 3′ intron region were amplified in one fragment using primers *trnK*685F (Hu et al. 2000; Lavin et al. 2000) and *trnK*2R* (Wojciechowski et al. 2004) for most samples. However, for members of tribe Detarieae, the *trnK*2R* primer often was replaced by a modified version, *trnK*2Rdet (5′-ACACGGCTTTCCCTATGTCTAC-3′). In addition, for more recalcitrant samples, the region was amplified in three fragments, using the following primer pairs: *trnK*685F and KC6 (5′-ATAARCCAATTATCCGAGCAT-3′), *matK*4La (Wojciechowski et al. 2004) and *matK*1932Ra (Wojciechowski et al. 2004), and *matK*1100L (Wojciechowski et al. 2004) and *trnK*2R* or *trnK*2Rdet. These same primers were used for direct sequencing, although we sometimes used additional primers for problematic samples: M4LR (5′-GAAGCGATGACCCACTTTCTAC-3′, M. F. Wojciechowski, unpublished data, 2001), KC4 (5′-TATGTGTCAGATRTACRAATAC-3′), KC5 (5′-CRATTAACATCTTCKGKAGTC-3′), and M1F (5′-CGGGAGTATATTTAGGACTC-3′). The chloroplast *trnL* (UAA) intron was amplified and sequenced as noted in Bruneau et al. (2001).

Phylogenetic analyses

Sequence alignment was initially performed using ClustalX (Thompson et al. 1997), with the default parameters. For the *matK* gene, we translated sequences into amino-acid sequences to guide the alignment. Both the *matK*/*trnK* and *trnL* alignments then were verified and modified manually where inconsistencies were found. To correct the alignment of the noncoding sequences we used the alignment criteria described by Kelchner (2000). Nonautapomorphic gaps were treated as separate presence/absence characters using the simple gap coding strategy of Simmons and Ochoterena (2000) as implemented in GapCoder (Young and Healy 2003). For the *matK*/*trnK* region, we included all of the *matK* gene, as well as most of the 3′-*trnK* intron region except for 141 nucleotides where alignment was ambiguous. In the *trnL* matrix, because of alignment ambiguities a total of 433 nucleotides were removed from the analyses, most of which make up a large and variable AT-rich insertion unique to the Detarieae.

Three matrices were analysed: the *matK*/*trnK* matrix, the *trnL* intron matrix, and the concatenated matrix of these two regions. The concatenated matrix included only species for which we had both *matK*/*trnK* and *trnL* sequences: the matrix included 248 sequences. All three matrices were analysed using parsimony and the concatenated matrix also was analysed using Bayesian methods. For the parsimony analyses, performed using PAUP* (version 4.0b10; Swofford 2001), an initial series of trees was generated by retaining a maximum of five most parsimonious trees per replicate from 1000 random addition replicates with the tree bisection-reconnection (TBR) branch swapping option. This initial set of most parsimonious trees then was used as starting trees for a heuristic analysis, with TBR branch swapping and retaining a maximum of 20 000 trees. Strict consensus trees were constructed for each analysis. Support values were generated by the bootstrap procedure using a heuristic search with TBR branch swapping, retaining a maximum of two trees for each of one sub-replicate for a total of 5000 bootstrap replicates.

The Bayesian analyses were implemented in a parallel version of Mr. Bayes (version 3.1.2; Huelsenbeck and Ronquist 2001) on a shared-memory multiprocessor computer (Altix 4700). A nucleotide substitution model was selected using the Akaike information criterion (AIC) as implemented in Modeltest (version 3.7; Posada and Crandall 1998); this was done for each matrix separately, including for the *trnK* 3′ intron region alone, as well as for the concatenated matrix. The Bayesian estimation consisted of two independent runs, each for 19 250 000 generations, sampling trees and parameters every 1000th generation. Each run consisted of 16 simultaneous Markov chains with eight swaps per generation. We used the $GTR + I + \Gamma$ nucleotide substitution model with the base frequency, substitution rates, and among-site variation variables estimated from the data. The partitions (*matK*, *trnK*, *trnL*) were allowed to evolve independently for the proportion of invariable sites, the shape parameter of the gamma distribution, the nucleotide frequencies and substitution rates ("unlinked"). Indels were coded as for the parsimony analyses and treated using the binary data (restriction site) model with variable coding as suggested by Ronquist et al. (2005). All sample points prior to reaching stationarity of the chains were discarded (equivalent to discarding the first 12 500 trees). Convergence was assessed by comparing majority rule consensus trees from

the two analyses and by using Tracer (version 1.3; Rambaut and Drummond 2007) to compare density plots of the estimated parameters and of the likelihoods from the two analyses. The posterior probabilities for individual clades were compared for congruence and summarized on a majority rule consensus.

Divergence time estimates

Divergence times and nucleotide substitution rates were obtained using age constraints derived from the fossil record. The age of the legume stem node was fixed at 65 Ma for calibration, but analyses also were implemented fixing the stem node at 70 Ma. A total of 18 minimum age fossil constraints were included that ranged from 24 to 60 Ma (Table 1). We chose to place all of these constraints at the stem, rather than crown nodes, in order not to bias the age estimates towards older ages. Certain of these fossil calibration points were used previously (Fougère-Danezan 2005; Lavin et al. 2005); others are used for the first time here.

The presence of a molecular clock was assessed for each partition and for the concatenated dataset using a likelihood ratio test to compare clock- and nonclock-like models on the Bayesian combined tree topology (Felsenstein 1981). This test was implemented in Hy-Phy (version 0.99 beta; Kosakovsky Pond et al. 2005). Because the constancy of nucleotide substitution rates among lineages was rejected in all cases, we used the Penalised Likelihood (PL) method (Sanderson 2002) that allows rate heterogeneity among lineages and the use of multiple fossil calibrations to estimate divergence times. The PL analysis, implemented in r8s (version 1.7; Sanderson 2003), was performed with the truncated Newton algorithm, five initial starts and five perturbed restarts on a fully resolved majority rule consensus tree obtained from the Bayesian analysis of the combined data where branch lengths consisted of the posterior means estimated from the two independent Bayesian runs. Sequence-based cross-validation was used to find an appropriate smoothing value (λ; Sanderson 2002). Smoothing values from 0.001 to 1000 were covered by increments of $10^{0.5}$. We obtained 95% confidence intervals on divergence time estimates by running PL analyses on 100 randomly sampled trees obtained at likelihood stationarity from the Bayesian analysis (posterior predictive method). Divergence times were estimated on all trees using the PL method in r8s and the central 95% of the age distribution for each node was used as a confidence interval. Because of the amount of computational time required to perform cross-validation on 100 large trees, the ideal smoothing value for the consensus tree was used for all trees in the posterior predictive method. Nucleotide substitution rates were calculated for the fully resolved Bayesian consensus, which was used to produce a tree where branch lengths are proportional to the substitution rate (“ratogram”).

A three-step fossil cross-validation procedure was used to assess the concordance or internal consistency of age estimates among fossils and to detect fossils that give ages that deviate significantly from the estimates obtained with the remaining fossils (Near and Sanderson 2004; Near et al. 2005). In this analysis, each fossil constraint was sequentially used as a single fixed constraint and the divergence times were estimated for the remaining nodes for which a fossil was available. The deviation between the estimated molecular ages and the fossil ages were compared for each fossil constraint, and an average squared deviation was calculated to rank the fossil calibrations in terms of their impact on the age estimates. The significance of change in variance of the average squared deviation before and after fossil calibration removal was determined using a one-tailed *F*-statistics as suggested by Near et al. (2005).

Results

Sequence characteristics

Sequence characteristics, lengths, number of indels, and number of variable characters are given in Table 2 for each of the regions studied. The *matK* gene was generally 1515–1530 bp in length (505–510 amino acids) in the Polygalaceae, Quillajaceae, Surianaceae, Cercideae, Detarieae, Dialiinae, *Duparquetia*, and Papilionoideae. In the Mimosoideae and other Cassieae and Caesalpinieae lineages, the *matK* gene was generally nine amino acids shorter (ca. 1500 bp in length). The *trnL* intron was 550 bp in length in most legumes and related families, but approximately 1180 bp in length in all of the Detarieae. A total of 1251 (43%) parsimony informative characters were obtained from the combined analysis, with a slightly higher proportion from the *matK/trnK* region (48%) than from the *trnL* intron (39%).

Phylogenetic analyses

All parsimony analyses resulted in over 20 000 trees, the maximum retained in memory. The consensus trees of the *matK*/*trnK* (not shown) and combined *matK*/*trnK* and *trnL* analyses (not shown) were better resolved than that of the *trnL* analysis alone. This latter topology is presented (see supplementary data[4], Fig. S1) because it includes a greater number of species and therefore indicates the position of taxa that could not be included in the analyses (both parsimony and Bayesian) of the combined data set. Although relationships among the major clades are very similar in the three parsimony analyses, we obtained better resolution and higher bootstrap values in the combined data analysis. Bootstrap supports for the parsimony analysis of the combined data are given in Figs. 1–4 on the Bayesian majority rule tree (see below). Statistics for all these analyses are given in Table 2.

In all of the analyses, the same major clades are resolved as monophyletic: Cercideae, Detarieae, Dialiinae, Papilionoideae, and a clade that includes all of the remaining legumes except the genus *Duparquetia*. This latter clade includes the monophyletic subfamily Mimosoideae, along with the majority of Caesalpinieae and Cassieae genera. The major differences among the three parsimony analyses are in the position of the first branching clades of the legumes. Although unresolved in the *trnL* strict consensus tree (Fig. S1[4]), individual *trnL* trees resolve these lineages with either *Duparquetia* and Detarieae together, or with Detarieae alone, as sister to all other Leguminosae. In the strict consensus of the *matK*/*trnK* analysis, relationships among these clades also are not resolved, but the individual trees place either *Duparquetia* or Cercideae as sister to all other Leguminosae. The combined *matK*/*trnK* + *trnL* parsimony analysis consistently resolves the relationship with Cercideae

Table 1. Fossils used as fixed (A) or minimum age (B–S) constraints in the divergence time analyses of the caesalpinioid legumes.

Node	Taxonomic identity and phylogenetic position	Age	Fossil organ, description and locality	References
A (Fig. 1)	Leguminosae fixed stem node	65 Ma	Previous molecular divergence time analyses	Magallón and Sanderson (2001); Lavin et al. (2005)
B (Fig. 1)	Crown node of the Leguminosae	60 Ma	Wood with vestured pits; Mali	Crawley (1988)
C (Fig. 1)	*Cercis*	34 Ma	Leaves and fruits from Sheep Rock Creek, Oregon, USA.; leaves from Florissant locality, USA	MacGinitie (1953); Herendeen and Manchester (2004),
D (Fig. 1)	*Bauhinia* s.l.	46 Ma	Lobed leaf with pulvinus at base of lamina, fan-like main veins radiating from the base; Tanzania[a]	Jacobs and Herendeen (2004, unpublished data)
E (Fig. 2)	*Hymenaea*	24 Ma	Flowers in amber; Dominican Republic	Hueber and Langenheim (1986)
F (Fig. 2)	MRCA *Prioria* and *Oxystigma*	24 Ma	*Prioria*-like flowers in amber; Dominican Republic[b]	Poinar and Poinar (1999)
G (Fig. 2)	*Daniellia* and *Brandzeia* clade	53 Ma	*Daniellia* wood in amber; Paris Basin, France	De Franceschi and De Ploëg (2003)
H (Fig. 3)	*Aphanocalyx*	46 Ma	Leaves; Tanzania	Herendeen and Jacobs (2000)
I (Fig. 3)	*Crudia*	45 Ma	Fruit and leaflets; SE USA[c]	Herendeen and Dilcher (1990)
J (Fig. 1)	Papilionoideae	55 Ma	*Barnebyanthus buchananensis* flowers; SE USA and Wyoming, USA	Crepet and Herendeen (1992); Herendeen and Wing (2001)
K (Fig. 1)	*Swartzia*	45 Ma	Fruits and leaflets; SE USA	Herendeen (1992)
L (Fig. 4)	*Arcoa*	34 Ma	*Prosopis linearifolia* leaves; Florissant locality, USA	MacGinitie (1953); Lavin et al. (2005)
M (Fig. 4)	*Acrocarpus*	45 Ma	Dehiscent fruit with wings along both sutures; SE USA	Herendeen (1992)
N (Fig. 4)	*Senna*	45 Ma	Fruits; SE USA and Mexico	Herendeen (1992); Calvillo-Canadell and Cevallos-Ferriz (2005)
O (Fig. 4)	*Mezoneuron*	45 Ma	Fruits; several sites SE and W USA	Herendeen and Dilcher (1991)
P (Fig. 4)	*Caesalpinia* s.s.	45 Ma	Fruits and leaflets; SE USA	Herendeen (1992)
Q (Fig. 4)	*Acacia*	46 Ma	Leaf; Tanzania[d]	Herendeen and Jacobs (2000)
R (Fig. 4)	*Dinizia*	45 Ma	*Eomimosoidea plumosa* flowers, leaves, and fruits; SE USA[e]	Crepet and Dilcher (1977); Herendeen and Dilcher (1990)
S (Fig. 4)	Ingeae	45 Ma	*Albizia* polyads; Egypt[f]	Guinet et al. (1987)

Note: All fossils were placed at the stem node of clades unless otherwise noted. The positions of these fossils in our phylogenetic analyses are noted in Figs. 1–4. MRCA, most recent common ancestor.

[a]Younger fossils with affinities to *Bauhinia* have been described (Calvillo-Canadell and Cevallos-Ferriz 2002; Chen and Zhang 2005) but are not used here.

[b]The flowers described by Poinar and Poinar (1999) are pedicellate, have 2 bracteoles at the base of the calyx, 4 or 5 sepals (unclear), they have 10 stamens, and appear apetalous. Although similar to *Prioria*, the flowers of present day *Prioria* are sessile and the genus *Prioria* is very similar to *Oxystigma* in terms of flower morphology (Breteler 1999), making it difficult to ascribe the fossil with certainty to one of the *Prioria* s. s. clade genera (sensu Fougère-Danezan et al. 2007). Therefore, as a conservative minimum age estimate, we placed this fossil constraint at the stem node of the entire *Prioria* s. s. clade. We use the same age as used for the *Hymenaea* constraint (see text).

[c]In addition, *Crudia* pollen is known from the Paleocene and younger sediments (Adegoke et al. 1978; Muller 1981; Graham 1985, 1989, 1992).

[d]Herendeen and Jacobs (2000) consider that this fossil may belong to the African clade of the polyphyletic *Acacia* s.l. (to which belongs the only *Acacia* sampled here, *Acacia caven*). In addition, polyads of *Acacia* have been reported from Eocene and younger sediments from Germany, Australia, New Zealand, Egypt, and Cameroon (Graham and Jarzen 1969; Mildenhall 1972; Guinet and Lugardon 1976; Salard-Cheboldaeff 1978; Muller 1981; Guinet 1981, 1990; Guinet et al. 1987; Graham 1992).

[e]The taxon was suggested to be most similar to *Dinizia*, which until recently was considered part of the Mimosoideae clade, but now is grouping with the *Dimorphandra* clade.

[f]Although not sampled here, based on more thoroughly sampled phylogenetic analyses of the Mimosoideae (Luckow et al. 2003), *Albizia* would occur in the clade defined as the most recent common ancestor of *Archidendron hirsutum* and *Zapoteca tetragona*. Calvillo-Canadell and Cevallos-Ferriz (2005) also describe a lower Oligocene leaflet from Puebla, Mexico as belonging to the genus *Inga*.

as sister to all other legumes, and with *Duparquetia* and Detarieae together as sister to the remaining Leguminosae.

Although the sampling for large genera is far from adequate, the *trnL* analysis (Fig. S1[4]) allows us to test, at least in part, the monophyly of certain of these genera. For the *Bauhinia* s.l. (*Barklya*, *Phanera*, *Lasiobema*, *Lysiphyllum*, *Tylosema*, *Gigasiphon*, *Bauhinia* s. str.), and *Caesalpinia* s.l. (*Coulteria*, *Erythrostemon*, *Guilandina*, *Libidibia*, *Mezoneuron*, *Poincianella*, *Tara*, *Caesalpinia* s. str.) segregate genera for which we were able to sequence multiple species, all are supported as monophyletic or do not contradict their monophyly (the analysis simply lacks resolution). Of the other genera for which we had multiple species, all are supported as monophyletic or paraphyletic (lack of resolution), except eight genera of the Detarieae, *Hymenostegia*, *Anthonotha*, *Macrolobium*, *Paloue*, *Copaifera*, *Berlinia*, *Plagiosiphon*, and *Tetraberlinia*.

The Bayesian majority rule tree (Figs. 1–4) for the combined matrix yielded a topology similar to that obtained from the parsimony analysis of the combined data. As with

Table 2. Sequence characteristics, parsimony and Bayesian analysis statistics, choice of evolutionary model and deviation from clock-like evolution for each of the matrices and subsets of the matrices analysed in a phylogenetic analysis of the caesalpinioid legumes.

	matK gene	3′-*trnK region*	Combined *matK*/*trnK* region	*trnL* intron	Combined *matK*/*trnK* and *trnL*
Number of sequences	259	259	259	376	248
Aligned length (bp)	1707	413	2120	1252	3372
Indels	25	2	27	55	82
Excluded characters	0	141 (34.1%)	141 (6.7%)	433 (34.6%)	574 (17.0%)
Variable characters	1198 (67.8%)	174 (63.5%)	1372 (67.3%)	536 (61.3%)	1853 (63.6%)
Parsimony informative characters	834 (47.2%)	138 (50.4%)	972 (47.7%)	342 (39.1%)	1251 (42.9%)
CI′, RI, length	CI′ = 0.37, RI = 0.86, L = 4375	CI′ = 0.42, RI = 0.85, L = 778	CI′ = 0.38, RI = 0.86, L = 5213	CI′ = 0.43, RI = 0.91, L = 1591	CI′ = 0.39, RI = 0.86, L = 6443
% GC content	28.9%	35.3%	29.6%	31.2%	30.4%
Evolutionary model, AIC	GTR+I+Γ	TVM+Γ	TVM+I+Γ	GTR+I+Γ	TVM+I+Γ
LHR test, clock-like evolution	Rejected*, χ^2 = 19587.1	Rejected*, χ^2 = 3873.67	Rejected*, χ^2 = 21434.1	Rejected*, χ^2 = 10273.8	Rejected*, χ^2 = 30191.6

Note: AIC, Akaike information criterion; LHR, likelihood ratio test of clock-like evolution.
*Rejected at $p < 0.001$.

the parsimony analyses, in the Bayesian analysis relationships among the first branching lineages of the legumes also are not well supported. The Detarieae clade is resolved as sister to a poorly supported clade (0.60 posterior probability (PP)) that includes *Duparquetia*, the Cercideae clade, and the clade comprising all of the remaining legumes. Although there is slightly more resolution in the majority rule Bayesian tree (and higher clade support values) than in the strict consensus of the parsimony analysis for the combined data, ensuing differences among the other legume clades are relatively minor (see Figs. 1–4).

Divergence time estimates

The likelihood ratio tests performed on individual datasets and on the combined matrix found significant rate heterogeneity among lineages and justified the use of the PL method for dating nodes (Table 2). A smoothing value parameter of $\lambda = 3.2$ produced the lowest error for the consensus tree from the Bayesian analysis in the sequence-based cross-validation procedure implemented in r8s. The mean age of divergence of the different nodes of interest are shown in Table 3, along with 95% confidence intervals obtained using the posterior predictive method. For five of eight nodes where the age obtained from the Bayesian consensus tree topology fell outside the confidence interval established with random Bayesian trees, the posterior probability of the clade was lower than one. The age of 12 of the constrained nodes reached their minimum in the PL analysis on the consensus Bayesian tree (noted with an asterisk in Table 3). Of these 12 nodes, 7 were those with the greatest difference between their ages estimated with and without constraints (nodes E, G, H, I, O, Q, S). Removing all constraints decreased the age of the nodes of interest by 7.2% to 90.4%; the age difference was particularly acute for constrained node H (Table 3). The fossil cross-validation method of Near and Sanderson (2004) suggested that none of the 18 fossil constraints was significantly inconsistent, as indicated by the change in variance of the average squared deviation between the molecular and fossil age estimate.

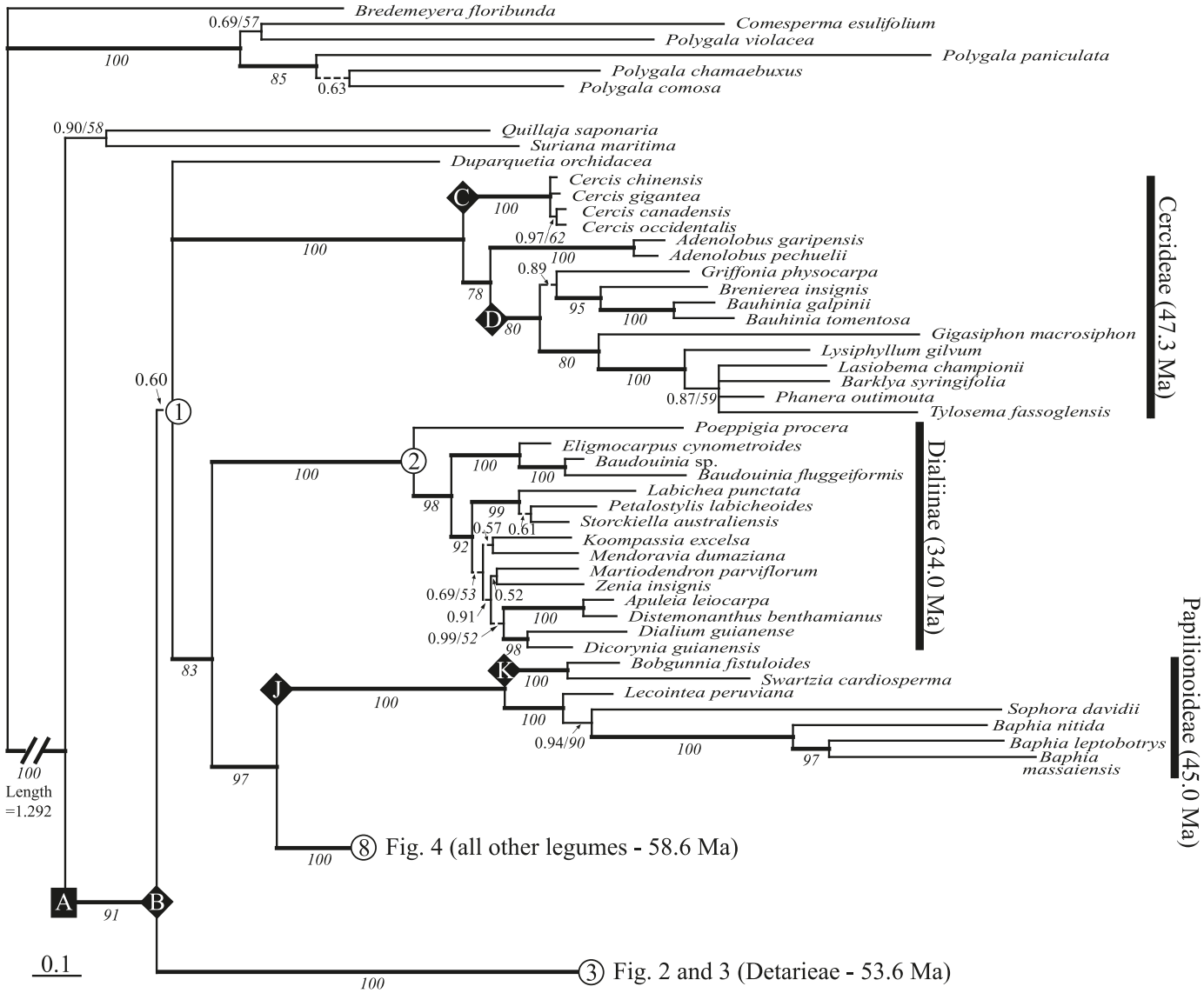
Fixing the stem node of the Legumes at 65 Ma resulted in a crown node age for the legume family at 64 Ma (Table 3). We also fixed the age of the stem node at 70 Ma to evaluate the effect of different stem node ages on our divergence time estimates for major lineages within the legumes. Although fixing the stem node at 70 Ma yielded an age estimate of 68.6 Ma for the Leguminosae, there is relatively little difference in the ages estimated for the other major lineages within the family (data not shown). With a legume stem node fixed at 65 Ma, the age of the major caesalpinioid clades were estimated at 34.0 to 56.5 Ma (Table 3; Fig. 5).

Discussion

Phylogenetic relationships and divergence time estimates

Despite slight differences obtained among the consensus trees of the parsimony and Bayesian analyses, a strongly supported pattern of relationships among legume lineages and genera emerges from analyses of the chloroplast *trnL* and *mat*/*trnK* regions. Relative to previously published *trnL* analyses for caesalpinioid legumes (e.g., Bruneau et al.

Figs. 1–4. Phylogenetic analysis of the chloroplast *mat*/3′-*trnK* region and *trnL* intron for caesalpinioid legumes. Bayesian majority rule consensus derived from 13 500 trees retained after reaching likelihood stationarity in two independent analyses. Posterior probabilities (mostly above branches) and bootstrap support values from the parsimony analysis (mostly below branches, in italics) of the same data are noted. Branches in bold are those supported by a 1.0 posterior probability; clades that are unresolved in the parsimony analysis are indicated by broken lines. Nodes labeled with letters are fixed calibration points (square, node A) or minimum age calibration points (diamond shape, nodes B–S), whereas numbered nodes represent those whose ages were estimated. Black lines to the right indicate clade names; grey lines indicate poorly supported clades. **Fig. 1.** Portion of the consensus tree showing relationships among the first diverging branches of the Leguminosae.

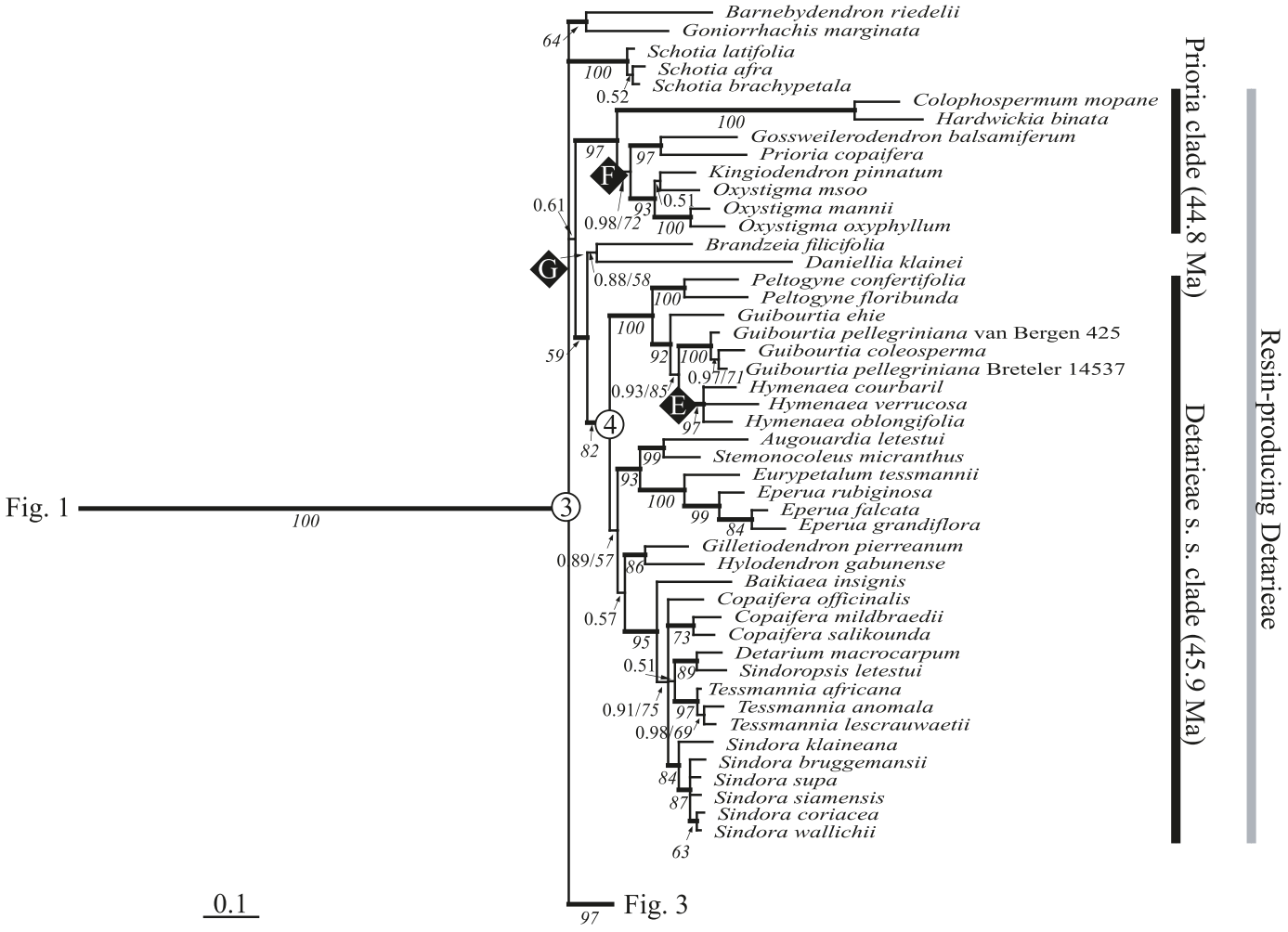


2001), the combination with *matK* data leads to better resolved relationships, even among closely related genera. Wojciechowski et al. (2004) emphasized the utility of using *matK* for examining relationships within the Leguminosae, but this had yet to be clearly demonstrated for caesalpinioid legumes. Although *matK*/*trnK* has certainly improved our understanding of phylogenetic relationships among these legumes, certain relationships remain weakly supported (e.g., relationships among the first diverging lineages of the legumes or among genera in the Detarieae). It is likely that independent, perhaps faster evolving, nuclear genes (e.g., Choi et al. 2006) will need to be examined and that a better integration of morphological and molecular data (e.g., Herendeen et al. 2003*a*) will be necessary to confidently assess relationships among these lineages and caesalpinioid genera.

Relationships at the base of the legume tree

Relationships among the first branches of the legume phylogeny are not well resolved. In the parsimony analysis of the combined *mat*/*trnK* and *trnL* data, Cercideae is sister to all other legumes, whereas in the Bayesian analysis, the Detarieae clade is sister to all remaining legumes. Although most recent taxonomic and phylogenetic evaluations of the

Fig. 2. Portion of the consensus tree showing relationships among the resin-producing Detarieae.



family consider Cercideae to be sister to all of the Leguminosae (e.g., Bruneau et al. 2001; Herendeen et al. 2003*a*; Wojciechowski 2003; Lewis and Forest 2005), the relative position of these early diverging lineages always has been ambiguous or poorly supported in phylogenetic analyses and depends on taxon sampling, locus sequenced and method of phylogenetic analysis (cf., Doyle et al. 2000; Bruneau et al. 2001; Kajita et al. 2001; Wojciechowski et al. 2004; Bello et al. 2008). Although each of the lineages is strongly supported as monophyletic, relatively few molecular or morphological (e.g., Herendeen et al. 2003*a*) characters support relationships among these clades.

The position of the genus *Duparquetia* is particularly problematic. In the combined parsimony analysis, *Duparquetia* always occurs as sister to the Detarieae, a relationship that is not supported by bootstrap values over 50%, and in the fully resolved majority rule Bayesian consensus tree *Duparquetia* occurs as sister to the Cercideae (Fig. 5), but only with very low clade support (Fig. 1). The monospecific West African genus *Duparquetia* was placed in its own subtribe Duparquetiinae (tribe Cassieae) by Irwin and Barneby (1981). Phylogenetic analyses of the Leguminosae based on morphological characters (Chappill 1995), suggested that *Duparquetia* occurred as a trichotomy with the Papilionoideae and a clade that includes all other Caesalpinioideae and Mimosoideae. *Duparquetia orchidaea* Baill. is highly unusual in its floral (Klitgaard et al. 2002; G. Prenner and B.B. Klitgaard, Royal Botanic Gardens, Kew, unpublished data, 2008) and pollen morphology (Banks et al. 2003) and it lacks vestured pits, a characteristic anatomical feature of the wood of all other Leguminosae except for species of tribe Cercideae and of most of the Dialiinae clade (Herendeen et al. 2003*a*). The uniqueness of this genus is thus reflected both in the long branch leading to *Duparquetia* in chloroplast DNA phylogenetic analyses (this study for *matK/trnK* and *trnL*; Bello et al. 2008 for *rbcL*) and in its morphology.

Our divergence time analyses, with a fixed legume stem node at 65 Ma, suggest an Early Paleocene crown node age for the Leguminosae (64.0 Ma; Table 3; Fig. 5). Lavin et al. (2005) fixed the stem node of the legumes at 1 Ma intervals between 60 and 70 Ma, and estimated the age of the legumes to vary from 59 to 64 Ma. Similarly, Magallón and Sanderson (2001) estimated the age of the legume crown group based on the first reliable fossil appearance for the family from Herendeen and Crane (1992) at 60 Ma. The somewhat older estimate for the age of the family seen here is a consequence, in part, of our positioning of the 60 Ma minimum age constraint based on fossil wood with vestured pits at the crown node of the legumes (node B, Fig. 1;

Fig. 3. Portion of the consensus tree showing relationships within the Amherstieae clade of the Detarieae.

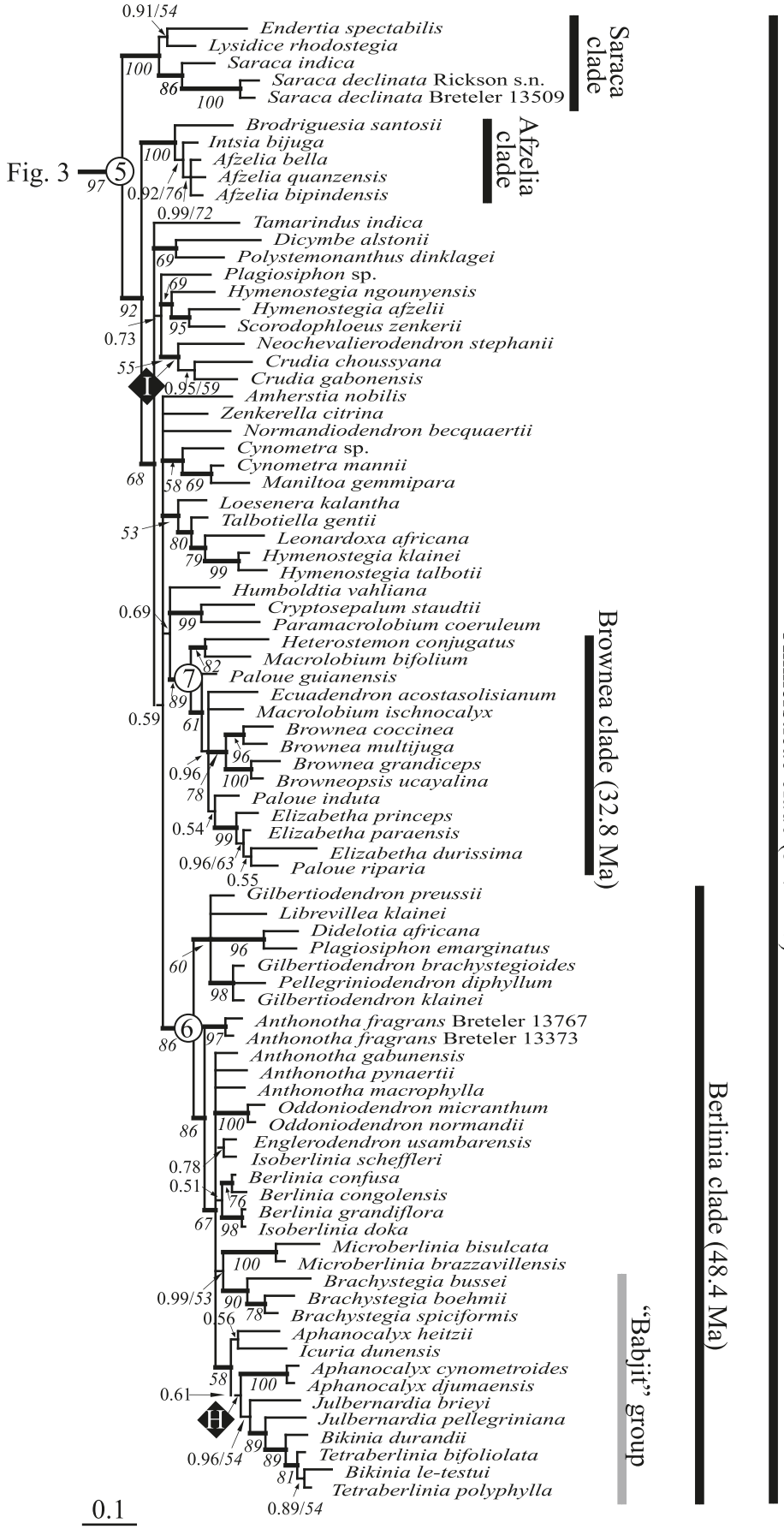


Fig. 4. Portion of the consensus tree showing relationships within a clade comprised of several Cassieae and Caesalpinieae clades and the Mimosoideae.

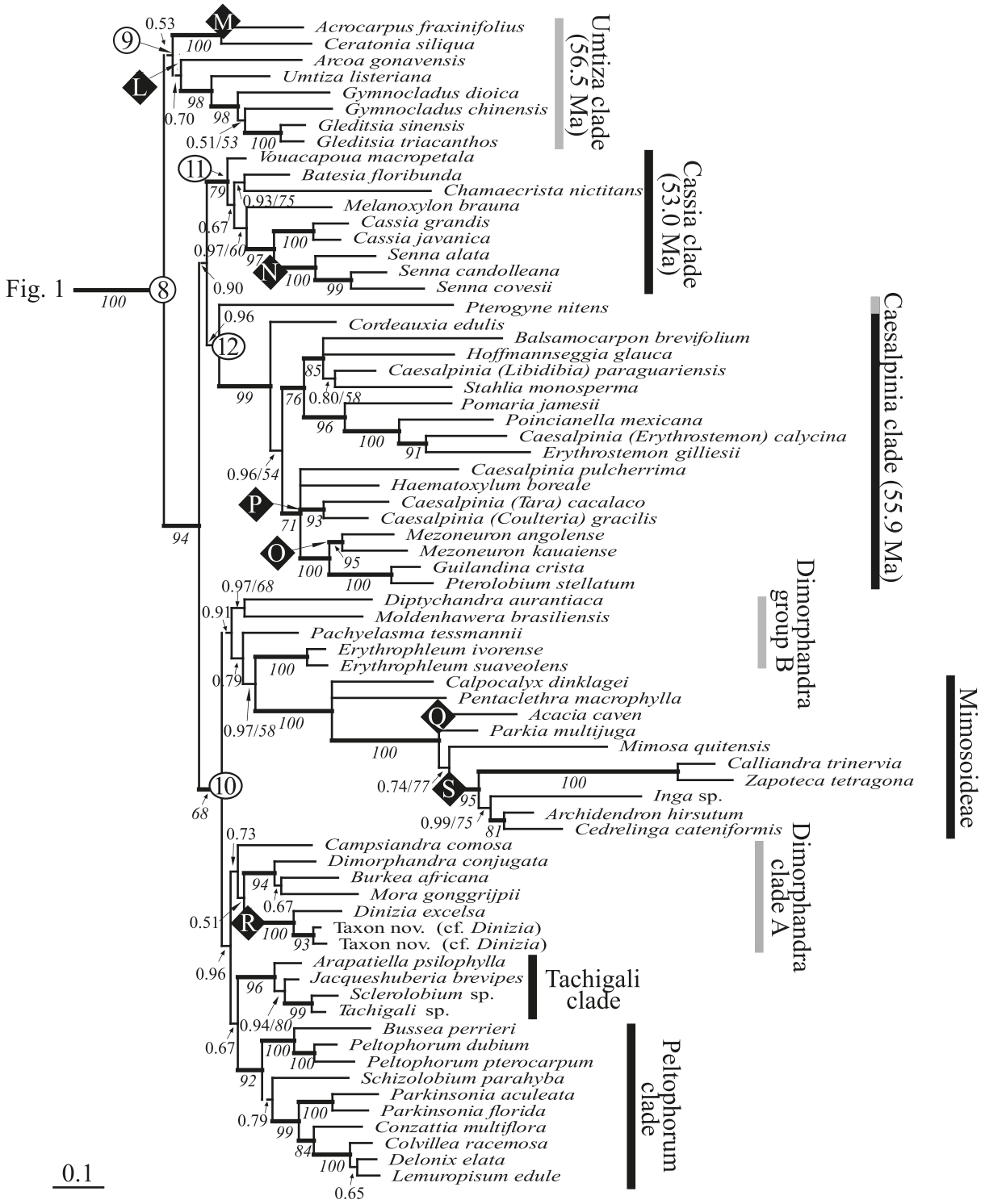


Table 1), which leads to methodological contradictions if the stem node is fixed at 60 Ma. Fossil caesalpinioid wood with well-documented vestured pits has been described by Crawley (1988) from the middle to upper Palaeocene (~60 Ma) of Mali. Vestured pits are an anatomical feature of the wood of most Leguminosae, but they are absent in the Cercideae, in most Dialiinae clade genera (all except *Poeppigia* and *Mendoravia*) and in *Duparquetia* (Herendeen 2000). Because of ambiguity in the optimization of the vestured pit character on our phylogeny, we assigned this fossil as a minimum constraint to the crown node of the Leguminosae (Fig. 1). Other late Paleocene and early Eocene fossil constraints (e.g., nodes G and J) placed at relatively deep nodes in the Leguminosae also contribute to our early Pale-

Table 3. Divergence time estimates of constrained (lettered) and unconstrained (numbered) nodes as determined by Penalised Likelihood analyses of caesalpinioid legumes.

Node	Posterior probability on consensus Bayesian tree	Constraint age Ma	Divergence time point estimate on consensus Bayesian tree (Ma)	Divergence time estimate from 100 random Bayesian trees			Divergence time estimate without minimum age constraints enforced (Ma)
				Mean (Ma)	SD	95% confidence interval	
A	1.00	65 (fixed)	—	—	—	—	—
B	1.00	60 (min.)	64.0	64.0	0.243	63.9–64.0	59.4
C	1.00	34 (min.)	47.3	47.3	0.424	47.2–47.4	28.1
D*	1.00	46 (min.)	46.0	46.0	0	46.0–46.0	26.1
E*	0.93	24 (min.)	24.0	24.7	1.369	24.4–25.0	5.9
F	1.00	24 (min.)	44.8	44.7	2.167	44.3–45.2	14.1
G*	1.00	53 (min.)	53.0	53.0	0	53.0–53.0	15.3
H*	0.61	46 (min.)	46.0	46.0	0	46.0–46.0	4.4
I*	1.00	45 (min.)	45.0	45.0	0	45.0–45.0	6.9
J	1.00	55 (min.)	61.2	61.3	0.455	61.2–61.4	49.9
K*	1.00	45 (min.)	45.0	45.0	0	45.0–45.0	32.5
L	0.70	34 (min.)	52.4	55.2	3.044	54.6–55.8	28.4
M*	1.00	45 (min.)	45.0	45.0	0.321	45.0–45.0	18.9
N*	1.00	45 (min.)	45.0	45.0	0	45.0–45.0	15.9
O*	1.00	45 (min.)	45.0	45.0	0	45.0–45.0	12.9
P	1.00	45 (min.)	48.3	47.8	1.457	47.5–48.1	18.5
Q*	1.00	46 (min.)	46.0	46.0	0.141	46.0–46.1	13.4
R*	0.51	45 (min.)	45.0	48.8	2.689	48.3–49.3	20.5
S*	0.74	45 (min.)	45.0	45.4	0.682	45.3–45.5	12.7
1	0.60	—	63.7	63.7	0.285	63.7–63.8	58.2
2	1.00	—	34.0	34.0	2.334	33.5–34.4	28.8
3	1.00	—	53.6	53.8	0.411	53.7–53.9	17.3
4	1.00	—	46.1	46.0	2.652	45.5–46.6	12.7
5	1.00	—	52.1	52.1	0.575	52.0–52.2	12.7
6	1.00	—	48.4	48.5	0.673	48.3–48.6	7.0
7	1.00	—	32.8	37.0	3.878	36.2–37.8	5.7
8	1.00	—	58.6	58.6	0.643	58.5–58.8	35.0
9	0.53	—	56.5	57.6	1.391	57.3–57.9	31.1
10	1.00	—	56.3	56.3	0.888	56.1–56.5	27.6
11	1.00	—	53.0	53.2	1.567	52.9–53.5	22.7
12	0.96	—	55.9	56.0	0.945	55.8–56.2	28.7

Note: Summary statistics for these nodes are provided based on 100 trees obtained by random sampling at likelihood stationarity from the Bayesian analysis of the combined data. The position of each of the nodes is indicated on Figs. 1–4. *, constrained nodes whose age estimates in the Penalised Likelihood analyses on the Bayesian consensus tree were the same as their minimum age.

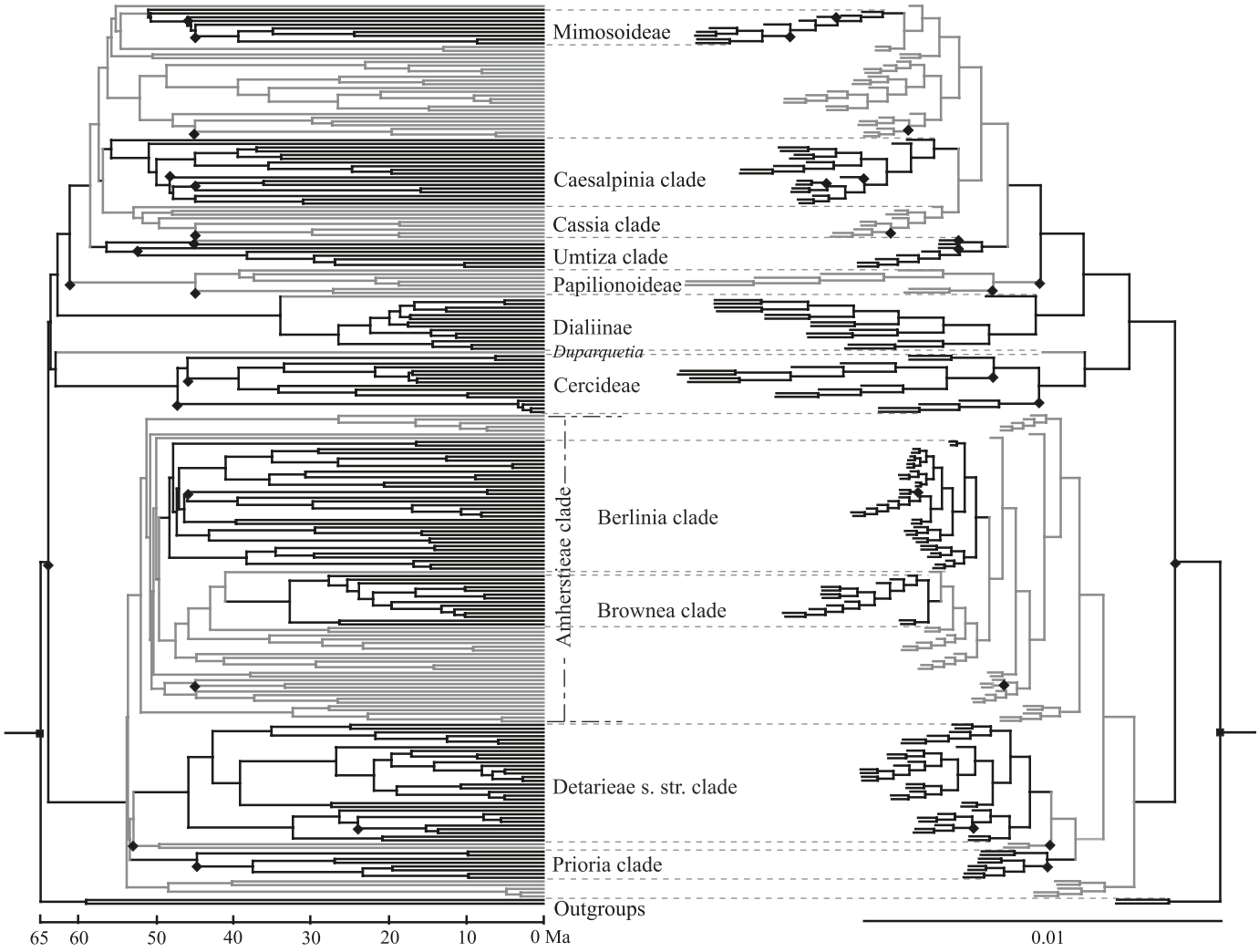
ocene estimate for the crown node of the legumes. Given the relatively poor fossil data available from the late Cretaceous and early Tertiary from low latitudes (Herendeen et al. 1992), fixing the age of the root may not be strongly justified assumption. However, nonclock methods calculate age estimates with difficulty unless at least one fixed (deep) node is included, regardless of the number of minimum age constraints (e.g., Sanderson 2002, 2004). The ambiguity in relationships among the first diverging lineages of the legumes (Detarieae, Cercideae, *Duparquetia*) also adds uncertainty to any age estimate for the family.

Cercideae clade

The monophyly of tribe Cercideae is well supported in all morphological (Chappill 1995; Herendeen et al. 2003*a*) and molecular (e.g., Käss and Wink 1996; Doyle et al. 1997, 2000; Bruneau et al. 2001) phylogenetic studies, as it is here (Fig. 1). Members of this tribe share a number of unique vegetative and floral morphological features that support its monophyly (e.g., Wunderlin et al. 1981). However, relationships among genera within the tribe have been more problematic, especially relative to the large and complex pantropical *Bauhinia* s.l.

Our analyses do not support the division proposed by Wunderlin et al. (1981), who recognised two subtribes within the Cercideae: the Cercidinae (*Cercis*, *Griffonia*, *Adenolobus*) and the Bauhiniinae (*Brenieria* and *Bauhinia* s.l.). Instead, our analyses place *Cercis* and *Adenolobus* as a grade, and *Griffonia* as sister to *Brenieria* and *Bauhinia* s. str. (see Fig. 1). As in our previous molecular analyses (Bruneau et al. 2001), the results presented here also bring into question the monophyly of the genus *Bauhinia* s.l. Rather, our analyses support the recent taxonomic treatment by Lewis and Forest (2005) who, based on preliminary molecular analyses and previous taxonomic treatments, recognise the genera *Barklya*, *Phanera*, *Lasiobema*, *Lysiphyllum*, *Tylosema*, and *Piliostigma* as distinct from *Bauhinia* s. str. Furthermore, all of the segregate genera for which we were

Fig. 5. Fossil-constrained divergence time and nucleotide substitution rate estimates for the caesalpinioid legumes as determined from the Penalised Likelihood method on a fully resolved Bayesian consensus tree of the combined *matK*/*trnK* and *trnL* data. Left side: divergence time estimates (chronogram). Right side: rates of nucleotide substitutions (ratogram).



able to sample multiple species (Figs. 1 and S1[4]) are supported as monophyletic (*Bauhinia* s. str., *Phanera*, *Lysiphyllum*). Studies are ongoing to more completely sample these large and complex genera and to test the monophyly of the other non-monospecific segregate genera (Sinou et al. 2007; F. Forest, Royal Botanic Gardens, Kew, unpublished data, 2005).

We estimate a mid-Eocene (47.3 Ma) age for the crown node of the Cercideae, an estimate that is relatively older than that obtained by Lavin et al. (2005; 34 Ma). The inclusion of the mid Eocene *Bauhinia* s.l. leaf fossil from Tanzania (Jacobs and Herendeen 2004) in our analysis (node D, Table 1) invariably leads to this older age estimate. Regardless, the ages estimated here are much younger than those suggested by Polhill et al. (1981), who recognised Cercideae as one of the first diverging lineages of the legumes, but with some relic genera dating back to 60 Ma. This has yet to be corroborated from the fossil record. More recently based on biogeographical analyses of the Leguminosae as a whole, Schrire et al. (2005*a*, 2005*b*) noted that the distribution of taxa in this clade, with north temperate taxa sister to succulent (deciduous dry forest), grass, and rain forest species, may be the consequence of widespread ancestral taxa that were present in the boreotropics (surrounding the Tethys seaway) and which subsequently dispersed into the Southern Hemisphere via dry succulent biome corridors to speciate in the grass and rain forest biomes.

Dialiinae clade

The Dialiinae clade, here defined to include subtribes Dialiinae and Labicheinae of the Cassieae as recognised by Irwin and Barneby (1981), together with the genus *Poeppigia*, is strongly supported as sister to the Papilionoideae plus the clade that includes the Mimosoideae, and most Caesalpinieae and Cassieae lineages (Fig. 1). The relationships are similar to those presented by Bruneau et al. (2001) based on *trnL* analyses, but the addition of *matK*/*trnK* data yields higher resolution, and additional genera have been sequenced since then.

As found in previous analyses (Bruneau et al. 2001), the monospecific neotropical genus *Poeppigia*, which was placed in its own generic group of the Caesalpinieae by Pol-

hill and Vidal (1981), is sister to this entire clade. *Poeppigia* has vestured pits in its wood, whereas they are absent in the rest of the Dialiinae clade (except for *Mendoravia*), suggesting a labile pattern of evolution of this character among the early diverging lineages of the legumes. Despite its unique combination of morphological characters, including pseudopapilionoid flowers, *Poeppigia* has axillary or terminal, cymose, paniculate inflorescences; a character that likely is a synapomorphy for the Dialiinae clade. The genus also has a narrowly winged fruit; a structure that occurs in a number of Dialiinae clade genera (Herendeen et al. 2003*a*). Two of the three Malagasy genera of this clade, *Baudounia* and *Eligmocarpus* (newly sequenced here), form a strongly supported sister group, sister to all other Dialiinae. The two Australian genera, *Petalostylis* and *Labichea*, plus *Storckiella* with species in Australia, Fiji, and New Caledonia, also always form a strongly supported clade, which occurs as sister to the remaining genera in the combined and *mat/trnK* alone analyses. The monospecific west African *Distemonanthus* and the South American *Apuleia* form a strongly supported sister pair, as do the pantropical *Dialium* and neotropical *Dicorynia*, but the relationships among these two clades and the other genera in the clade are not well resolved. The large and heterogenous genus, *Dialium*, is supported as monophyletic, albeit with limited species sampling (but see Fig. S1[4]). The relationship between the monospecific Chinese genus *Zenia* and neotropical *Martiodendron*, and between the Malagasy *Mendoravia* and the Southeast Asian *Koompassia* are both weakly supported.

Among the caesalpinioid lineages, the Dialiinae clade was estimated to be the youngest (34 Ma; Figs. 1 and 5). This is the only main caesalpinioid lineage for which no clear fossil evidence is available to calibrate the divergence time estimates. Calvillo-Canadell and Cevallos-Ferriz (2005) described a fossil fruit from Oligocene deposits (30 Ma) from Mexico that they attributed to the genus *Apuleia*. This fossil was not used as a minimum age constraint in our analysis because of the general characteristics of the fossil fruit, and the high variation in fruit morphology in the extant species of the South American *Apuleia*, which do not easily allow assignment to a particular genus. Regardless, if correctly identified, the age of this fossil is not inconsistent with the estimated crown node age for the Dialiinae. This relatively young clade includes several small genera that occur on different continents, and, for example, suggests at least two independent New World – Africa or New World – Madagascar dispersal events that merit closer study.

Detarieae clade

The tribe Detarieae, as circumscribed by Mackinder (2005), is strongly supported as monophyletic in our analyses (Fig. 2) and the combination of *matK*/*trnK* and *trnL* data has resulted in better-supported hypotheses of relationships within this clade, relative to previous studies (Bruneau et al. 2000, 2001). The Detarieae comprises mostly African genera, but also includes New World and Asian taxa, with several continental disjuncts between close generic pairs and a few within genera. Morphologically, this tribe is highly diverse and classification systems have been modified numerous times over the past decades to accommodate accumulating information on this group (e.g., Léonard 1957; Cowan and Polhill 1981*a*, 1981*b*; Breteler 1995; Wieringa 1999; Mackinder 2005). The broad pattern that emerges here, is one in which the weakly supported monophyletic resin-producing Detarieae (sensu Fougère-Danezan et al. 2007), which includes the *Prioria* and the Detarieae s. str. clades, are sister to a large Amherstieae clade (sensu Bruneau et al. 2001). The position of the South African genus *Schotia*, as well as of the South American *Goniorrhachis* and *Barnebydendron* (which often form a clade), are unresolved, but in individual trees these genera always occur among the first diverging branches in the Detarieae clade.

Within the resin-producing Detarieae, the relationships observed follow those described by Fougère-Danezan et al. (2007), based on phylogenetic analyses of the *trnL-F* region and of nuclear ribosomal ITS data. Some slight inconsistencies occur as to the relative position of the *Prioria* clade (Fig. 2). In the *trnL* analysis, this clade is weakly supported as sister to all other Detarieae (Fig. S1[4]). However, in the *matK*/*trnK* and combined analyses, the *Prioria* clade occurs as sister to the Detarieae s. str. clade, thus (weakly) supporting as monophyletic the resin-producing Detarieae clade. Morphologically, the resin-producing Detarieae are extremely variable, particularly in their floral morphology, with numerous differences in sepal, petal, and stamen number, but most members of this clade produce bicyclic diterpenes, a characteristic unique to this group (Fougère-Danezan et al. 2007). The West African *Daniellia* and Malagasy *Brandzeia* are poorly supported as sister to the Detarieae s. str. clade. These two genera are grouped with *Colophospermum* and *Hardwickia* (*Prioria* clade) in the ITS analyses of Fougère-Danezan et al. (2007), in conflict with the chloroplast data. Fougère-Danezan et al. (2007) consider the pattern of relationships recovered by the ITS data to be anomalous, possibly attributable to paralogy, nucleotide compositional bias or past hybridization events.

Within the Amherstieae clade, the addition of *matK*/*trnK* data results in slightly more resolution than that found in previous analyses (Bruneau et al. 2000, 2001), but few strongly supported clades are consistently resolved (Fig. 3). Two well-supported clades of mainly Asian genera, the *Saraca* clade (*Saraca*, *Endertia*, and *Lysidice*) and the *Afzelia* clade (*Afzelia*, *Intsia*, and *Brodriguesia*) are successive sister lineages to the remaining members of the Amherstieae clade. Within the Amherstieae, the New World *Brownea* clade also is supported as monophyletic. This clade, which includes *Brownea*, *Browneopsis*, *Ecuadendron*, *Heterostemon*, and the polyphyletic genera *Macrolobium* and *Paloue* (Redden and Herendeen 2006), groups all of the exclusively New World genera of the Amherstieae clade (except for American species of *Crudia* and *Cynometra*, which do not group with the *Brownea* clade). Few other clades are resolved within the Amherstieae clade and some of the larger genera are not supported as monophyletic. The large and complex African genus *Hymenostegia* is polyphyletic, with certain species occurring with *Scorodophloeus*, and others in a clade with *Leonardoxa*, *Talbotiella*, and *Loesenera*. In the *trnL* analyses, where we have higher species level sampling, all *Cynometra* species group with *Maniltoa*, but neither genus is supported as monophyletic. The genus *Lebruniodendron*, not included in previous molecular analyses, occurs in the Amherstieae clade, but in an unresolved

position in the *trnL* analysis (not sequenced for *matK*/*trnK*).

Within the Amherstieae, an additional clade is resolved in all analyses: the "Macrolobieae" clade of Bruneau et al. (2000, 2001). Bruneau et al. (2001) used this terminology to follow the classification of Breteler (1995), but because this clade does not include the genus *Macrolobium*, this designation is confusing and thus here we rename this group the *Berlinia* clade. Within this entirely African *Berlinia* clade, the genera *Didelotia*, *Plagiosiphon* p. p., *Librevillea*, *Gilbertiodendron*, and *Pellegriniodendron* form a weakly supported monophyletic group that is sister to the remaining genera. Based on a combined morphological and molecular analysis, Wieringa and Gervais (2003) delimited the "babjit" clade to include the genera *Brachystegia*, *Aphanocalyx*, *Bikinia*, *Julbernardia*, *Icuria*, and *Tetraberlinia*. In our analyses, the babjit clade of Wieringa and Gervais (2003) is not supported as monophyletic, because it does not include the genus *Brachystegia*, which instead occurs as sister to *Microberlinia* in an unresolved position in the *Berlinia* clade.

The crown node estimate for the Detarieae is 53.6 Ma, with estimates of 45.9 Ma for the Detarieae s. str., 44.8 Ma for the *Prioria* clade, and 52.1 Ma for the mostly African Amherstieae clade. The Detarieae clade includes a number of fossil constraints, several of which are relatively old, and it is also the clade that appears to be the most affected in terms of age estimates when fossil constraints are removed (Table 3). This likely is a consequence of the low rate of plastid evolution in the Detarieae relative to that found in most other caesalpinioid lineages (Fig. 5), which may confound rate modeling (c.f., Hugall et al. 2007). This implies that fossil constraints are particularly important to improve divergence time estimates within the Detarieae. The slowly evolving chloroplast genome in this group also has resulted in poorly resolved relationships at the generic level, particularly in the Amherstieae clade. In contrast to the chloroplast genome, various phylogenetic studies based on nuclear genes have noted what appears to be rapid evolution of the nuclear genome, with duplication events reported in low copy nuclear genes (e.g., Archambault and Bruneau 2004), as well as in ribosomal loci (e.g., Wieringa and Gervais 2003). Although rates of morphological evolution have not been studied, the relatively old divergence time estimated suggests that the extreme diversity in floral morphology observed in this group (e.g., Tucker 2002*a*, 2002*b*) has been present for nearly 55 Ma. The early Eocene age for this clade also adds an element of complexity for biogeographical analyses, with some deeply diverging continental disjunctions suggestive of vicariance events.

Umtiza clade

The *Umtiza* clade, recognised by Herendeen et al. (2003*b*) in their combined morphological and molecular analyses, is only weakly supported as monophyletic in the Bayesian analyses of the combined data (Fig. 4) and it is not resolved in any of the parsimony analyses of the molecular data alone. Despite their disparate biogeographical distribution, the seven genera of the *Umtiza* clade share a number of unique morphological features and several genera are dioecious, with small greenish flowers, a condition not often encountered in the Caesalpinioideae (Herendeen et al. 2003*b*). The temperate genera *Gleditsia* and *Gymnocladus*, and the South African *Umtiza*, form a strongly supported monophyletic group, supported by a number of morphological features. Similarly, the grouping of the Mediterranean genus *Ceratonia*, with the Southeast Asian *Acrocarpus* and Malagasy *Tetrapterocarpon* (*trnL* only), is also supported by several morphological characters. However, the relationship of these two clades with the Caribbean genus *Arcoa* is poorly resolved, with *Arcoa* occurring either as sister to the *Ceratonia* + *Acrocarpus* + *Tetrapterocaropn* clade (Fig. 4) or as sister to the *Gleditsia* + *Gymnocladus* + *Umtiza* clade (Herendeen et al. 2003*a*).

Among the caesalpinioid lineages, the *Umtiza* clade is one of the oldest clades with an estimated Late Paleocene crown node age (56.5 Ma, node 9, Figs. 4 and 5). This small clade includes two fossil constraints, both of which are well documented. However, as this group is not always supported as monophyletic based on plastid data some caution is necessary in interpreting divergence times for the *Umtiza* clade. Regardless, the relatively old age for this group suggests that vicariance events might partly explain the observed generic-level continental disjunctions, even though oceanic dispersal most likely also is necessary to account for the wide distribution of this clade (Schrire et al. 2005*b*).

The *Umtiza* clade is sister to a large clade (node 8, Fig. 4) consisting of the Mimosoideae and the remaining genera of the polyphyletic tribes Caesalpinieae and Cassieae. Five main caesalpinioid clades are recognised in this group, but relationships among them are poorly resolved. Though needing to be substantiated by additional, well-characterized fossil evidence and by molecular phylogenetic analyses of nuclear genes, the Late Paleocene – Early Eocene estimates obtained for members of this large clade appear quite consistent within our analyses.

Cassia clade

As sampled here, the *Cassia* clade includes six genera. The monophyletic genus *Senna* occurs as sister to the monophyletic *Cassia* (both with limited sampling), and this group is sister to the monospecific Brazilian genus *Melanoxylon*. The *Batesia* and *Chamaecrista* sister pair and the genus *Vouacapoua* are supported as members of this clade, but their relationships to the *Senna* + *Cassia* + *Melanoxylon* clade are poorly resolved. The genera *Cassia*, *Senna*, and *Chamaecrista* have long been considered close relatives and together were placed in the subtribe Cassiinae of the Cassieae (e.g., Irwin and Barneby 1981). Recent analyses of the genus *Senna* by Marazzi et al. (2006) also support the genus as monophyletic and sister to *Cassia*, but their sampling of other genera was too limited to assess relationships with other Cassieae and Caesalpinieae genera. *Batesia*, *Melanoxylon*, and *Vouacapoua* had been placed in the *Peltophorum* Group of the Caesalpinieae by Polhill and Vidal (1981) and Polhill (1994), but phylogenetic analyses of chloroplast DNA data by Haston et al. (2003, 2005) did not resolve these genera as part of a core *Peltophorum* Group (see below). Haston et al. (2005) noted that the genus *Vouacapoua* may not be monophyletic, with only one species, *Vouacapoua macropetala* Sandwith, occurring as nested in a position similar to that recovered here. Haston et al. (2005) also find support for a clade that includes the South

American genera, *Batesia*, *Melanoxylon*, and *Recordoxylon* (the latter not sampled here). However, *Cassia*, *Senna*, and *Chamecrista* were not included in their analyses, making it difficult to compare relationships within the *Cassia* clade. At least some of the genera in the *Cassia* clade have stomata on both surfaces of the leaflets, a feature that is restricted to certain clades within the caesalpinioid legumes (Herendeen et al. 2003*a*). In the Caesalpinioideae, the ability to nodulate is restricted to a few genera only, two of which, *Chamaecrista* and *Melanoxylon*, are members of this clade (Sprent 2007). The *Cassia* clade genera all have singly pinnate leaves and a number of them have heteromorphic stamens with anthers that dehisce by terminal pores.

The mostly New World *Cassia* clade has an estimated Early Eocene crown node age (53.0 Ma). Although numerous *Cassia* fossil leaflets have been reported in the literature, many of these have characteristics typical of many legume genera and therefore are not clearly identifiable to the genus *Cassia* or even to members of this clade (Herendeen et al. 1992).

Caesalpinia clade

The *Caesalpinia* clade includes all of the *Caesalpinia* s. l. segregate genera (*Coulteria*, *Erythrostemon*, *Guilandina*, *Libidibia*, *Mezoneuron*, *Poincianella*, *Tara*) recognised by Lewis (2005), as well as *Cordeauxia*, *Balsamocarpon*, *Hoffmannseggia*, *Stahlia*, *Pomaria*, *Haematoxylum*, *Pterolobium*, *Stuhlmannia*, and *Zuccagnia*, the latter two sequenced only for *trnL*. The genus *Pterogyne* is sister to all of these genera in all analyses, but always with low bootstrap support. In contrast, *Cordeauxia* is strongly supported as sister to all other remaining *Caesalpinia* clade genera. The remaining genera are resolved into two distinct clades. One comprises a well-supported group with *Balsamocarpon*, *Hoffmannseggia*, *Stahlia*, *Libidibia*, and *Zuccagania* (the latter only in the *trnL* analysis), sister to a clade that includes *Erythrostemon*, *Poincianella*, and *Pomaria*. The second clade includes *Guilandina* and *Pterolobium* as a sister pair, *Tara* and *Coulteria* as a sister pair, and the genera *Mezoneuron*, *Caesalpinia* s. str. and *Haematoxylum*. Certain of these generic relationships were also found by Simpson et al. (2003), based on *trnL* sequences supplemented with morphological data, but because of differences in sampling and low resolution, comparisons between the two studies are difficult. The generic composition of the two main clades found here is similar to that resolved by Lewis and Schrire (1995), based on morphological characters (although the positions within the two main clades of the genera *Haematoxylum* and *Tara* (as *Russellodendron* in Lewis and Schrire (1995)) are reversed), but differ from the two main groups noted by Lewis (2005) in placing the three genera *Erythrostemon*, *Poincianella*, and *Pomaria*, with the *Balsamocarpon* to *Zuccagnia* clade, rather than with the *Guilandina* to *Haematoxylum* clade.

Several more *Caesalpinia* s. l. species were sequenced in the *trnL* analysis and although the relationships are not well resolved in that analysis, most of the segregate genera either form monophyletic groups (*Coulteria*, *Erythrostemon*, *Libidibia*, *Mezoneuron*, *Poincianella*, *Tara*) or at least show no conflict in their position (simply unresolved: *Caesalpinia* s. str., *Guilandina*) (Fig. S1[4]).

The crown node of the *Caesalpinia* clade was estimated to be Late Paleocene in age (55.9 Ma). Two Middle Eocene fossil constraints are used in this clade, one of which is based on fossil fruits attributed to the genus *Mezoneuron* (node O, Table 1; Herendeen and Dilcher 1991; Lavin et al. 2005). The molecular age estimate for this node is 71% lower than the fossil-constrained age estimate, suggesting that fossil constraints in this clade also are critically important to accurately evaluate divergence times.

Peltophorum clade

The well-supported *Peltophorum* clade comprises eight genera, and is fully concordant with the core *Peltophorum* Group recognised by Haston et al. (2003, 2005). Generic-level relationships proposed by Haston et al. (2005), based on chloroplast *trnL-F*, *rbcL*, and *rps16* sequences are also entirely supported in our study. *Bussia* and *Peltophorum* form a strongly supported monophyletic group, as do *Delonix*, *Lemuropisum*, *Colvillea*, *Conzattia*, and *Parkinsonia*. *Schizolobium* is weakly supported (Fig. 4) as sister to this latter clade or it is unresolved with respect to these two clades (all parsimony analyses). No unique morphological synapomorpies are apparent for the *Peltophorum* clade, although together the genera share a combination of features, such as pinnate leaves, generally yellow petals, and narrow seeds (Haston et al. 2005). Each of the two strongly supported clades, however, have clear morphological synapomorpies (Haston et al. 2005) despite their wide geographical distributions, with African and Malagasy taxa grouped with South American genera in both clades.

Tachigali clade

The *Peltophorum* clade is sister to a group that includes *Arapatiella*, *Jacqueshuberia*, *Sclerolobium*, and *Tachigali*; a group of South American genera that was resolved in the Haston et al. (2005) study and named the *Tachigali* Group (Fig. 4). *Tachigali* and *Sclerolobium* have long been considered close relatives, and are now generally considered to be congeneric (e.g., Barneby 1996; H. van der Werff, Missouri Botanical Garden, unpublished data, 2008). Certain morphological and pollen features also support a close relationship of these two genera with *Arapatiella* and *Jacqueshuberia* (Haston et al. 2005). The genus *Diptychandra* had been placed in the *Sclerolobium* Group along with *Sclerolobium* and *Tachigali* by Polhill and Vidal (1981), but our analyses suggest it might be best placed with some of the *Dimorphandra* Group genera (see below).

Dimorphandra Group

The *Peltophorum* and *Tachigali* clades together are weakly supported as sister to a poorly supported clade comprising some members of the *Dimorphandra* Group (clade A, Fig. 4): *Burkea*, *Dimorphandra*, *Mora*, *Dinizia* (the latter transferred from the Mimosoideae in agreement with Luckow et al. (2000, 2003)) and two samples of a new taxon thought to be closely related to *Dinizia* and *Dimorphandra* (G.P. Lewis, unpublished data, 2007) (Fig. 4). *Campsiandra* is weakly supported as sister to this *Dimorphandra* Group p. p. clade. Four other *Dimorphandra* Group genera, *Diptychandra*, *Moldenhawera*, *Pachyelasma*, and *Erythrophleum*, together occur as a paraphyletic grade

(Group B, Fig. 4) at the base of the Mimosoideae, and this entire clade (Mimosoideae + *Dimorphandra* p. p. grade) occurs as sister to the grouping of the *Peltophorum*, *Tachigali*, and *Dimorphandra* p. p. clades, with strong support in the Bayesian analysis. Polhill and Vidal (1981) and Polhill (1994) placed 10 genera in the *Dimorphandra* Group, but this disparate group of genera has not been supported as monophyletic in any molecular analysis (e.g., Luckow et al. 2000; Bruneau et al. 2001). Several *Dimorphandra* Group genera have characteristics similar to taxa in the Mimosoideae (e.g., bipinnate leaves, alternate leaflets, small regular flowers; Luckow et al. 2000), supporting the close relationship found for *Diptychandra*, *Moldenhawera*, *Pachyelasma*, *Erythrophleum*, and the Mimosoideae. All other caesalpinioid genera known to nodulate occur in one or the other of these *Dimorphandra* Groups (i.e., *Dimorphandra*, *Campsiandra*, *Moldenhawera*, *Erythrophleum*) or in the *Tachigali* clade (*Sclerolobium*) (Sprent 2007).

Divergence time analyses and fossil evidence in caesalpinioid legumes

The PL divergence time estimates yielded ages that ranged from 34 to 56.5 Ma for the age of crown nodes of each of the major caesalpinioid lineages (Fig. 5). For all of the caesalpinioid crown nodes, the age estimates presented here are slightly older than those suggested by Lavin et al. (2005). Although this may be the consequence of fixing the stem node of the Leguminosae at a slightly older age, because the age of the fixed stem node (65 vs. 70 Ma) had little effect on the age estimates for other nodes, we consider that the difference is more likely the consequence of taxon sampling density and of the greater number of fossil caesalpinioid minimum age constraints used here.

Contrary to that found for the caesalpinioid legumes, the age estimate for the Papilionoideae crown clade (45.0 Ma) is younger than that found by Lavin et al. (2005), who reported an age of 58.6 Ma. However, our sampling of Papilionoideae is limited and we use many fewer fossil calibration points, both of which can lead to younger age estimates for a group (e.g., Magallón and Sanderson 2005). The age estimated here for the Mimosoideae crown node (45 Ma) is similar to the 42 Ma found by Lavin et al. (2005). Based on late Paleocene to early Eocene boundary fossil flowers described as *Protomimosoidea buchananensis* from southeastern USA (Crepet and Taylor 1985, 1986), Lavin et al. (2005) constrained the stem clade of the Mimosoideae to a minimum age of 55 Ma. However, these flowers share some features with genera of the *Dimorphandra* Group (Herendeen et al. 1992) and because of the polyphyletic nature of the *Dimorphandra* Group, and uncertainty as to its placement, we did not use this fossil in our analysis. However, this fossil could be considered as an upper bound for the crown node of the Mimosoideae plus *Dimorphandra*, *Tachigali*, and *Peltophorum* clade node (node 10, Fig. 4), a node that we estimated based on the other fossil evidence to be 56 Ma (Table 3). Thus, despite poorly resolved relationships among these caesalpinioid clades and the Mimosoideae, we can posit a late Paleocene age for this entire clade (node 8) and a mid Eocene age for the Mimosoideae.

Several methods have been proposed for assessing the divergence times of clades (reviewed in Magallón 2004; Sanderson et al. 2004; Rutschmann 2006), but often only a few fossils are available or clearly assignable to particular clades or nodes. Fossils are critically needed as calibration points when estimating divergence times among lineages, but several errors can be encountered when using fossils to calibrate lineages in divergence time estimations (Crepet et al. 2004; Near and Sanderson 2004; Forest et al. 2005; Rutschmann et al. 2007). Among these are errors in the taxonomic identification and assignment of fossils to particular clades, the positioning of fossils at either the stem or crown nodes, and the age attributed to the fossils. In addition, the choice of calibration points and the age assigned to a calibration point may have a significant effect on the estimate of ages of other nodes (e.g., Soltis et al. 2002; Magallón and Sanderson 2005). Here we have a number of fossils that have been well studied and that clearly can be assigned to particular clades based on synapomorphic characters. Although none of the 18 fossil constraints appears to be internally inconsistent with any of the other fossil constraints, several fossil nodes are particularly sensitive to the presence or absence of fossil calibration points in their age estimates and merit critical assessment (Table 3). Among these are the two Mimosoideae fossils (nodes Q and S, Fig. 4) and most likely results from our critical undersampling of this subfamily, leading to poor assignments to particular nodes of the Mimosoideae. One is the *Mezoneuron* fossil constraint (node O) that was described above and four others (nodes H, I, E, G) are associated with the Detarieae.

The fossil calibration point that we found to be most sensitive to the presence or absence of fossil constraints was that based on the *Aphanocalyx* bifoliolate fossil leaves described from the Middle Eocene of Tanzania by Herendeen and Jacobs (2000). Although we placed a minimum age constraint of 46 Ma for the stem node of the *Aphanocalyx* lineage, removing the fossil constraints resulted in a molecular age estimate of 4.4 Ma (node H, Fig. 3; Table 3). This fossil occurs in a kimberlite-associated lacustrine sediment whose age was determined by isotope analyses of crystals deposited at the time of formation (Herendeen and Jacobs 2000). The age of the deposit is likely reliable. Furthermore, although hypotheses on the phylogenetic position of *Aphanocalyx* may be altered slightly with new character evidence, this genus clearly belongs to the *Berlinia* clade, and more specifically to the babjit clade of Wieringa and Gervais (2003). *Aphanocalyx* is a genus of 14 species concentrated in West tropical Africa, but with two species in Zambesian and South Central Africa (Wieringa 1999; Mackinder 2005). The Detarieae in general are important forest elements in Africa and 48 of 82 genera are confined to Africa (including Madagascar), with a further three genera being pantropical and five more distributed in either Africa and Asia, or Africa and the neotropics (Mackinder 2005). Although the leaf morphology observed in the Tanzania fossil is clearly most similar to species of the extant genus *Aphanocalyx*, because of a poor fossil record, it may not be possible to ascribe this fossil to any extant African genus with certainty and it is possible that it belongs to a now extinct African taxon. Regardless, the wide discrepancy between the strictly molecular and fossil-calibrated age estimates for this node more likely reflects the generally

poor fossil record for African taxa (and more specifically for the Amherstieae clade) rather than problems inherent with the *Aphanocalyx* fossil. The incompleteness of the fossil record results in a critical undersampling of calibration points, which are essential for corroborating divergence time estimates, particularly across large phylogenetic trees with significant amounts of rate heterogeneity as seen here (Fig. 5). More and better fossil evidence for this portion of the Detarieae is needed if the age of this lineage is to be ascertained.

The minimum age for the *Crudia* lineage (45 Ma, node I, Fig. 3; Table 3) is also much older than the age estimated for this node without any fossil constraints (6.9 Ma). This minimum age fossil is based on fossil fruit and leaflets from the Middle Eocene of the Mississippi Embayment in southeastern USA, one of the best studied macrofossil sites in the USA (Herendeen and Dilcher 1990). The assignment of this fossil was based on careful evaluation of morphological and anatomical characters in both the fossil and extant legumes. The genus *Crudia* comprises 50 to 55 species that occur in Africa, America, and Asia. Herendeen and Dilcher (1990) considered the Mississippi Embayment fossils to be most similar to extant African or American species of *Crudia*. Although not well sampled in any molecular analysis to date, the genus *Crudia* does appear to be monophyletic, being well characterized morphologically by its laterally compressed, woody, dehiscent fruits, the valves usually covered with a dense rusty brown indumentum, and distinctive racemose inflorescences with apetalous flowers and reflexing sepals that reveal an erect pubescent ovary. The position of this genus in the Amherstieae clade is poorly resolved, as is the position of the majority of the genera in this clade. The slow rate of evolution of the chloroplast genome in the Detarieae and particularly in the Amherstieae clade may result in underestimates of the ages based on the uncalibrated molecular data in this clade (Figs. 3 and 5), if rate modeling is not entirely adequate.

The other two fossil constrained nodes that are most affected by the presence of fossil constraints are based on amber deposits, but from different regions of the world. The minimum age constraint for the *Daniellia* lineage was found to be 72.2% younger when no fossil constraints are available (15.3 Ma vs. 53 Ma, node G, Fig. 2; Table 3). This fossil, recently described by De Franceschi and De Ploëg (2003) from the early Eocene of the Paris Basin, is based on well-preserved fossil wood and flowering material. This amber deposit was dated using palynological evidence and chemically is similar to Dominican amber, produced by the genus *Hymenaea* (De Franceschi and De Ploëg 2003). Although both the wood and flower are morphologically similar to the extant species of the African genus *Daniellia*, these fossils could be attributable to other resin-producing Detarieae. The wood has the resin canals and vestured pits typical of the resin-producing Detarieae. The position of *Daniellia* (and its sister genus the Malagasy *Brandzeia*) is not well resolved in any molecular phylogenetic analysis (see Fougère-Danezan et al. 2007), but there is no doubt that this clade represents one of the first diverging lineages within the resin-producing Detarieae.

Similarly, although younger in age than the *Daniellia* fossil constraint, the constrained *Hymenaea* stem node (24 Ma) was estimated at 5.9 Ma without any fossil constraints in the analysis (node E, Fig. 2; Table 3). Graham (1992) reported late Eocene as the age of the *Hymenaea* fossils, and Lavin et al. (2005) used 34 Ma as a minimum age constraint for the stem node of *Hymenaea*. The age of the Dominican amber is problematic because there are differing opinions as to whether the amber has been reworked from the original depositional sediments and how much older it might be than the Miocene age deposits in which the amber is found (Iturralde-Vinent and MacPhee 1996). Because of the ambiguity in the age of the Dominican Amber deposits we use an intermediate age of 24 Ma which corresponds to the mean of older age estimate reported by Poinar et al. (1993). Regardless, the age estimated here without any fossil constraints is younger than the youngest age estimates for Dominican amber (Langenheim 2003). Interestingly, the age for the *Prioria* lineage stem node is estimated to be 44.7 Ma, much older than the minimum age constraint imposed on this lineage based on fossils from Dominican amber. We therefore consider that the age estimated using the fossil constraint is probably more reliable than an unconstrained estimate (see Rutschmann et al. 2007).

For nodes of interest where no fossil constraints were used, removing all the fossil constraints affected more strongly the age estimates for the Detarieae and sub-clades within (nodes 3–7; 68%–86% younger) than the age estimates for the other caesalpinioid clades (nodes 1–2, 8–12; 9%–57% younger). Although the age, identification and placement of certain of the fossils might be questionable, future studies are unlikely to alter significantly our evaluation. Rather, rate heterogeneity among lineages likely leads to underestimates of ages when fossil constraints are not available. Overall, the abundance of the fossil record for the caesalpinioid legumes serves to substantiate the lower to mid-Eocene ages estimated for these early diverving lineages of the legumes and provides a well-corroborated phylogenetic framework for future biogeographical and morphological analyses.

Acknowledgements

We thank Élyse-Ann Faubert, Marie-Pierre Gauthier, Vincent Manzanilla, Sébastien Renaut and Catherine Viel for laboratory assistance; Félix Forest, Marty Wojciechowski, Matt Lavin, Angelica Bello, Marie Fougère-Danezan and Karen Redden for providing unpublished sequences; Simon Joly and Vincent Manzanilla for help with phylogenetic analyses; Hannah Banks and Peter Gasson for their input in validating dates and identification of fossils; and Simon Joly, Matt Lavin, Marty Wojciechowski and the IRBV systematics group for comments that helped improve the manuscript. We are greatly indebted to numerous collaborators who, over the past several years, have helped us with field work and have provided plant material from their own field expeditions. In particular we thank Frans Breteler, Jan Wieringa, Barbara Mackinder, Colin Hughes, Karen Redden, Bente Klitgaard, John Clark, Rolland Ranaivojaona, Tefy-harison Andriamihajarivo, and Martin Cheek for their continued contributions. The caesalpinioid working group also thanks the staff at several herbaria for their hospitality and loans: US, K, MT, BM, P, NY, MO, WAG. We kindly acknowledge the Réseau québécois de calcul de haute performance for use of the computer clusters. This project was

funded by grants to A.B. from Natural Sciences and Engineering Research Council (Canada) and the Québec Ministère des Relations internationales, and by an National Science Foundation (USA) grant to P.S.H., G.P.L., and A.B.

References

Adegoke, O.S., Jan du Chene, R.E., Agumana, A.E., and Ajayi, P.O. 1978. Palynology and age of the Kerri-Kerri Formation, Nigeria. Rev. Esp. Micropaleontol. **10**: 267–283.

Archambault, A., and Bruneau, A. 2004. Phylogenetic utility of the *LEAFY/FLORICAULA* gene in the Caesalpinioideae (Leguminosae): gene duplication and a novel insertion. Syst. Bot. **29**: 609–626. doi:10.1600/0363644041744374.

Banks, H., Klitgaard, B.B., Lewis, G.P., Crane, P.R., and Bruneau, A. 2003. Pollen and the systematics of tribes Caesalpinieae and Cassieae (Caesalpinioideae: Leguminosae). *In* Advances in legume systematics. Part 10. Higher level phylogenetics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 95–122.

Barneby, R.C. 1996. Neotropical Fabales at NY: asides and oversights. Brittonia, **48**: 174–187. doi:10.2307/2807811.

Bello, M.A., Bruneau, A., Forest, F., and Hawkins, J.A. 2008. Elusive relationships within order Fabales: phylogenetic analyses using *matK* and *rbcL* sequence data. Syst. Bot. In press.

Breteler, F.J. 1995. The boundary between Amherstieae and Detarieae (Caesalpinioideae). *In* Advances in legume systematics. Part 7. Phylogeny. *Edited by* M.D. Crisp and J.J. Doyle. Royal Botanic Gardens, Kew, Surrey, UK. pp. 53–61.

Breteler, F.J. 1999. A revision of *Prioria*, including *Gossweilerodendron*, *Kingiodendron*, *Oxystigma*, and *Pterygopodium* (Leguminosae – Caesalpinioideae – Detarieae) with emphasis on Africa. Agric. Univ. Wagening. Pap. **99**: 1–61.

Bruneau, A., and Breteler, F.J., Wieringa, J.J., Gervais, G.Y.F., and Forest, F. 2000. Phylogenetic relationships in tribes Macrolobieae and Detarieae as inferred from chloroplast *trn*L intron sequences. *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 121–149.

Bruneau, A., Forest, F., Herendeen, P.S., Klitgaard, B.B., and Lewis, G.P. 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trn*L intron sequences. Syst. Bot. **26**: 487–514.

Calvillo-Canadell, L., and Cevallos-Ferriz, S.R.S. 2002. *Bauhcis moranii* gen. et sp. nov. (Cercideae, Caesalpinieae), an Oligocene plant from Tepexi de Rodriguez, Puebla, Mex., with leaf architecture similar to *Bauhinia* and *Cercis*. Rev. Palaeobot. Palynol. **122**: 171–184. doi:10.1016/S0034-6667(02)00135-5.

Calvillo-Canadell, L., and Cevallos-Ferriz, S.R.S. 2005. Diverse assemblage of Eocene and Oligocene leguminosae from Mexico. Int. J. Plant Sci. **166**: 671–692. doi:10.1086/430096.

Chappill, J.A. 1995. Cladistic analysis of the Leguminosae: the development of an explicit hypothesis. *In* Advances in legume systematics. Part 7. Phylogeny. *Edited by* M.D. Crisp and J.J. Doyle. Royal Botanic Gardens, Kew, Surrey, UK. pp. 1–9.

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrén, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn, G.H., Graham, S.W., Barret, S.C.H., Dayanandan, S., and Albert, V.A. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Ann. Mo. Bot. Gard. **80**: 528–580. doi:10.2307/2399846.

Chen, Y.F., and Zhang, D.X. 2005. *Bauhinia larsenii*, a fossil legume from Guangxi, China. Bot. J. Linn. Soc. **147**: 437–440. doi:10.1111/j.1095-8339.2005.00373.x.

Choi, H.K., Luckow, M.A., Doyle, J., and Cook, D.R. 2006. Development of nuclear gene-derived molecular markers linked to legume genetic maps. Mol. Genet. Genom. **276**: 56–70. doi:10.1007/s00438-006-0118-8.

Cowan, R.S., and Polhill, R.M. 1981*a*. Detarieae DC. (1825). *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 117–134.

Cowan, R.S., and Polhill, R.M. 1957*b*. Amherstieae Benth. emend. J. Léon. (1957). *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 135–142.

Crawley, M. 1988. Palaeocene wood from the Republic of Mali. Bull. Br. Mus. (Nat. Hist.) Geol. **44**: 3–14.

Crepet, W.L., and Dilcher, D.L. 1977. Investigations of angiosperms from the Eocene of North America: a mimosoid inflorescence. Am. J. Bot. **64**: 714–725. doi:10.2307/2441724.

Crepet, W.L., and Herendeen, P.S. 1992. Papilionoid flowers from the early Eocene of southeastern North America. *In* Advances in legume systematics. Part 4. The fossil record. *Edited by* P.S. Herendeen and D.L. Dilcher. Royal Botanic Gardens, Kew, Surrey, UK. pp. 43–55.

Crepet, W.L., and Taylor, D.W. 1985. The diversification of the Leguminosae: first fossil evidence of the Mimosoideae and Papilionoideae. Science (Wash.), **228**: 1087–1089. doi:10.1126/science.228.4703.1087.

Crepet, W.L., and Taylor, D.W. 1986. Primitive mimosoid flowers from the Paleocene-Eocene and their systematic and evolutionary implications. Am. J. Bot. **73**: 548–563. doi:10.2307/2444261.

Crepet, W.L., Nixon, K.C., and Gandolfo, M.A. 2004. Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceus deposits. Am. J. Bot. **91**: 1666–1682. doi:10.3732/ajb.91.10.1666.

De Franceschi, D., and De Ploëg, G. 2003. Origine de l'ambre des faciès sparnaciens (Éocène inférieur) du Bassin de Paris: le bois de l'ambre producteur. Geodiversitas, **25**: 633–647.

Doyle, J.J., and Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**: 11–15.

Doyle, J.J., Doyle, J.L., Ballenger, J.A., Dickson, E.E., Kajita, T., and Ohashi, H. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. Am. J. Bot. **84**: 541–554. doi:10.2307/2446030.

Doyle, J.J., Chappill, J.A., Bailey, C.D., and Kajita, T. 2000. Towards a comprehensive phylogeny of legumes: evidence from *rbcL* sequences and non-molecular data. *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau, Royal Botanic Gardens, Kew, Surrey, UK. pp. 1–20.

Felsenstein, J. 1981. A likelihood approach to character weighting and what it tells us about parsimony and compatibility. Biol. J. Linn. Soc. **16**: 183–196. doi:10.1111/j.1095-8312.1981.tb01847.x.

Forest, F., Bruneau, A., Hawkins, J.A., Kajita, T., Doyle, J.J., and Crane, P.R. 2002. The sister of the Leguminosae revealed: phylogenetic relationships in the Fabales determined using *trnL* and *rbcL* sequences. *In* Proceedings of Botany 2002 — Botany in the Curriculum: Integrating Research and Teaching [online]. 2–

4 August 2002. Madison, Wisc. Available from bsa2002.scientific-conference.net/section12/abstracts/101.shtml [accessed 11 February 2008].

Forest, F., Savolainen, V., Chase, M.W., Lupia, R., Bruneau, A., and Crane, P.R. 2005. Teasing apart molecular- versus fossil-based error estimates when dating phylogenetic trees: a case study in the birch family (Betulaceae). Syst. Bot. **30**: 118–133. doi:10.1600/0363644053661850.

Fortunato, R.H. 2005. Mimozygantheae. *In* Legumes of the world. *Edited by* G. Lewis, B. Schrire, B. Mackinder, and M. Lock. Royal Botanic Gardens, Kew, Surrey, UK. pp. 184–185.

Fougère-Danezan, M. 2005. Phylogénie moléculaire et morphologique des Detarieae résinifères (Leguminosae: Caesalpinioideae). Contribution à l'étude de l'histoire biogéographique des légumineuses. Ph.D. thesis, Université de Montréal (Montréal, Que.) and Université Paul Sabatier (Toulouse III, Toulouse, France).

Fougère-Danezan, M., Maumont, S., and Bruneau, A. 2003. Phylogenetic relationships in resin-producing Detarieae inferred from molecular data and preliminary results for a biogeographic hypothesis. *In* Advances in legume systematics. Part 10. Higher level phylogenetics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 161–180.

Fougère-Danezan, M., Maumont, S., and Bruneau, A. 2007. Relationships among resin producing Detarieae s.l. (Leguminosae) as inferred by molecular data. Syst. Bot. **32**: 748–761.

Gervais, G.Y.F., and Bruneau, A. 2002. Phylogenetic analysis of a polyphyletic African genus of Caesalpinioideae (Fabaceae): *Monopetalanthus* Harms. Plant Syst. Evol. **235**: 19–34. doi:10.1007/s00606-002-0222-0.

Graham, A. 1985. Studies in neotropical paleobotany. IV. The Eocene communities of Panama. Ann. Mo. Bot. Gard. **72**: 504–534. doi:10.2307/2399101.

Graham, A. 1989. Studies in neotropical paleobotany. VII. The Lower Miocene communities of Panama — The La Boca Formation. Ann. Mo. Bot. Gard. **76**: 50–66. doi:10.2307/2399342.

Graham, A. 1992. The current status of the legume fossil record in the Caribbean region. *In* Advances in legume systematics. Part 4. The fossil record. *Edited by* P.S. Herendeen and D.L. Dilcher. Royal Botanic Gardens, Kew, Surrey, UK. pp. 161–167.

Graham, A., and Jarzen, D.M. 1969. Studies in neotropical paleobotany. I. The Oligocene communities of Puerto Rico. Ann. Mo. Bot. Gard. **56**: 308–357. doi:10.2307/2394849.

Guinet, Ph. 1981. Mimosoideae: the characters of their pollen grains. *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 835–855.

Guinet, Ph. 1990. The genus *Acacia* (Leguminosae: Mimosoideae); its affinities as borne out by its pollen characters. Plant Syst. Evol. **5**(Suppl.): 81–90.

Guinet, Ph., and Lugardon, B. 1976. Diversité des structures de l'exine dans le genre *Acacia* (Mimosaceae). Pollen Spores, **18**: 483–511.

Guinet, Ph., El Sabrouty, M.N., Soliman, H.A., and Omran, A.M. 1987. Étude des caractères du pollen des légumineuses-Mimosoideae des sédiments Tertiaires du nord-ouest de l'Égypte. Mém. Trav. E.P.H.E. Inst. Montpellier, **17**: 159–171.

Haston, E.M., Lewis, G.P., and Hawkins, J.A. 2003. A phylogenetic investigation of the *Peltophorum* group (Caesalpinieae: Leguminosae). *In* Advances in legume systematics. Part 10. Higher level phylogenetics. *Edited by* A. Bruneau and B.B. Klitgaard. Royal Botanic Gardens, Kew, Surrey, UK. pp. 149–159.

Haston, E.M., Lewis, G.P., and Hawkins, J.A. 2005. A phylogenetic reappraisal of the *Peltophorum* group (Caesalpinieae: Leguminosae) based on the chloroplast *trnL-F*, *rbcL* and *rps16* sequence data. Am. J. Bot. **92**: 1359–1371. doi:10.3732/ajb.92.8.1359.

Herendeen, P.S. 1992. The fossil history of the Leguminosae from the Eocene of southeastern North America. *In* Advances in legume systematics. Part 4. The fossil record. *Edited by* P.S. Herendeen and D.L. Dilcher. Royal Botanic Gardens, Kew, Surrey, UK. pp. 85–160.

Herendeen, P.S. 2000. Structural evolution in the Caesalpinioideae (Leguminosae). *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 45–64.

Herendeen, P.S., and Crane, P.R. 1992. Early caesalpinioid fruits from the Palaeogene of southern England. *In* Advances in legume systematics. Part 4. The fossil record. *Edited by* P.S. Herendeen and D.L. Dilcher. Royal Botanic Gardens, Kew, Surrey, UK. pp. 57–68.

Herendeen, P.S., and Dilcher, D.L. 1990. Reproductive and vegetative evidence for the occurrence of *Crudia* (Leguminosae, Caesalpinioideae) in the Eocene of Southeastern North-America. Bot. Gaz. **151**: 402–413. doi:10.1086/337840.

Herendeen, P.S., and Dilcher, D.L. 1991. *Caesalpinia* subgenus *Mezoneuron* (Leguminosae, Caesalpinioideae) from the Tertiary of North America. Am. J. Bot. **78**: 1–12. doi:10.2307/2445223.

Herendeen, P.S., and Jacobs, B.F. 2000. Fossil legumes from the Middle Eocene (46.0 Ma) Mahenge Flora of Singida, Tanzania. Am. J. Bot. **87**: 1358–1366. doi:10.2307/2656727. PMID:10991905.

Herendeen, P.S., and Manchester, S.R. 2004. Fruits and foliage of Cercis from the Late Eocene of Oregon. *In* Proceedings of Botany 2004 — Alpine Diversity: Adapted to the Peaks [online]. 31 July –5 August 2004. Snowbird resort, Salt Lake city, Utah. Available from www.botanyconference.org [accessed 11 February 2008].

Herendeen, P.S., and Wing, S. 2001. Papilionoid legume fruits and leaves from the Paleocene of northwestern Wyoming. *In* Proceedings of Botany 2001 — Plants and People [online]. 12–16 August 2001. Albuquerque Convention Center, Albuquerque, New Mexico. Available from www.botany2001.org/section7/abstracts/26.shtml [accessed 11 February 2008]..

Herendeen, P.S., Crepet, W.L., and Dilcher, D.L. 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. *In* Advances in legume systematics. Part 4. The fossil record. *Edited by* P.S. Herendeen and D.L. Dilcher. Royal Botanic Gardens, Kew, Surrey, UK. pp. 303–316.

Herendeen, P.S., Bruneau, A., and Lewis, G.P. 2003a. Phylogenetic relationships in the caesalpinioid legumes: a preliminary analysis based on morphological and molecular data. *In* Advances in legume systematics. Part 10. Higher level phylogenetics. *Edited by* B. B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 37–62.

Herendeen, P.S., Lewis, G.P., and Bruneau, A. 2003*b*. Floral morphology in caesalpinioid legumes: testing the monophyly of the "*Umtiza*" clade. Int. J. Plant Sci. **164**: S393–S407. doi:10.1086/376881.

Hu, J.-M., Lavin, M., Wojciechowski, M.F., and Sanderson, M.J. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trnK/matK* sequences and its implications for evolutionary patterns in Papilionoideae. Am. J. Bot. **87**: 418–430. doi:10.2307/2656638. PMID:10719003.

Hueber, F.M., and Langenheim, J. 1986. Dominican amber tree had African ancestors. Geotimes, **31**: 8–10.

Huelsenbeck, J.P., and Ronquist, F.R. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics, **17**: 754–755. doi:10.1093/bioinformatics/17.8.754. PMID:11524383.

Hugall, A.F., Foster, R., and Lee, M.S.Y. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Syst. Biol. **56**: 543–563. doi:10.1080/10635150701477825. PMID:17654361.

Ireland, H., Pennington, R.T., and Preston, J. 2000. Molecular systematics of the Swartzieae. *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 217–231.

Irwin, H.S., and Barneby, R.C. 1981. Cassieae. *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 97–106.

Iturralde-Vinent, M.A., and MacPhee, R.D.E. 1996. Age and paleogeographical origin of Dominican amber. Science (Wash.), **273**: 1850–1852. doi:10.1126/science.273.5283.1850.

Jacobs, B.F., and Herendeen, P.S. 2004. Eocene dry climate and woodland vegetation in tropical Africa reconstructed from fossil leaves from northern Tanzania. Palaeogreogr. Palaeoclimatol. Palaeoecol. **213**: 115–123.

Joly, S., Starr, J.R., Lewis, W.H., and Bruneau, A. 2006. Polyploid and hybrid evolution in roses east of the Rocky Mountains. Am. J. Bot. **93**: 412–425. doi:10.3732/ajb.93.3.412.

Kajita, T., Ohashi, H., Tateishi, Y., Bailey, C. D., and Doyle, J. J.· 2001. *RbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. Syst. Bot. **26**: 515–536.

Käss, E., and Wink, M. 1996. Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcL*-sequences. Biochem. Syst. Ecol. **24**: 365–378. doi:10.1016/0305-1978(96)00032-4.

Kelchner, S.A. 2000. The evolution of non-coding chloroplast DNA and its application to plant systematics. Ann. Mo. Bot. Gard. **87**: 482–498. doi:10.2307/2666142.

Klitgaard, B.B., Forest, F., Bruneau, A., and Banks, H. 2002. *Duparquetia* (Leguminosae: Caesalpinioideae): an enigmatic basal legume with orchid-like flowers. *In* Proceedings of Flowers: diversity, development and evolution. 5–7 July 2002. Institute of Systematic Botany, University of Zurich, Zurich, Switzerland..

Kosakovsky Pond, S.L., Frost, S.D.W., and Muse, S.V. 2005. HyPhy: hypothesis testing using phylogenies [online]. Bioinformatics, **2**: 676–679. Available from www.hyphy.org/[accessed 26 November 2007].

Langenheim, J.H. 2003. Plant resins: chemistry, evolution, ecology, and ethnobotany. Timber Press, Portland, Ore.

Lavin, M., Thulin, M., Labat, J.-N., and Pennington, R.T. 2000. Africa, the odd man out: molecular biogeographic studies of dalbergioid legumes (Fabaceae) suggest otherwise. Syst. Bot. **25**: 449–467. doi:10.2307/2666689.

Lavin, M., Schrire, B.D., Lewis, G.P., Pennington, R.T., Delgado-Salinas, A., Thulin, M., Hughes, C.E., Beyra Matos, A., and Wojciechowski, M.F. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **359**: 1509–1522. doi:10.1098/rstb.2004.1536. PMID:15519969.

Lavin, M., Herendeen, P.S., and Wojciechowski, M.F. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. Syst. Biol. **54**: 575–594. doi:10.1080/10635150590947131. PMID:16085576.

Léonard, J. 1957. Genera des Cynometreae et des Amherstieae africaines (Leguminosae-Caesalpinioideae). Mémoires de la Classe des sciences, Académie Royale des sciences, des lettres et des beaux-arts de Belgique, memoires 8, **30**(2): 3–312.

Lewis, G.P. 2005. Caesalpinieae. *In* Legumes of the world. *Edited by* G. Lewis, B. Schrire, B. Mackinder and M. Lock. Royal Botanic Gardens, Kew, Surrey, UK. pp. 127–161.

Lewis, G.P., and Forest, F. 2005. Cercideae. *In* Legumes of the world. *Edited by* G. Lewis, B. Schrire, B. Mackinder and M. Lock. Royal Botanic Gardens, Kew, Surrey, UK. pp. 57–67.

Lewis, G.P., and Schrire, B.D. 1995. A reappraisal of the *Caesalpinia* group (Caesalpinioideae: Caesalpinieae) using phylogenetic analysis. *In* Advances in legume systematics. Part 7. Phylogeny. *Edited by* M.D. Crisp and J.J. Doyle. Royal Botanic Gardens, Kew, Surrey, UK. pp. 41–52.

Lewis, G.P., Schrire, B., Mackinder, B., and Lock, M. 2005. Legumes of the world. Royal Botanic Gardens, Kew, Surrey, UK.

Luckow, M., White, P.J., and Bruneau, A. 2000. Relationships among the basal genera of mimosoid legumes. *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 165–180.

Luckow, M., Miller, J.T., Murphy, D.J., and Livshultz, T. 2003. A phylogenetic analysis of the Mimosoideae (Leguminosae) based on chloroplast DNA sequence data. *In* Advances in legume systematics. Part 10. Higher level systematics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 197–220.

MacGinitie, H.D. 1953. Fossil plants of the Florissant Beds, Colorado. Carnegie Institution of Washington Publications. Publ. **599**: i-iii, 1–198.

Mackinder, B. 2005. Tribe Detarieae. *In* Legumes of the world. *Edited by* G. Lewis, B. Schrire, B. Mackinder, and M. Lock. Royal Botanic Gardens, Kew, Surrey, UK. pp. 68–109.

Magallón, S.A. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. Int. J. Plant Sci. **165**: S7–S21. doi:10.1086/383336.

Magallón, S.A., and Sanderson, M.J. 2001. Absolute diversification rates in angiosperm clades. Evolution, **55**: 1762–1780. PMID:11681732.

Magallón, S.A., and Sanderson, M.J. 2005. Angiosperm divergence times: The effect of genes, codon positions, and time constraints. Evolution, **59**: 1653–1670. doi:10.1554/04-565.1. PMID:16329238.

Marazzi, B., Endress, P.K., De Queiroz, L.P., and Conti, E. 2006. Phylogenetic relationships within Senna (Leguminosae, Cassiinae) based on three chloroplast DNA regions: Patterns in the evolution of floral symmetry and extrafloral nectarines. Am. J. Bot. **93**: 288–303. doi:10.3732/ajb.93.2.288.

Mildenhall, D.C. 1972. Fossil pollen of *Acacia* type from New Zealand. N.Z. J. Bot. **10**: 485–494.

Muller, J. 1981. Fossil pollen records of extant angiosperms. Bot. Rev. **47**: 1–142. doi:10.1007/BF02860537.

Near, T.J., and Sanderson, M.J. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. Philos. Trans. R. Soc. Lond. B Biol. Sci. **359**: 1477–1483. doi:10.1098/rstb.2004.1523. PMID: 15519966.

Near, T.J., Meylan, P.A., and Shaffer, H.B. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. Am. Nat. **165**: 137–146. doi:10.1086/427734. PMID:15729646.

Pennington, R.T., Klitgaard, B.B., Ireland, H., and Lavin, M. 2000. New insights into floral evolution of basal Papilionoideae from molecular phylogenies. *In* Advances in legume systematics. Part 9. *Edited by* P.S. Herendeen and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 233–248.

Pennington, R.T., Lavin, M., Ireland, H., Klitgaard, B.B., Preston, J., and Hu, J.-M. 2001. Phylogenetic relationships of basal papilionoid legumes based upon sequences of the chloroplast intron *trnL*. Syst. Bot. **26**: 537–556.

Poinar, G., and Poinar, R. 1999. The amber forest. A reconstruction of a vanished world. Princeton University Press, Princeton, N.J.

Poinar, H.N., Cano, R.J., and Poinar, G.O.J. 1993. DNA from an extinct plant. Nature (London), **363**: 677. doi:10.1038/363677a0.

Polhill, R.M. 1981. Papilionoideae. *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 191–208.

Polhill, R.M. 1994. Complete synopsis of legume genera. *In* Phytochemical dictionary of the Leguminosae. Vol. 1. Plants and their constituents. *Edited by* F.A. Bisby, J. Buckingham, and J.B. Harborne. Chapman and Hall, London, UK. pp. xlix–liv.

Polhill, R.M., and Vidal, J.E. 1981. Caesalpinieae. *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 81–96.

Polhill, R.M., Raven, P.H., and Stirton, C.H. 1981. Evolution and systematics of the Leguminosae. *In* Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 1–26.

Posada, D., and Crandall, K.A. 1998. ModelTest: testing the model of DNA substitution [online]. Bioinformatics, **14**: 817–818. Available from darwin.uvigo.es/software/modeltest.html [accessed 25 September 2007]. doi:10.1093/bioinformatics/14.9.817. PMID:9918953.

Rambaut, A., and Drummond, A.J. 2007. Tracer version 1.4 [online]. Available from tree.bio.ed.ac.uk/software/tracer/ [accessed 30 July 2007]. .

Redden, K.M., and Herendeen, P.S. 2006. Morphology and phylogenetic analysis of *Paloue* and related genera in the *Brownea* clade (Caesalpinioideae: Leguminosae). Int. J. Plant Sci. **167**: 1229–1246. doi:10.1086/508065.

Ronquist, F., Huelsenbeck, J.P., and van der Mark, P. 2005. Mr. Bayes 3.1 manual [online]. Available from mrbayes.csit.fsu.edu/mb3.1_manual.pdf [accessed 14 February 2008].

Rutschmann, F. 2006. Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. Divers. Distrib. **12**: 35–48. doi:10.1111/j.1366-9516.2006.00210.x.

Rutschmann, F., Eriksson, T., Abu Salim, K., and Conti, E. 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. Syst. Biol. **56**: 591–608. doi:10.1080/10635150701491156. PMID: 17654364.

Salard-Cheboldaeff, M. 1978. Sur la palynoflore Maestrichtienne et Tertiaire du Bassin sédimentaire Littoral du Cameroun. Pollen Spores, **20**: 215–260.

Sanderson, M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. **19**: 101–109. PMID:11752195.

Sanderson, M.J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock [online]. Bioinformatics, **19**: 301–302. Available from loco.biosci.arizona.edu/r8s/ [accessed 24 July 2007]. doi:10.1093/bioinformatics/19.2.301. PMID:12538260.

Sanderson, M.J. 2004. r8s: version 1.7. User's manual [online]. Available from loco.biosci.arizona.edu/r8s/r8s1.7.manual.pdf [accessed 30 December 2004].

Sanderson, M.J., Thorne, J.L., Wikstrom, N., and Bremer, K. 2004. Molecular evidence on plant divergence times. Am. J. Bot. **91**: 1656–1665. doi:10.3732/ajb.91.10.1656.

Schrire, B.D., Lavin, M., and Lewis, G.P. 2005*a*. Global distribution patterns of the Leguminosae: insights from recent phylogenies. *In* Plant diversity and complexity patterns: local, regional and global dimensions. *Edited by* I. Friis and H. Balslev. Danske Videnskabernes Selskab, Biologiske Skrifter, **55**: 375–422.

Schrire, B.D., Lewis, G.P., and Lavin, M. 2005*b*. Biogeography of the Leguminosae. *In* Legumes of the world. *Edited by* G. Lewis, B. Schrire, B. Mackinder, and M. Lock. Royal Botanic Gardens, Kew, Surrey, UK. pp. 21–54.

Simmons, M.P., and Ochoterena, H. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. **49**: 369–381. doi:10.1080/10635159950173889. PMID:12118412.

Simpson, B.B., and Lewis, G.P. 2003. New combinations in *Pomaria* (Caesalpinioideae: Leguminosae). Kew Bull. **58**: 175–184. doi:10.2307/4119360.

Simpson, B.B., Larkin, L.L., and Weeks, A. 2003. Progress towards resolving the relationships of the *Caesalpinia* Group (Caesalpinieae: Caesalpinioideae: Leguminosae). *In* Advances in legume systematics. Part 10. Higher level systematics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 123–148.

Sinou, C., Bruneau, A., Forest, F., and Lewis, G.P. 2007. Phylogenetic analysis of the genus *Bauhinia* (Leguminosae: Caesalpiniodeae: Cercideae) [online]. *In* Proceedings of Botany and Plant Biology 2007 Joint Congress. 7–11 July 2007. Chicago, Ill. Available from www.2007.botanyconference.org/engine/search/index.php?func=detail&aid=1135 [accessed 11 February 2008].

Soltis, D.E., Soltis, P.S., Nickrent, D., Johnson, L.A., Hahn, W.H., Hoot, S.B., Sweere, J.A., Kuzoff, R.B., Kron, K.A., Chase, M.W., Swensen, S.M., Zimmer, E.A., Chaw, S.-M., Gillespie, L.J., Kress, W.J., and Systsma, K.J. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. Ann. Mo. Bot. Gard. **84**: 1–49. doi:10.2307/2399952.

Soltis, P.S., Soltis, D.E., Savolainen, V., Crane, P.R., and Barraclough, T.G. 2002. Rate heterogeneity among lineages of Tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc. Natl. Acad. Sci. U.S.A. **99**: 4430–4435. doi:10.1073/pnas.032087199. PMID: 11917101.

Sprent, J.I. 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. New Phytol. **174**: 11–25. doi:10.1111/j.1469-8137.2007.02015.x. PMID:17335493.

Swofford, D.L. 2001. PAUP* – Phylogenetic analysis using parsimony (and other methods). Version 4.0. Sinauer Associates, Sunderland, Mass.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**: 4876–4882. doi:10.1093/nar/25.24.4876.

Tucker, S.C. 2002*a*. Comparative floral ontogeny in Detarieae (Leguminosae: Caesalpinioideae). 1. Radially symmetrical taxa lacking organ suppression. Am. J. Bot. **89**: 875–887. doi:10.3732/ajb.89.6.875.

Tucker, S.C. 2002*b*. Comparative floral ontogeny in Detarieae (Leguminosae: Caesalpinioideae). 2. Zygomorphic taxa with petal and stamen suppression. Am. J. Bot. **89**: 888–907. doi:10.3732/ajb.89.6.888.

Tucker, S.C. 2003. Floral development in legumes. Plant Physiol. **131**: 911–926. doi:10.1104/pp.102.017459. PMID:12644644.

Wieringa, J.J. 1999. *Monopetalanthus* exit. A systematic study of *Aphanocalyx*, *Bikinia*, *Icuria*, *Michelsonia* and *Tetraberlinia* (Leguminosae, Caesalpinioideae). Agric. Univ. Wagening. Pap. **99–4**: 1–320.

Wieringa, J.J., and Gervais, G.Y.F. 2003. Phylogenetic analyses of combined morphological and molecular data sets on the *Aphanocalyx-Bikinia-Tetraberlinioa* group (Leguminosae, Caesalpinioideae, Detarieae s.l.). *In* Advances in legume systematics.

Part 10. Higher level systematics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew. pp. 181–196.

Wojciechowski, M.F. 2003. Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective. *In* Advances in legume systematics. Part 10. Higher level systematics. *Edited by* B.B. Klitgaard and A. Bruneau. Royal Botanic Gardens, Kew, Surrey, UK. pp. 5–35.

Wojciechowski, M.F., Lavin, M., and Sanderson, M.J. 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. Am. J. Bot. **91**: 1846–1862. doi:10.3732/ajb.91.11.1846.

Wunderlin, R., Larsen, K., and Larsen, S.S. 1981. Cercideae. Advances in legume systematics. Part 1. *Edited by* R.M. Polhill and P.H. Raven. Royal Botanic Gardens, Kew, Surrey, UK. pp. 107–116.

Young, N.D., and Healy, J. 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics, **4**: 6. doi:10.1186/1471-2105-4-6. PMID:12689349.