American Journal of Botany 91(11): 1846–1862. 2004.

A PHYLOGENY OF LEGUMES (LEGUMINOSAE) BASED ON ANALYSIS OF THE PLASTID *MATK* GENE RESOLVES MANY WELL-SUPPORTED SUBCLADES WITHIN THE FAMILY[1]

MARTIN F. WOJCIECHOWSKI,[2,5] MATT LAVIN,[3] AND MICHAEL J. SANDERSON[4]

[2]School of Life Sciences, Arizona State University, Tempe, Arizona 85287-4501 USA; [3]Department of Plant Sciences, Montana State University, Bozeman, Montana 59717 USA; and [4]Section of Evolution and Ecology, University of California, Davis, California 95616 USA

Phylogenetic analysis of 330 plastid *matK* gene sequences, representing 235 genera from 37 of 39 tribes, and four outgroup taxa from eurosids I supports many well-resolved subclades within the Leguminosae. These results are generally consistent with those derived from other plastid sequence data (*rbcL* and *trnL*), but show greater resolution and clade support overall. In particular, the monophyly of subfamily Papilionoideae and at least seven major subclades are well-supported by bootstrap and Bayesian credibility values. These subclades are informally recognized as the *Cladrastis* clade, genistoid sensu lato, dalbergioid sensu lato, mirbelioid, millettioid, and robinioid clades, and the inverted-repeat-lacking clade (IRLC). The genistoid clade is expanded to include genera such as *Poecilanthe*, *Cyclolobium*, *Bowdichia*, and *Diplotropis* and thus contains the vast majority of papilionoids known to produce quinolizidine alkaloids. The dalbergioid clade is expanded to include the tribe Amorpheae. The mirbelioids include the tribes Bossiaeeae and Mirbelieae, with Hypocalypteae as its sister group. The millettioids comprise two major subclades that roughly correspond to the tribes Millettieae and Phaseoleae and represent the only major papilionoid clade marked by a macromorphological apomorphy, pseudoracemose inflorescences. The robinioids are expanded to include *Sesbania* and members of the tribe Loteae. The IRLC, the most species-rich subclade, is sister to the robinioids. Analysis of the *matK* data consistently resolves but modestly supports a clade comprising papilionoid taxa that accumulate canavanine in the seeds. This suggests a single origin for the biosynthesis of this most commonly produced of the nonprotein amino acids in legumes.

Key words: caesalpinioid legumes; Leguminosae; *matK*; mimosoid legumes; papilionoid legumes; phylogeny.

The legume family is the third largest family of angiosperms (Mabberley, 1997) with approximately 730 genera and over 19 400 species worldwide (Lewis et al., in press). Legumes are second only to Poaceae (the grasses) in agricultural and economic importance. The family includes horticultural varieties and many species harvested as crops and for oils, fiber, fuel, timber, medicines, and chemicals. Ranging in habit from large trees to annual herbs, the family is well represented throughout temperate and tropical regions of the world (Rundel, 1989). The Leguminosae is particularly diverse, however, in tropical forests with a seasonally dry aspect and temperate shrublands tailored by xeric climates. Legumes are noticeably absent to poorly represented in mesic temperate habitats, including many arctic and alpine regions and the understory of cool temperate forests. The predilection of legumes for semi-arid to arid habitats is related to a nitrogen-demanding metabolism, which is thought to be an adaptation to climatically variable or unpredictable habitats whereby leaves can be produced economically and opportunistically (McKey, 1994). Indeed, nitrogen fixation via root-nodulating symbiotic bacteria is just one of several ways (in addition to associations with arbuscular mycorrhizae, ectomycorrhizae, and uptake of inorganic nitrogen compounds) in which legumes obtain high levels of nitrogen to meet the demands of their metabolism (Sprent, 1994, 2001). All legumes play an important role in the terrestrial nitrogen cycle regardless of whether they form root nodules (Sprent, 2001). Considered to be a tropical family with perhaps a late Cretaceous origin (65–70 Mya), the Leguminosae has an abundant and continuous fossil record since the Tertiary (Crepet and Taylor, 1985, 1986; Crepet and Herendeen, 1992; Herendeen et al., 1992). The occurrence of diverse assemblages of taxa representing all three subfamilies at multiple localities dating from the middle to upper Eocene, especially the Mississippi Embayment of southeastern North America, suggests that most major lineages of woody legumes (except for the tribe Cercideae) were present and had diversified extensively by this time (Herendeen et al., 1992).

Reconstructing the phylogenetic relationships of the Leguminosae is essential for understanding the origin and diversification of this ecologically and economically important family of angiosperms. Comprehensive phylogenetic analyses of Leguminosae began with the plastid gene *rbcL* (Doyle, 1995; Käss and Wink, 1995, 1996; Doyle et al., 1997) following the early, widespread use of this gene for phylogenetic studies of land plant relationships (e.g., Chase et al., 1993). Among the conclusions that emerged, the monophyly of the Fabales (sensu Angiosperm Phylogeny Group, 2003) and the sister rela-

[1] Manuscript received 18 December 2003; revision accepted 10 June 2004.

The authors thank M. Crisp, H. C. de Lima, A. Delgado-Salinas, P. Fritsch, P. S. Herendeen, C. Hughes, D. Kelch, J.-N. Labat, A. Liston, R. Pasquet, R. T. Pennington, A. S. Salywon, K. P. Steele, M. Thulin, N. Weeden, curators of the ARIZ, ASU, DAV, K, MO, MONT, and UC/JEPS herbaria, R. Puente (Desert Botanical Garden), E. Simms (UC Berkeley Botanical Garden), E. Sandoval and T. Metcalf (UC Davis Botanical Conservatory), and B. Hall (UC Santa Cruz Arboretum) for kindly providing genomic DNAs, seeds, leaf materials, and specimen loans for many of the species included in this study. A. Bruneau, P. S. Herendeen, and R. T. Pennington provided unpublished *matK* sequences for selected taxa. We especially thank K. P. Steele for many helpful suggestions, and A. Bruneau and B. Schrire for comments that improved this paper. This study was supported in part by grants from the National Science Foundation to M. J. S. and M. F. W. (DEB-9407824) and to M. L. (DEB-0075202) and funds provided by Arizona State University to M. F. W.

[5] E-mail: mfwojciechowski@asu.edu.

tionship of legumes to Polygalaceae, Surianaceae, and the rosaceous genus *Quillaja* Molina were very strongly supported (Doyle et al., 2000). Second, the monophyly of Leguminosae is consistently resolved although not as strongly as for the Fabales (Doyle et al., 2000; Kajita et al., 2001). Third, while monophyly of mimosoid legumes (subfamily Mimosoideae) is well supported by the *rbcL* data (Doyle et al., 2000), a more extensive sampling of the subfamily suggested certain mimosoid genera, *Dinizia* Ducke and *Piptadeniastrum* Brenan, were unresolved with respect to related caesalpinioid outgroups (Luckow et al., 2000). Fourth, the subfamily Caesalpinioideae (caesalpinioids) is consistently resolved as paraphyletic with respect to mimosoids and papilionoids (e.g., Polhill et al., 1981; Doyle et al., 2000; Kajita et al., 2001), although several well-supported subclades have been detected in recent studies of this subfamily; for example, the tribe Cercideae, resolved as the sister clade to the rest of the family (Doyle et al., 2000), the tribe Detarieae sensu lato (s.l.), distributed principally in tropical Africa and including approximately half of the genera in the Caesalpinioideae (Bruneau et al., 2001; Herendeen et al., 2003a), and the ''*Umtiza*'' clade (Herendeen et al., 2003b). Lastly, the traditionally recognized subfamily Papilionoideae (sensu Polhill, 1981a, 1994) is consistently resolved as monophyletic, albeit with only modest support (e.g., Doyle et al., 1997; Kajita et al., 2001).

The Papilionoideae has received the most attention, if only because it is the largest and most widespread of the three legume subfamilies with an estimated 476 genera and 13 860 species (Lewis et al., in press). Papilionoids traditionally have been diagnosed by traits that now are considered synapomorphies of the subfamily. These include wood with predominantly paratracheal axial parenchyma that is usually storied; vessels with alternate vestured pits and simple perforation plates; absence of bipinnate leaves; unidirectional initiation of sepals, petals, and stamens; clawed petals; and a seed testa with a hilar valve and no pleurogram (Polhill, 1981a; Tucker, 1987a, 2002; Tucker and Douglas, 1994; Chappill, 1995; Gasson, 2000). These many distinctions have sometimes resulted in papilionoids being ranked at the familial level (e.g., Hutchinson, 1964; Takhtajan, 1969). Moreover, support for the monophyly of Papilionoideae has not changed with family-wide molecular phylogenetic analyses involving the plastid *rbcL* locus (Käss and Wink, 1995, 1996, 1997; Doyle et al., 1997, 2000; Kajita et al., 2001) or *trnL* intron (Pennington et al., 2001).

Despite insights gained into the higher-level relationships of the family from studies of the *rbcL* locus, and to a lesser extent the *trnL-F* region, many issues in legume phylogeny remain unresolved (reviewed in Wojciechowski, 2003). This is particularly true for the relationships within the caesalpinioid and mimosoid subgroups and among some of the major papilionoid clades, the genistoids, dalbergioids, millettioids-phaseoloids, and Hologalegina (e.g., Crisp et al., 2000; Hu et al., 2000; Wojciechowski et al., 2000; Lavin et al., 2001). More variable nucleotide sequences are needed to improve the resolution of and support for the major clades within legumes. The most promising is the plastid gene *matK*, which has been shown by several recent studies on different papilionoid subgroups to provide excellent resolution among closely related genera (Hu et al., 2000; Lavin et al., 2001, 2003; Steele and Wojciechowski, 2003). Here we draw on these recent studies and a large number of new *matK* sequences as part of a more extensive phylogenetic analysis of the family, with particular emphasis on the major clades of the Papilionoideae.

MATERIALS AND METHODS

Taxon sampling—Complete *matK* gene sequences from 330 taxa were included in this study, representing 235 genera of legumes as recognized by Polhill (1994) and four outgroup taxa from Fabales (*Polygala*, *Suriana*, *Quillaja*) and Rosales (*Vauquelinia*). Sampling was most extensive in papilionoids (Papilionoideae), including representatives from 29 of the 30 tribes and 179 of the 451 genera. In contrast, 28 of 151 genera (four of four tribes) of caesalpinioids (Caesalpinioideae) and 28 of 76 genera (four of five tribes) of mimosoids (Mimosoideae) were sampled. Representatives of segregate genera *Cercidium* Tul. (*Parkinsonia*), *Brachypterum* Benth. and *Paraderris* (Miq.) R. Geesink (*Derris*), *Poissonia* Baill. (*Coursetia*), *Philenoptera* Benth. (*Lonchocarpus*), *Calia* Teran & Berland (*Sophora*), and the newly described *Maraniona* (Hughes et al., 2004) were included. This study samples extensively in traditionally circumscribed tribes Aeschynomeneae (17/26 genera), Dalbergieae (15/17), all genera of Amorpheae (8), Robinieae (12), most genera of Trifolieae (6/7), and Vicieae (4/5). Representatives of only two, monogeneric tribes, Mimozygantheae Burkart (Mimosoideae) and Euchresteae (Nakai) Ohashi (Papilionoideae), were not sampled for this analysis. Appropriate outgroup taxa from Polygalaceae, Surianaceae, Quillajaceae, and Rosaceae were chosen based on results of recent molecular phylogenetic studies of eurosids using *rbcL-atpB*-18S nuclear ribosomal DNA (Soltis et al., 2000), *rbcL* alone (Soltis et al., 1995; Kajita et al., 2001), *matK* (Steele et al., 2000), and the *trnL-F* region (Persson, 2001).

Sequences from 140 taxa are formally reported here for the first time, complete with voucher specimen and database accession information, although a few of them have been used in part for phylogenetic analyses presented previously (Wojciechowski et al., 2000). Papers by Hu et al. (2000), Lavin et al. (2001, 2003), Luckow et al. (2003), Miller et al. (2003), McMahon and Hufford (2004), Steele and Wojciechowski (2003), Riley-Hulting et al. (2004), and Thulin et al. (in press) provide sampling information for approximately 190 *matK* sequences from subgroups of the taxa included here and should be consulted for more details. The sources of plant material and GenBank accession numbers for *matK* sequences from all taxa included in this paper are provided in the Appendix (see Supplemental Data accompanying the online version of this article).

DNA sequence data—The data presented here were gathered in our laboratories using similar methods. Genomic DNAs were isolated from field-collected, greenhouse-grown plants, silica-dried and herbarium material using the procedure of Doyle and Doyle (1987) or using DNeasy Plant Minikits (Qiagen, Valencia, California, USA). Polymerase chain reaction (PCR) amplifications were performed using Taq and Platinum Taq DNA polymerases (Life Technologies, Gaithersburg, Maryland, USA) as described previously (Wojciechowski et al., 1999; Lavin et al., 2000). For most of the newly sequenced taxa, double-stranded copies of the *matK* gene and the flanking 3′ *trnK* intron region were amplified using primers trnK685F and trnK2R*; typical reaction conditions were 2 min at 95°C for denaturation, followed by 35 cycles of 30 s at 95°C, 30–60 s at 55–57° C for annealing, 2 min 30 s at 72°C for primer extension, then followed by a final 7 min incubation at 72°C. Amplification products were purified and then sequenced using these same primers and others listed in Table 1. DNA sequencing was performed on Applied Biosystems 377 and 3100 sequencers (Applied Biosystems, Foster City, California, USA) at the University of California (DBS Sequencing Facility, Davis, California, USA), Iowa State University (DNA Sequencing Facility, Ames, Iowa, USA), Northwoods DNA (Becida, Minnesota, USA), and Arizona State University (DNA Laboratory, Tempe, Arizona, USA). Sequencer output files were assembled into contigs and edited using the program Sequencher 4.1 (GeneCodes, Ann Arbor, Michigan, USA) before alignment.

Primers for the PCR amplification and sequencing of the *trnK*/*matK* region from legumes (Table 1) were originally designed by one of us (M. F. Wojciechowski) using published primer sequences (Steele and Vilgalys, 1994; Johnson and Soltis, 1995), which were modified based on the sole legume *matK*

TABLE 1. Sequences of oligonucleotide primers used for PCR amplification and sequencing of the plastid *matK* gene in legumes. Sequences given are all 5′ to 3′; forward and reverse refer to direction with respect to *matK* coding sequence.

Primer	Sequence
trnK685F (forward)	GTATCGCACTATGTATCATTTGA
matK4La (forward)	CCTTCGATACTGGGTGAAAGAT
matK1100L (forward)	TTCAGTGGTACGGAGTCAAATG
matK1932Ra (reverse)	CCAGACCGGCTTACTAATGGG
matK832R (reverse)	TTGCATAGAAATAGATTCGCTCAAA
trnK2R* (reverse)	CCCGGAACTAGTCGGATGG

sequence available at the time (*Pisum sativum*; Boyer and Mullett, 1988); various primers have been subsequently modified further to work more specifically with certain groups of taxa (e.g., Lavin et al., 2000; Riley-Hulting et al., 2004). Use of these primers generally resulted in 100% overlap in bidirectional sequencing of the entire *matK* gene, and most of the 3′ flanking *trnK* intron sequence, from these taxa.

The *matK* sequences were initially assembled into a data matrix by first translating a small set of representative DNA sequences (25 taxa) to their corresponding amino acid sequences, which were then aligned using ClustalX (Thompson et al., 1997) using standard pairwise and multiple alignment parameters settings (default gap penalty parameters and Gonnet weight matrix). The amino acid alignment was then used as a template to align the corresponding DNA sequences with insertions and deletions (indels) at equivalent positions. The remaining DNA sequences were added to this primary alignment as they became available, and the data matrix was manually readjusted as necessary to allow for additional indels. Indels were numerous but none were ambiguous with respect to their placement in the final aligned data matrix (see below). All sequences were essentially complete except for 20 taxa that had 50 or more nucleotides missing at one of the ends. All new sequences have been deposited in GenBank, and the final data matrix has been deposited in TreeBASE, study accession S1968 (http://www.treebase.org/).

Phylogenetic analyses—Phylogenetic analyses utilized maximum parsimony (MP) and Bayesian approaches. All parsimony analyses were performed using *PAUP** (version 4.0b10; Swofford, 2002). Multiple tree searches were conducted using heuristic search options that included SIMPLE, CLOSEST, or RANDOM addition sequences (1000 replicates) holding five trees per replicate, and tree bisection-reconnection (TBR) branch swapping, with retention of multiple parsimonious trees (MAXTREES = 5000–10 000). A nucleotide substitution model was selected using AIC implemented in Modeltest (version 3.06; Posada and Crandall, 1998). Bayesian analyses were performed using MrBayes version 3.0 (Huelsenbeck and Ronquist, 2001). Multiple Metropolis-coupled Markov chain Monte Carlo analyses were run with random or user-defined starting points for each run. Parameters for the Akaike information criteria (AIC)-selected GTR + Γ + I model were estimated using the default value of four Markov chains and the ''temperature'' parameter set to 0.2. Markov chains of 2 000 000 generations each were sampled every 100–10 000 generations, which was sufficient to distinguish the burn-in from stationarity phase. Log likelihood values for sampled trees stabilized after approximately 200 000 generations. Clade support was assessed using both nonparametric bootstrap resampling (Felsenstein, 1985) and Bayesian posterior probabilities (Huelsenbeck et al., 2002). Nonparametric bootstrap proportions were estimated from 100 to 500 bootstrap replicates incorporating heuristic parsimony searches using addition sequence and branch-swapping options as in our standard parsimony analyses. Bayesian posterior probabilities were estimated as the proportion of trees sampled after ''burn-in'' that contained each of the observed bipartitions. Although posterior probabilities may be over-credible as measures of clade support (Suzuki et al., 2002) and currently are a subject of debate (Wilcox et al., 2002; Alfaro et al., 2003), we have observed that high Bayesian posterior probabilities often support nodes that are otherwise also detected by parsimony strict consensus.

Thirty-seven indels of 1–4 amino acids each were identified in the complete data set, corresponding to 219 nucleotide positions. These sites, and a 25-nucleotide region surrounding a primer-binding site (positions 541–565) that was missing in about 80 taxa, were excluded from all analyses (244 total characters). Each of the 37 indels was treated as a separate character, and states were scored according to the presence or absence (1 or 0) of a sequence within an indel region (1711 total characters). Sensitivity analyses (cf. Whiting et al., 1997) involving the inclusion or exclusion of these recoded indel characters were performed to determine the effect on tree topology and clade support.

RESULTS

Characteristics of the matK sequences—The *matK* gene in legumes ranges from 1476 bases (492 amino acids) in *Erythrostemon gilliesii* to 1545 bases (515 amino acids) in length in several dalbergioid taxa (e.g., *Adesmia* and *Pictetia*). The final *matK* data set includes 1674 aligned positions with 1042 potentially parsimony-informative characters (73%, which excludes the 244 indels and missing characters) among the 330 taxa analyzed. Of the total 552 420 characters in the data set, missing data accounted for 1.3% while indels and other excluded positions accounted for 14.6% (80 500 bases). By comparison, the complete legume *rbcL* data set (Kajita et al., 2001) is 1404 aligned bases in length (with no indels) but contains fewer potentially informative characters (530 among the 319 sequences). That study, with a total of 242 sequences sampled from 194 legume genera and a similar emphasis on Papilionoideae but with fewer representatives from the other two subfamilies (24 Caesalpinioideae, 6 Mimosoideae) than the *matK* data set, contains sequences from a large number of taxa that are identical or very closely related to those in the *matK* data set. Comparative analyses of these genes (M. Lavin, P. S. Herendeen, and M. F. Wojciechowski, unpublished manuscript) show levels of sequence divergence up to ninefold lower for *rbcL* than for *matK*, and substitutions are distributed less uniformly among the three codon positions in *rbcL*. For example, the substitution rate at the third codon position in *rbcL* is 10 times that of the second position, whereas the third position in *matK* shows twice the rate of the second position. Previous comparisons of rates of substitution in *matK* vs. *rbcL* have yielded similar patterns of variation in other angiosperm groups (e.g., Steele and Vilgalys, 1994; Manos and Steele, 1997).

Within Fabales, pairwise distances (calculated across all sites as uncorrected *p* values in *PAUP**) in *matK* sequences ranged from a maximum of nearly 17% among outgroups and caesalpinioids, to nearly 7% among mimosoids, 12% among caesalpinioids, and just over 19% among papilionoids. Within monophyletic genera for which we have more than three species sampled, pairwise distances varied between 0.8–1.9% in *Astragalus* and 0.9–4.3% in *Sesbania*. Within major papilionoid clades, pairwise distances varied from 0.1% to almost 11% within the genistoid clade, 0.3–12% in Loteae + Robinieae; 0.0% (in *Lens*) to nearly 11% in the IRLC; 0.2–13% in dalbergioids, and 1.4–17.9% in the millettioids.

Phylogenetic reconstruction—Multiple heuristic searches of 1042 parsimony informative nucleotide characters, excluding indels, consistently converged on a large number of equally most parsimonious trees (maximum set saved = 10 000) with a minimum length of 8397 steps, a consistency index (CI) of 0.288 excluding uninformative characters, and a retention index (RI) of 0.791. The strict consensus tree of 5000 representative equally most parsimonious (MP) trees is highly re-

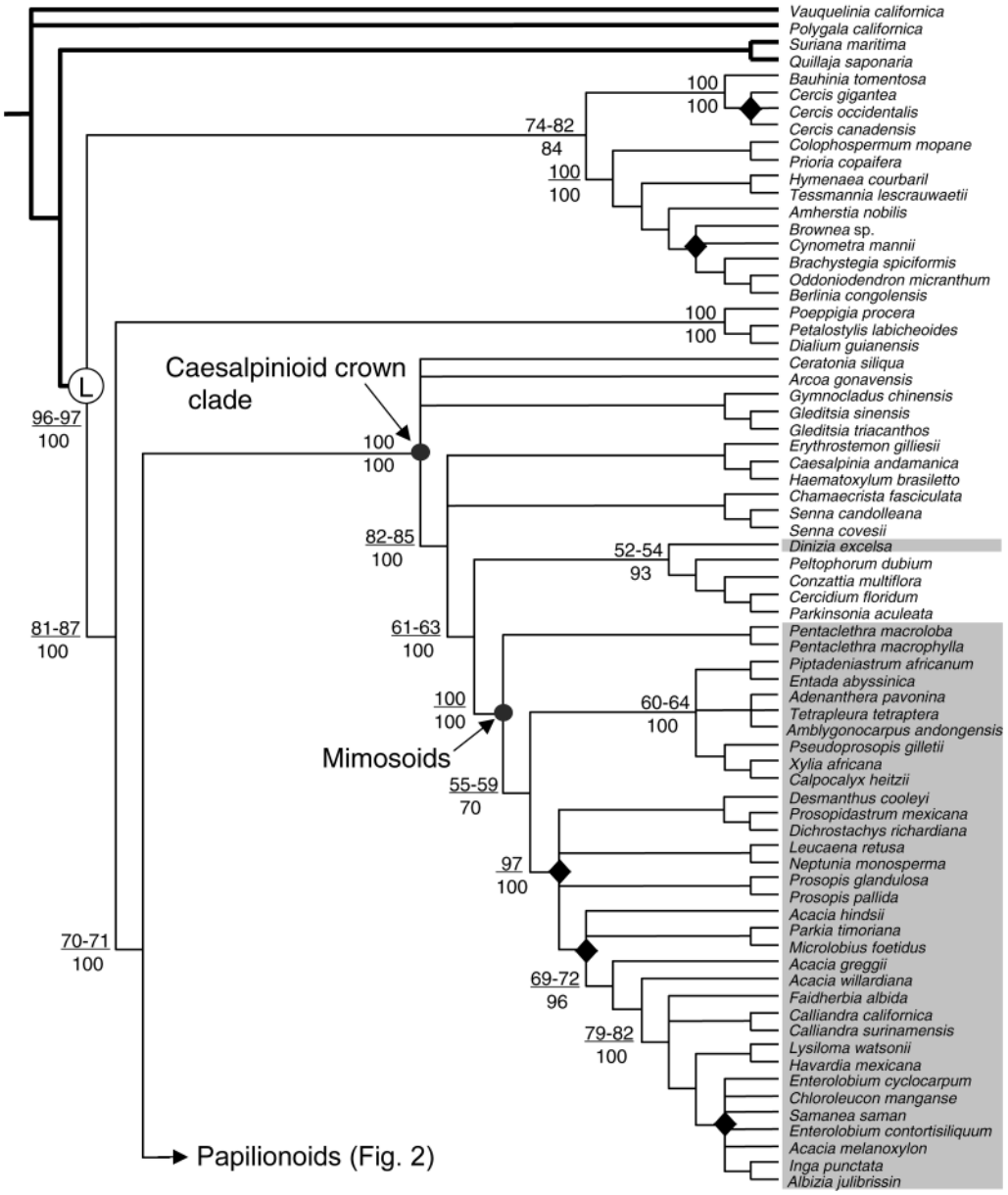

Fig. 1. Phylogeny of Leguminosae based on parsimony analyses of plastid *matK* gene sequences. Phylogenetic relationships among the three subfamilies of Leguminosae (crown clade, ''L''), and within subfamilies Caesalpinioideae and Mimosoideae. Outgroup lineages are indicated by bold lines. Members of Mimosoideae are indicated by gray boxes. Tree shown is strict consensus of 5000 equally most parsimonious trees (length = 8397 steps, consistency index = 0.288, retention index = 0.791) derived from heuristic search analyses of 330 *matK* sequences. Nodes designated by a diamond were not resolved in a 50% majority-rule consensus of the same set of 5000 equally most parsimonious trees. Nonparametric bootstrap proportions and Bayesian posterior probabilities from separate analyses (individual or range) are indicated above and below branches or immediately to left of appropriate node, respectively. Values are given for most nodes for which support values from both analyses were greater than 50%. *Vauquelinia* (Rosaceae) was designated as the outgroup for all analyses. Major papilionoid clades informally named here are indicated by a filled circle.

solved and presented in Figs. 1–5. The semi-strict consensus of the same set of most parsimonious trees resolves only three nodes that are not present in the strict consensus, while in the 50% majority-rule tree a total of only seven nodes were not fully resolved. Branching order and support values for the major clades of legumes resolved by these *matK* data were very similar in the maximum parsimony and Bayesian analyses (Figs. 1–6). To illustrate the heterogeneity in estimated branch lengths, a phylogram representation of a typical Bayesian tree (sampled post burn-in) is shown in Fig. 6.

Analysis of the *matK* data confirms results from earlier studies in that the family is a monophyletic group, papilionoids and mimosoids, excluding *Dinizia* (tribe Mimoseae), are monophyletic and nested within a paraphyletic Caesalpinioideae. All mimosoids and the majority of the caesalpinioid tribes Caesalpinieae and Cassieae comprise a strongly supported clade (Fig. 1) that is the sister group to papilionoids. Seven major clades and a number of minor clades within papilionoids are also highly supported (Figs. 2–5). In spite of this, relationships among certain of the clades, especially the gen-

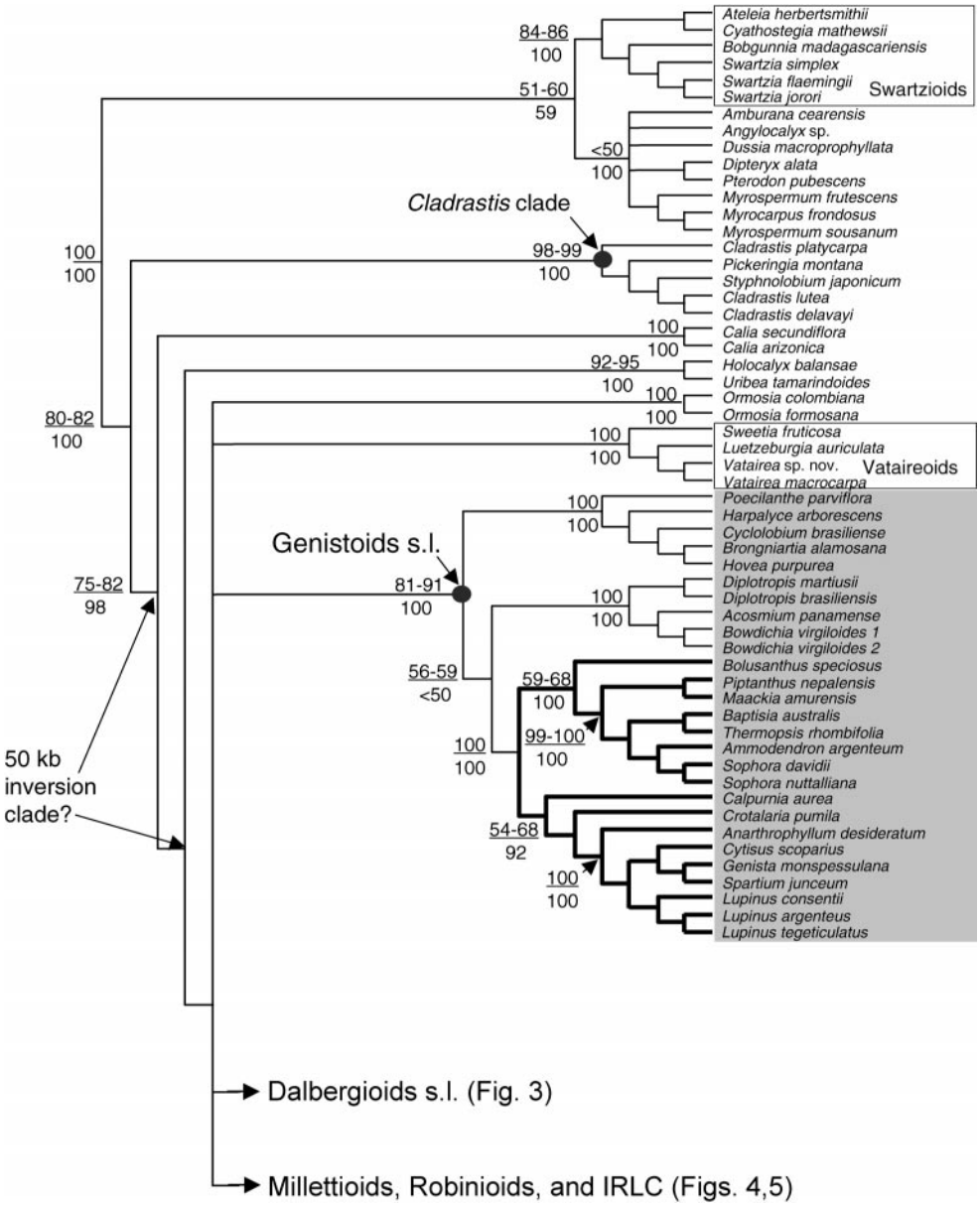

Fig. 2. Phylogenetic relationships of the Genistoid sensu lato clade, as well as other papilionoids including many of those once placed in the tribes Swartzieae and Sophoreae. The ''core genistoids'' (sensu Crisp et al., 2000) clade is indicated by bold lines. Nodes consistent with the presence of a 50-kb inversion in the plastid DNA genome are indicated by arrows. See Fig. 1 for details.

istoid s.l. and the dalbergioid s.l. clades, remain unresolved. Parsimony analyses suggest the dalbergioid s.l. clade branches before the genistoid s.l. clade, whereas Bayesian analyses suggest the genistoid s.l. clade is the sister group to the dalbergioid s.l. clade plus the remaining papilionoids (i.e., *Baphia* clade, mirbelioids, millettioids, and Hologalegina).

Of the 37 indel characters, 12 were synapomorphic for clades identified in the maximum parsimony and Bayesian analyses. For example, two single-amino-acid insertions, one at positions 421–423 and a second at positions 1498–1500, were synapomorphies for the papilionoid clade. Similarly, one-, two-, or three-amino-acid insertions/deletions uniquely mark each of the *Sweetia-Vatairea* clade (Fig. 2), New World *Lupinus* (Fig. 2), dalbergioid s.l. clade (Fig. 3), the genus *Sesbania* (Fig. 5), and the *Caragana* plus *Hedysarum* clade (Fig. 5). Inclusion of the indels as additional characters had little effect on phylogenetic relationships, based on comparison of the strict consensus topology derived from analysis of the data set containing the indel characters (data not shown) to that presented in Figs. 1–5. The exception involved *Platycyamus regnellii* which was resolved as sister to the clade defined by the MRCA of *Apios americana* and *Phaseolus vulgaris* (Fig. 4), in analyses that included the indel characters. Likewise, addition of the indel characters had little effect on bootstrap proportions for nodes receiving support in the 50% bootstrap consensus tree (<5% difference; data not shown).

Phylogenetic criteria for papilionoid clade nomenclature—We have used four criteria for recognizing and informally naming major clades within the Leguminosae, which are

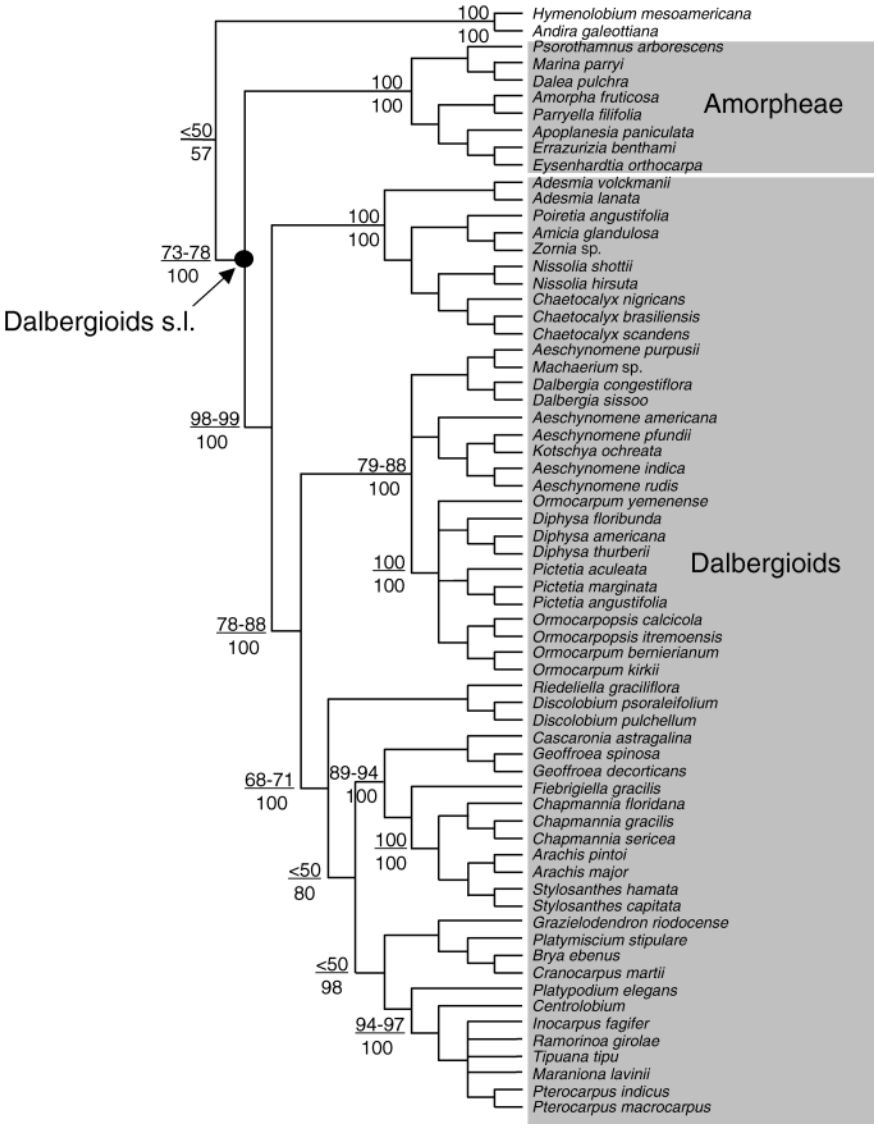

Fig. 3. Phylogenetic relationships in Dalbergieae sensu lato clade: Amorpheae and Dalbergioid subclades. See Fig. 1 for details.

consistent with formal node-based definitions under a system of phylogenetic nomenclature (de Queiroz and Gauthier, 1994). First, groups are resolved as monophyletic in strict consensus analyses. Second, bootstrap proportions greater than 70% support the clade of interest. Third, taxonomic sampling within the clade is diverse and/or extensive. Lastly, results are at least approximately congruent with that obtained by other studies (i.e., in showing support for clades that correspond to those informally recognized here). The following clade names are used throughout the discussion. The ''Caesalpinioid crown'' clade includes all the Mimosoideae and members of tribes Caesalpinieae and Cassieae of subfamily Caesalpinioideae that form the sister group to them and is defined as the least inclusive clade that contains *Ceratonia siliqua* and *Albizia julibrissin*. The ''papilionoid'' clade is equivalent to the subfamily Papilionoideae and is delimited by the most recent common ancestor (MRCA) of *Swartzia simplex* and *Vicia faba*. Within papilionoids, the ''*Cladrastis*'' clade is delimited by the MRCA of *Cladrastis platycarpa* and *Cladrastis lutea*, which renders the genus *Cladrastis* paraphyletic with respect to *Pickeringia* and *Styphnolobium* (Fig. 2). The ''genistoid s.l.'' clade is delimited by the MRCA of *Poecilanthe parviflora* and *Lupinus argenteus* (Fig. 2). The ''dalbergioid s.l.'' clade comprises all descendants of the MRCA of *Amorpha fruticosa* and *Pterocarpus indicus* (Fig. 3). The ''mirbelioid'' clade is delimited by the MRCA of *Daviesia latifolia* and *Gompholobium minus* (Fig. 4). The ''millettioid'' clade is delimited by the MRCA of *Xeroderris stuhlmannii* and *Phaseolus vulgaris* (Fig. 4). The ''robinioid'' clade comprises all descendents of the MRCA of *Robinia pseudoacacia* and *Lotus japonicus* (Fig. 5). The ''inverted repeat-lacking'' clade (IRLC) is delimited by the MRCA of *Glycyrrhiza lepidota* and *Vicia faba* (Fig. 5). The robinioids and IRLC are sister groups and comprise ''Hologalegina.''

DISCUSSION

The phylogenetic analyses of the legume *matK* sequences presented here achieve our main goal of reconstructing a robust molecular phylogeny for the Leguminosae, with the par-

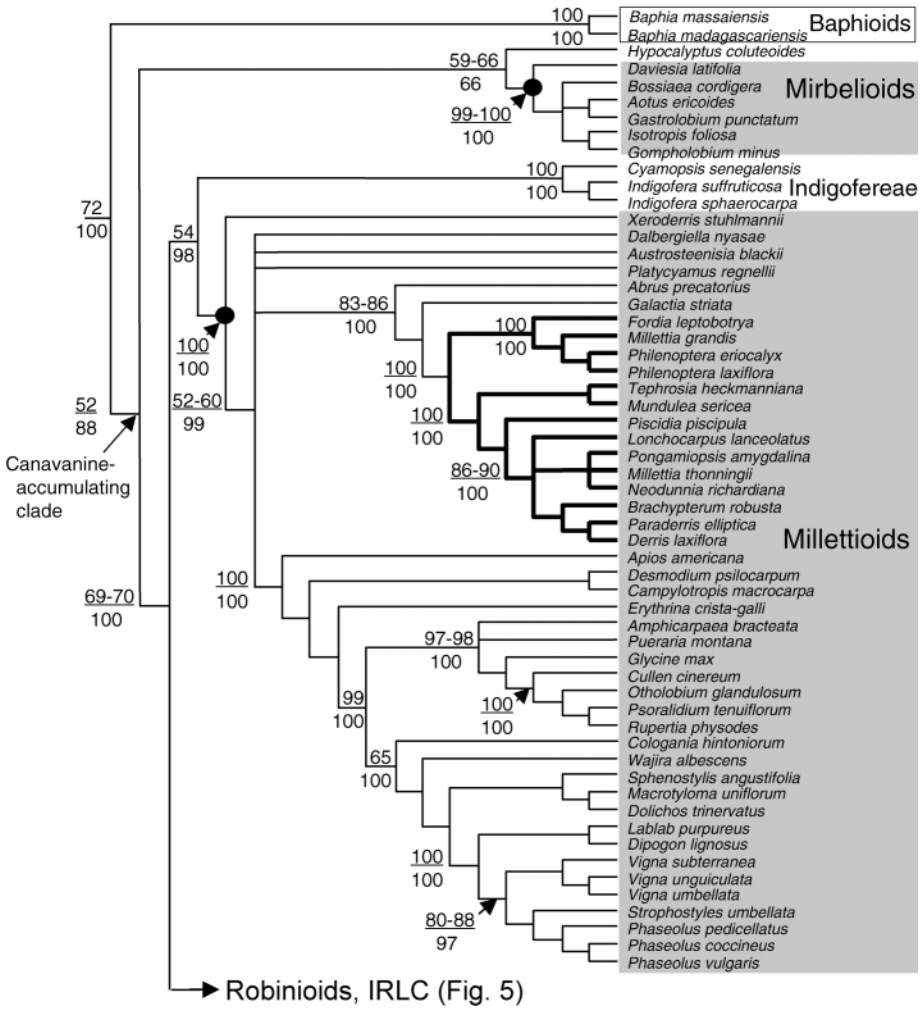

Fig. 4. Phylogenetic relationships in Millettioid clade, Indigofereae, Hypocalypteae, Mirbelioid, and Baphioid clades; memberships are indicated by boxes. The ''core Millettieae'' clade (sensu Hu et al., 2000) is indicated by bold lines. Arrows are used to specify nodes with indicated support values. See Fig. 1 for details.

ticular finding that the identity and inter-relationships of many clades within papilionoids are for the first time well resolved. The monophyly of the family and details of relationships among the outgroups, as well as the identity and relationships among the various caesalpinioid and mimosoid subgroups are only briefly addressed in this study as they are the subject of intensive, ongoing investigation by others (e.g., Luckow et al., 2000, 2003; Bruneau et al., 2001; Herendeen et al., 2003a, b). Regardless, when combined with results from these recent studies, this study reveals much promise for *matK* sequences in resolving all of the major clades within Fabales.

Consistent with recent results using the *rbcL* gene (Kajita et al., 2001) and *trnL* intron (Bruneau et al., 2001; Herendeen et al., 2003a) sequences, analyses of *matK* sequences support the monophyly of the family Leguminosae (Fig. 1) and the paraphyly of subfamily Caesalpinioideae. Even with our limited sampling of caesalpinioid genera, at least four well-supported clades emerge from this analysis that correspond to major caesalpinioid clades detected by analyses of *trnL* intron sequences alone (Bruneau et al., 2001), *trnL* sequences combined with morphological characters (Herendeen et al., 2003a, b), or *rbcL* sequences (Kajita et al., 2001). The well-supported clade represented by *Colophospermum*, *Hymenaea*, and *Berlinia* (Fig. 1) corresponds to the ''Detarieae s.l.'' clade of Herendeen et al. (2003a; see their fig. 3). Our finding of a sister relationship of this clade to that comprising *Bauhinia* and *Cercis* (the tribe Cercideae clade) is unexpected yet interesting; it may be an artifact of poor sampling among these caesalpinioid taxa but to our knowledge has not been observed in any previous study. The genera *Petalostylis* and *Poeppigia* form a well-supported clade that corresponds to the ''Dialiinae s.l.'' clade of Herendeen et al. (2003a). The sister relationship of this latter clade to all remaining Leguminosae (i.e., papilionoids, mimosoids, and closely related caesalpinioids) is also detected by these earlier analyses. A well-supported clade including the remaining caesalpinioids, all mimosoids and papilionoids corresponds to ''clade A'' of Bruneau et al. (2001; see their fig. 6), while the *Umtiza* + *Caesalpinia* + Mimosoideae clade of Herendeen et al. (2003a; see their fig. 2), is well supported with *matK* sequences (Fig. 1). We refer to this clade as the Caesalpinioid crown. The *Umtiza* subclade (Herendeen et al., 2003b) is sister to the rest of the Caesalpinioid crown clade and is here represented by *Arcoa*, *Ceratonia*, *Gymnocladus*, and *Gleditsia*. This subclade is not supported

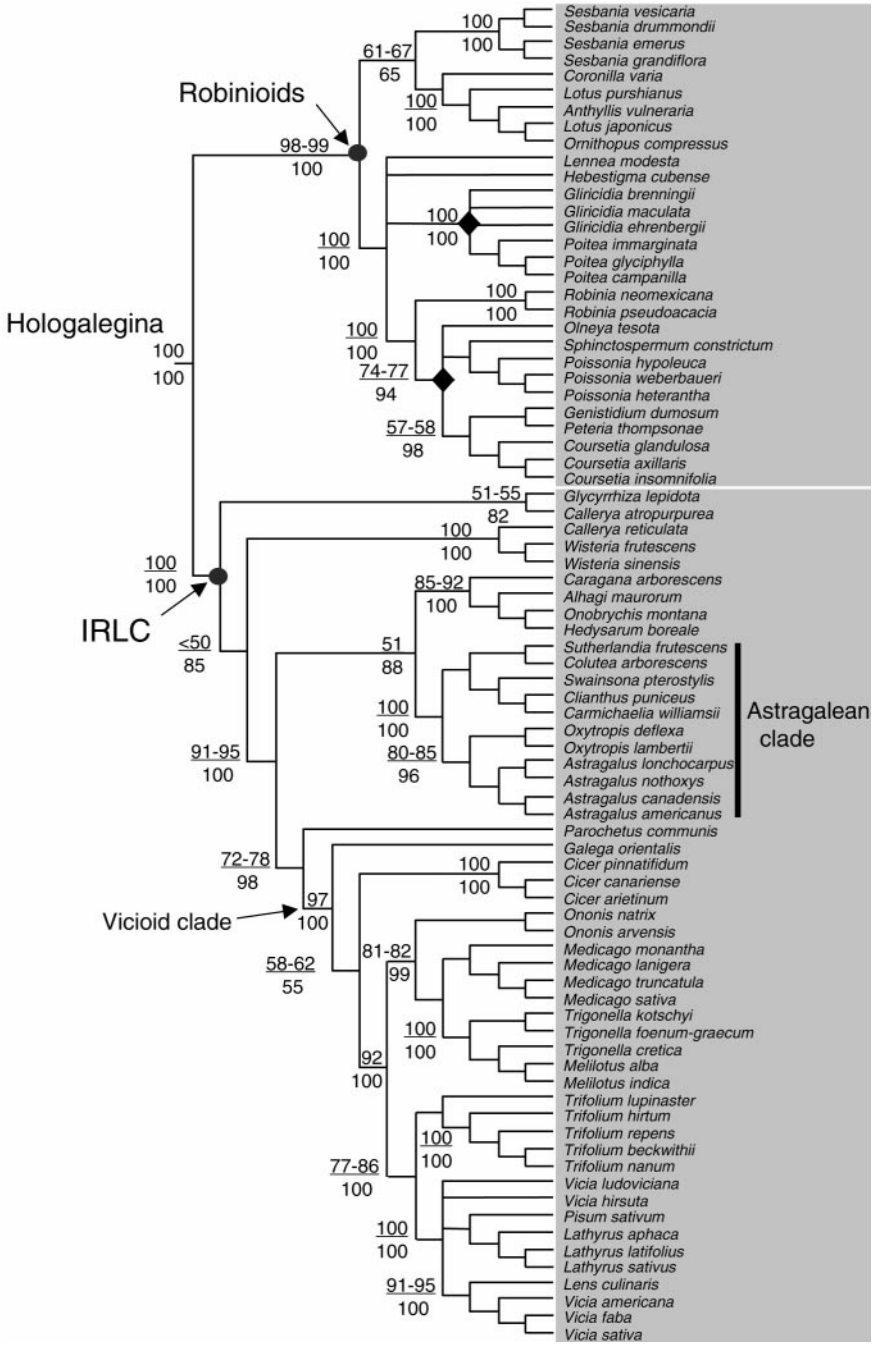

Fig. 5. Phylogenetic relationships in Hologalegina clade: the Robinioid clade and the IRLC. See Fig. 1 for details.

by parsimony bootstrap (<50%) in our analyses, however. Regardless, the constituents and relationships of the *Umtiza* subclade within the Caesalpinioid crown clade detected in this study are in agreement with and further substantiate the findings of Herendeen et al. (2003a, b).

Although sampling of taxa from Mimosoideae was limited, our results generally agree well with those of Luckow et al. (2003), especially with respect to the monophyly of at least the vast majority of the genera traditionally assigned to the Mimosoideae and paraphyly of constituent tribes. With the exception of *Dinizia*, which appears more closely related to certain caesalpinioids than to mimosoids on the basis of morphological and molecular evidence, as concluded by Luckow et al. (2003), the rest of the taxa sampled from this subfamily are resolved as monophyletic with high support in our analyses (Fig. 1). Furthermore, our results clearly show *Piptadeniastrum* nested within the mimosoid clade, confirming recent results by Luckow et al. (2003). The mimosoid clade generally shows poorly resolved relationships or at least short internal branch lengths compared to other clades of Leguminosae (Fig. 6). This suggests either a slow down in the rate of substitution or a relatively recent diversification of most of the extant mem-

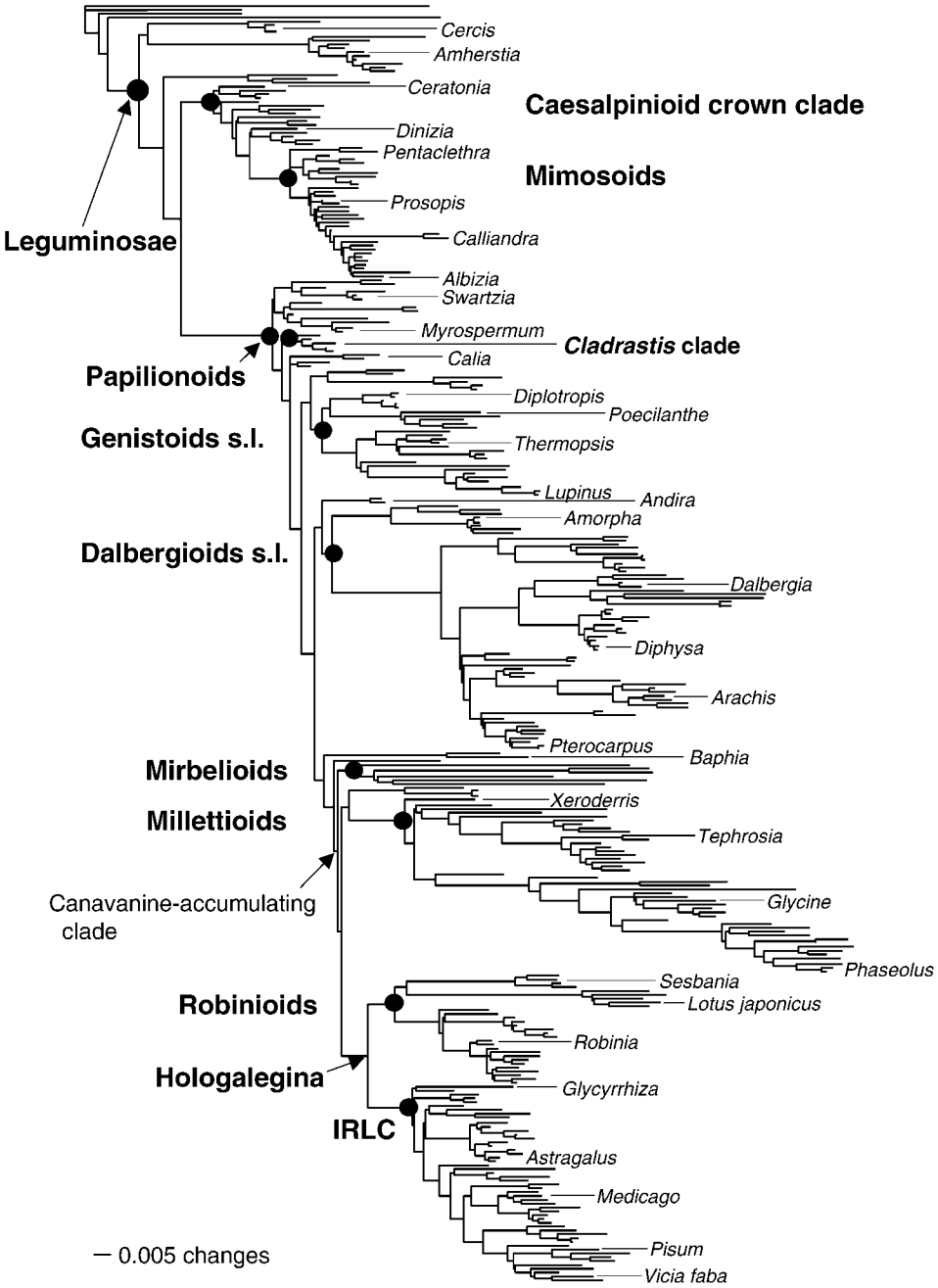

Fig. 6. Representative Bayesian tree sampled according to posterior probabilities from an analysis of 330 legume *matK* sequences. Estimated branch lengths (under the GTR + Γ + I model) are shown; scale is indicated at bottom. Major subclades of Leguminosae, and representative taxa, are indicated by filled circles.

bers of mimosoids. These alternative hypotheses are being addressed elsewhere (Lavin et al., in press).

In contrast to the caesalpinioids and mimosoids, our results have significant implications with regard to papilionoid phylogenetics. In all our analyses, papilionoids are strongly supported as monophyletic (Fig. 1) compared to previous *rbcL* studies where papilionoids were resolved as monophyletic but with relatively low statistical support (e.g., Kajita et al., 2001; 57% bootstrap and 62% parsimony jackknife). Similar to the findings of other studies involving broad sampling of caesalpinioid legumes (e.g., Herendeen et al., 2003a), papilionoids are not resolved as sister to an isolated caesalpinioid lineage, as are the mimosoids, but rather are nested among the major caesalpinioid clades as an early branch in the legume phylogeny (Fig. 1). In marked contrast to the most recent *rbcL* analysis in which most major clades within papilionoids were weakly resolved (fig. 3 of Kajita et al., 2001), the *matK* strict consensus is very highly resolved (Figs. 2–5). Furthermore, both bootstrap proportions and Bayesian posterior probabilities for the major subclades often exceed 95%. The results presented here provide some of the best evidence to date in support of relationships among the major papilionoid subclades,

which heretofore have been largely unresolved by cladistic analyses of DNA sequences data.

Consistent with the results of Doyle et al. (1997) and Pennington et al. (2001), the *matK* phylogeny resolves certain representatives of Swartzieae and Sophoreae as the sister group to the rest the subfamily. The clade that forms the sister group to all remaining papilionoids, here delimited by the MRCA of *Swartzia simplex* and *Myrospermum sousanum* (Fig. 2), is unexpected in that it now includes representatives of a number of disparate lineages such as *Angylocalyx* and Dipterygeae (*Dipteryx* and *Pterodon*) that had been poorly resolved or supported in previous studies (e.g., Pennington et al., 2001). One of two subclades of this clade includes *Swartzia* and recent segregate *Bobgunnia*, *Ateleia*, and *Cyathostegia*, and corresponds to the ''swartzioid'' clade of Ireland et al. (2000) and Pennington et al. (2001). The other contains *Amburana*, *Angylocalyx*, *Dipterygeae*, *Dussia*, *Myrocarpus*, and *Myrospermum*. The resolution of this larger clade of morphologically eclectic genera as sister to the rest of Papilionoideae suggests that the swartzioid clade of Pennington et al. (2001) could be expanded to encompass the majority of papilionoid genera that lack the 50-kb inversion in the plastid DNA genome.

With respect to the rest of the papilionoid subgroups, our sampling is much more extensive. The following seven well-supported clades resolved in this study are thus validated not only by extensive sampling, but also by the resolution of these subclades in other recent studies. These seven are the *Cladrastis* clade, the genistoid s.l., the dalbergioid s.l., the mirbelioids, the millettioids, the robinioids, and the inverted-repeat-lacking clades, the last two of which comprise Hologalegina. Even if resolved by previous studies, relationships among these major papilionoid subclades have been heretofore resolved at best with only weak support (e.g., Hu et al., 2000; Kajita et al., 2001; Lavin et al., 2001; Pennington et al., 2001).

The Cladrastis clade—The genera *Cladrastis* and *Styphnolobium* traditionally have been classified in Sophoreae s.s. (Polhill, 1981b) whereas *Pickeringia* Nuttall has been classified in tribe Thermopsideae (Turner, 1981). These three genera form a well-supported clade in all our analyses. While a sister group relationship of *Cladrastis* and *Styphnolobium* has been observed in previous molecular studies (e.g., Doyle et al., 1997; Pennington et al., 2001) and is notable biogeographically because both genera exhibit East Asian–North American disjunctions, this study is the first to suggest a close relationship of these two genera with *Pickeringia*. *Pickeringia* is a monotypic genus restricted to the sclerophyllous chaparral vegetation of the California Floristic Province of western North America (Raven and Axelrod, 1995). The strongly supported position of this genus in the *Cladrastis* clade confirms Polhill's (1981b) initial prediction and Sousa and Rudd's (1993) subsequent conclusion of a close relationship between these three genera based on floral (bracts at base of inflorescence) and chromosomal similarities ($n = 14$; Goldblatt, 1981; Palomino et al., 1993). This clade is also supported by results from cladistic analyses of nuclear ribosomal DNA internal transcribed spacers (nrDNA ITS) sequence data (M. F. Wojciechowski, unpublished data). This placement of *Pickeringia* reveals that Thermopsideae sensu Yakovlev (Turner, 1981) is not monophyletic, contrary to molecular evidence presented previously (Crisp et al., 2000). Furthermore, the absence of quinolizidine alkaloids in *Pickeringia* of the type that is characteristic of other Thermopsideae (Turner, 1981) is consistent with the *matK* results. The presence of quinolizidine alkaloids, a prominent group of secondary metabolites once considered to be widely distributed in papilionoid legumes (Kinghorn and Balandrin, 1984), now appear to be of systematic significance only for the ''genistoids'' (see next paragraph). In a recent analysis (Kite and Pennington, 2003), the failure to detect similar alkaloids in extracts of *Cladrastis* and *Styphnolobium* is also in accordance with the phylogenetic position of these taxa based on *trnL* sequence data (Pennington et al., 2001) and the result presented here. The disjunct distribution of the *Cladrastis* clade in warm temperate to tropical regions of the Northern Hemisphere is common to many other legume groups (e.g., *Gleditsia*, *Gymnocladus*, *Desmodium*, *Lespedeza*, etc.; Schrire et al., in press).

The genistoid s.l. clade—The genistoids include the many genera traditionally classified in the tribes Genisteae, Thermopsideae, Euchresteae, Crotalarieae, Liparieae, Podalyrieae, and Sophoreae s.s. (Käss and Wink, 1997; Crisp et al., 2000). The concept of a ''genistoid alliance'' was circumscribed by Polhill (1981a, 1994) who brought together for the first time this group of putatively related, predominantly Southern Hemisphere tribes that have been considered relatively isolated among quinolizidine-alkaloid-accumulating papilionoid legumes. The alliance as recognized by Polhill comprises four separate lineages. One includes the predominantly Northern Hemisphere Genisteae sensu stricto (s.s.), Euchresteae, and Thermopsideae together with certain Sophoreae (*Sophora* group). A second involves the mainly southern African Crotalarieae, Liparieae, and Podalyrieae, and now segregate tribe Hypocalypteae. A third comprises the endemic Australasian Bossiaeeae and Mirbelieae. The fourth includes the Neotropical–Australian Brongniartieae (including the *Templetonia* group). Early molecular phylogenetic analyses by Käss and Wink (1996, 1997) suggested that most species of the alliance formed a monophyletic group with some certain members of Sophoreae (i.e., some but not all species of *Maackia* Rupr. & Maxim. and *Sophora* L.) near the base of papilionoids. The analysis of Doyle et al. (1997) suggested the genistoids were polyphyletic and formed three clades, the largest of which approximates the genistoid clade of Käss and Wink. The monogeneric Euchresteae (*Euchresta*) has been shown by subsequent analyses (Kajita et al., 2001) to be nested within a *Sophora* s.s. clade, while Liparieae has been formally included within Podalyrieae (Schutte and van Wyk, 1998), a placement verified by analyses of *rbcL* and nrDNA ITS sequences (Käss and Wink, 1997; Kajita et al., 2001; van der Bank et al., 2002).

Crisp et al. (2000) confirmed the polyphyly of the genistoids sensu Polhill (1981a), but suggested that this name be restricted to a well-supported ''core genistoids'' group, from Africa and Eurasia, that comprises the majority of the tribes that made up Polhill's genistoid alliance. This clade is strongly supported by the *matK* data (Fig. 2; corresponds to the clade delimited by the MRCA of *Bolusanthus speciosus* and *Spartium junceum*). In addition, the *matK* phylogeny corroborates results of other studies in resolving a core genistoid clade nested within a larger genistoid clade. A more inclusive group, referred here to as the genistoid s.l. clade and well supported by *matK* sequence analysis (Fig. 2), includes the Brongniartieae (sensu Crisp and Weston, 1987; Thompson et al., 2001), *Poecilanthe* and *Cyclolobium* of Millettieae (Hu et al., 2000, 2002), and a number of largely woody Neotropical genera of Sophoreae such as *Acosmium* Schott, *Bolusanthus* Harms, *Bowdichia*

Kunth, *Cadia* Forssk., *Diplotropis* Bentham, and most likely *Ormosia* Jackson (Kajita et al., 2001), *Dicraeopetalum* Harms, *Clathrotropis* Harms, and *Platycelyphium* Harms (Pennington et al., 2001), several of which have not been sampled for *matK* sequences. The monophyly of the genistoid s.l. clade as defined here is also supported by the taxonomic distribution of quinolizidine alkaloids (e.g., Kinghorn and Balandrin, 1984; van Wyk, 2003). All taxa known to accumulate these alkaloids, with the exception of *Calia* (Kite and Pennington, 2003) and *Ormosia* (Kinghorn and Balandrin, 1984), are members of the genistoid s.l. clade as defined here. While the relationship of these particular taxa to this clade is not definitively resolved by our analyses (Figs. 2, 6), Pennington et al. (2001) did find weak support at least for the inclusion of *Osmosia* within an ''expanded'' genistoids. Further resolution and sampling of these taxa as well as *Holocalyx*, *Uribea*, and the vataireoids, may yet show quinolizidine alkaloids to be a nonmolecular synapomorphy for an expanded genistoid clade.

The dalbergioid s.l. clade—The dalbergioid legumes, a mostly pantropical group of papilionoids, was originally circumscribed by a combined data approach to include 44 genera and ca. 1100 species from the tribes Aeschynomeneae, Adesmieae, subtribe Bryinae of Desmodieae, and tribe Dalbergieae except *Andira*, *Hymenolobium*, *Vatairea*, and *Vataireopsis* (Lavin et al., 2001). In addition, this clade is diagnosed apomorphically by the presence of the aeschynomenoid root nodule (Sprent, 2001). Although the position of the dalbergioid clade within the Papilionoideae was not well resolved or supported in previous studies using *rbcL* and *trnL* (Kajita et al., 2001; Pennington et al., 2001), there was preliminary evidence that its sister group included the predominantly North American temperate tribe Amorpheae (Lavin et al., 2001). Our results, like those of McMahon and Hufford (2004), consistently show Amorpheae as the sole sister clade to the dalbergioid clade even if parsimony bootstrap support for this relationship is moderate (Fig. 3). Thus, the dalbergioid clade (sensu Lavin et al., 2001) is now expanded to encompass this tribe and is herein referred to collectively as the dalbergioid s.l. clade. The similarity in base chromosome number among dalbergioids and genera of Amorpheae, where $x = 10$ is apparently ancestral with derived cases of aneuploidy (e.g., $x = 9$ and $x = 8$; Goldblatt, 1981), supports this decision. Furthermore, the glandular punctate leaves and indehiscent, single-seeded pods (derived from a two-plus-ovulate ovary) of Amorpheae, once thought to indicate a relationship with the genera of Psoraleeae (e.g., Stirton, 1981), are found variously within dalbergioids. The sister relationship of the primarily tropical American *Andira* plus *Hymenolobium* to the dalbergioid s.l. clade is very weakly supported but resolved in both the parsimony strict consensus (Fig. 3) and Bayesian analyses (Fig. 6). This relationship, though consistent with results of Lavin et al. (2001; see their fig. 5) needs to be further investigated with additional sampling of *Andira* and *Hymenolobium* species.

The mirbelioid clade—The endemic Australasian tribes Bossiaeeae and Mirbelieae comprise a fifth clade of c. 31 genera and 750 species within the papilionoids, although *matK* sampling is still quite limited from this group. The analyses of Doyle et al. (1997) and Crisp et al. (2000) provided the first molecular evidence, albeit not well supported by bootstrap analysis, that validated Crisp and Weston's (1987) hypothesis of a monophyletic Mirbelieae-Bossiaeeae group. Recent *trnL* intron and nrDNA ITS sequence analyses (Crisp and Cook, 2003) suggest Bossiaeeae is nested within a paraphyletic Mirbelieae. Although the monophyly of Mirbelieae–Bossiaeeae is well-supported by both bootstrap and Bayesian analyses of *matK* sequences (Fig. 4), neither of these tribes is resolved as monophyletic, consistent with the results of Crisp et al. (2000) and Crisp and Cook (2003).

The *matK* analyses provide the first unequivocal evidence for a sister group relationship between Mirbelieae-Bossiaeeae and the tribe Hypocalypteae (Schutte and van Wyk, 1998). Although this relationship receives only modest support in bootstrap and Bayesian analyses, it is consistently resolved (Fig. 4). A sister group relationship of Hypocalypteae, rather than nested within the Australasian Mirbelieae or Bossiaeeae, is also more consistent with their respective geographic distributions. The taxonomic position of *Hypocalyptus* Thunberg, a genus of three species geographically confined to the Cape region of South Africa, has been uncertain since Bentham (1837), but is historically considered linked to various tribes of the genistoid alliance. A cladistic analysis of morphological and biochemical characters (Schutte and van Wyk, 1998) resulted in the tribal ranking of Hypocalypteae and the determination that it was more closely related phylogenetically to Millettieae. Crisp et al. (2000) however, resolved a sister relationship of Hypocalypteae to Indigofereae using combined *rbcL* and nrDNA ITS sequences, although bootstrap support was lacking. Neither of these alternative positions fundamentally conflicts with the position detected using *matK* because neither finds any support from parsimony bootstrap analyses (Bayesian analyses were not performed in these two studies). Regardless, Hypocalypteae and Mirbelieae-Bossiaeeae collectively, or perhaps individually, occupy a sister position to the rest of the papilionoids that accumulate nonprotein amino acids in seed (i.e., Indigofereae plus the millettioids and the robinioids plus the IRLC; Fig. 4).

The millettioid clade—This clade includes all genera of the tribes Millettieae, Abreae, Phaseoleae, and Psoraleeae, plus Desmodieae subtribes Desmodiinae and Lespedezinae (Lavin et al., 1998; Hu et al., 2000, 2002; Kajita et al., 2001), with Indigofereae as its moderately supported sister group (Fig. 4). The predominantly tropical, woody Old World tribe Millettieae has been considered to be a transitional link from the ''less advanced'' elements of Dalbergieae and Sophoreae to putatively ''more advanced'' Old World tribes like Phaseoleae and Galegeae (Geesink, 1981, 1984; Polhill, 1981a). These authors used the term ''advanced'' to indicate a high degree of fusion of stamens and keel petals, as well as the accumulation of nonprotein amino acids in seeds rather than alkaloids. Geesink (1984) went so far as to consider Millettieae the paraphyletic ''stem'' group from which all other ''advanced'' papilionoids branched. Early *rbcL* analyses (Doyle et al., 1997) suggested the polyphyly of both Millettieae and Phaseoleae, and this was subsequently confirmed using nuclear phytochrome gene sequences (Lavin et al., 1998), *trnK*/*matK* sequences (Hu et al., 2000), and a combined analysis of *rbcL*, *matK*, and nrDNA ITS data (Hu, 2000). Regardless, the emerging pattern of relationships derived from these studies and from this *matK* analysis is that most of the constituent genera of Millettieae and Phaseoleae clade fall out in two very well-supported subclades (Fig. 4). The first, previously referred to as ''core Millettieae'' (Hu et al., 2000), comprises the majority of Millettieae and includes the large genera *Millettia* Wight & Arn. (c.

150 spp.), *Lonchocarpus* Kunth (c. 130 spp.), *Derris* Lour. (50–60 spp.) and *Tephrosia* Pers. (c. 350 spp.), while the majority of Phaseoleae dominates the second. Although shown here (Fig. 4) as sister lineages to core Millettieae, relationships of the monogeneric Abreae and certain Phaseoleae such as *Galactia* and presumably the other genera of subtribes Galactinae and Diocleinae (see Lewis et al., in press), are not yet resolved with certainty with respect to the core Millettieae clade. Additionally, certain genera classified in Millettieae, including *Xeroderris* and *Platycyamus*, have unresolved or weakly supported relationships with respect to the two main millettioid clades.

The large Phaseoleae clade resolved by *matK* includes tribes Desmodieae (except subtribe Bryinae) and Psoraleeae, in agreement with other studies (e.g., Kajita et al., 2001; Hu et al., 2002). Lackey's (1981) subtribal classification of Phaseoleae, the largest tribe of legumes in number of genera, is not congruent with the monophyletic subclades detected in this and other analyses (e.g., Kajita et al., 2001). The delimitation of the subtribe Phaseolinae is restricted to the descendants of the MRCA of *Wajira* and *Phaseolus* (Fig. 4), and excludes other genera once included, such as *Psophocarpus* and *Otoptera*, which are now known to be more closely related to genera in Glycininae (Thulin et al., in press; A. Delgado-Salinas, M. Lavin, M. Thulin, and N. Weeden, unpublished data). In spite of the findings of Lee and Hymowitz (2001), Glycininae is also not monophyletic and includes certain genera of subtribe Phaseolinae and probably all genera of tribe Psoraleeae (Kajita et al., 2001; A. Delgado-Salinas, M. Lavin, M. Thulin, and N. Weeden, unpublished data).

In this and most previous analyses (e.g., Kajita et al., 2001), tribe Indigofereae emerges as the moderately supported sister group to the millettioid clade (Fig. 4). The genera of Indigofereae have an inflorescence of the simple racemose type (each node bearing a single flower), whereas the tribes Abreae, Desmodieae, Millettieae, Phaseoleae, and Psoraleeae, which form the millettioid clade, share an unusual type of inflorescence, the pseudoraceme (see Tucker, 1987a, b). Indeed, the pseudoracemose inflorescence is found only and in all members of the millettioid clade, rendering it the most readily morphologically distinguished of the newly circumscribed large subclades of Papilionoideae. In addition, typical chromosome numbers in Indigofereae ($n = 8, 7, 6$; Goldblatt, 1981) are different from that which predominate in the tribes comprising the millettioids (e.g., Millettieae, Phaseoleae; $n = 11, 12$; Goldblatt, 1981). Whether the millettioid clade should be expanded to include the Indigofereae is as yet undecided, especially given its sister relationship is only moderately supported, but the lack of any nonmolecular evidence to support such a relationship argues against this.

The robinioid clade—Tribes Loteae (including Coronilleae; sensu Polhill, 1994) and Robinieae (sensu Lavin and Sousa, 1995) comprise an expanded robinioid clade, which is distributed primarily in the Northern Hemisphere of the New World, Europe, and Africa. Within this clade, a monophyletic *Sesbania* L. (Robinieae) is weakly supported as sister to Loteae and these collectively form the sister group to the remaining members of Robinieae (Fig. 5). This position of *Sesbania* thus renders Robinieae paraphyletic, a finding first suggested by a preliminary phylogenetic analysis of *matK* sequences (Wojciechowski et al., 2000) and confirmed with a recent study utilizing exhaustive sampling for *matK*, *trnL* intron, and nrDNA ITS sequences (Lavin et al., 2003). Allan et al. (2003) have provided molecular evidence for the monophyly of the largely Eurasian–North American Loteae and the paraphyly of the large genus *Lotus*. The robinioid clade is here expanded from that described by Lavin et al. (2003) to encompass *Sesbania* and Loteae. Multiple lines of molecular evidence strongly support the monophyly of this collective group. For example, surveys for the presence of the inverted repeat in plastid DNA genomes among papilionoid legumes first suggested Loteae was distinct from the other temperate herbaceous tribes (Lavin et al., 1990) and its sister group relationship to *Sesbania* was only more recently determined (Wojciechowski et al., 2000).

The IRLC—The inverted-repeat-lacking clade or IRLC (Wojciechowski et al., 1999, 2000) includes most members of Polhill's (1981a) temperate herbaceous group. This group comprises all members of tribes Carmichaelieae, Cicereae, Hedysareae, Trifolieae, Vicieae, and Galegeae, but not Loteae (and Coronilleae), as well as at least three genera of Millettieae, *Afgekia* Craib, *Callerya* Endlicher, and *Wisteria* Nutt. The IRLC was essentially the first clade of legumes to be distinguished on the basis of a molecular synapomorphy, loss of one copy of the 25-kilobase inverted repeat in the plastid genome (Lavin et al., 1990; Liston, 1995), in addition to a number of morphological features shared by members of this group including a predominantly herbaceous habit, epulvinate compound leaves, and base chromosome numbers of $n = 7$ or $n = 8$ (Polhill, 1981a). The monophyly of the IRLC has been consistently detected in all subsequent cladistic analyses of molecular sequence data (Sanderson and Wojciechowski, 1996; Doyle et al., 1997; Käss and Wink, 1997; Hu et al., 2000, 2002; Wojciechowski et al., 2000; Kajita et al., 2001). Within the IRLC, the genera *Afgekia* (not included in this analysis), *Callerya*, and *Wisteria*, formerly classified in the tribe Millettieae, along with *Glycyrrhiza* L. of Galegeae, form a paraphyletic grade with respect to the remaining IRLC (Fig. 5). The ''vicioid'' subclade of the IRLC includes many of the particularly important agricultural genera such as *Cicer* L., *Lathyrus* L., *Lens* Mill., *Medicago* L., *Melilotus* Mill., *Pisum* L., *Trifolium* L., and *Vicia* L. (Wojciechowski et al., 2000; Steele and Wojciechowski, 2003). The vicioids are morphologically the most distinctive subclade of the IRLC, apomorphically characterized by craspedidromous leaflets and consistently well supported as monophyletic with *Parochetus* (Trifolieae) as the sister to all other vicioid taxa (Fig. 5). Other major clades within the IRLC include the Astragalean clade (Wojciechowski et al., 1999, 2000), defined as all descendants of the MRCA of *Astragalus americanus* and *Clianthus puniceus* (Fig. 5), and the hedysarioid clade (Sanderson and Wojciechowski, 1996; Wojciechowski et al., 2000), defined as all descendants of the MRCA of *Hedysarum boreale* and *Caragana arborescens* (Fig. 5).

Hologalegina—The robinioids and the IRLC comprise the largest of the well-marked papilionoid subclades, Hologalegina (Wojciechowski et al., 2000). This clade includes over 4800 species that make up the vast majority of legumes presently distributed in temperate regions of the world. The robinioids-IRLC dichotomy had been detected in studies with *rbcL* (Doyle et al., 1997; Kajita et al., 2001) but resolution was weakly supported. Furthermore, the placement of *Bolusanthus* (Sophoreae) as the sister group to Robinieae in those studies stands in marked disagreement with the results presented here

and in most other studies that consistently place *Bolusanthus* within the genistoids (Lavin et al., 1998; Hu et al., 2000, 2002; Pennington et al., 2001). Hologalegina and its two subclades are very strongly supported in our *matK* analyses (Fig. 5). Relationships within each of the subclades have received considerable support from analysis of *matK* sequences, and the circumscription of several of the subclades has been the subject of recent studies (Wojciechowski et al., 2000; Allan et al., 2003; Lavin et al., 2003).

Minor papilionoid clades—One of the surprising and well-supported papilionoid subclades to be identified in the *trnL* analyses of Ireland et al. (2000) and Pennington et al. (2001), the vataireoid clade, also finds strong support in our analyses (i.e., the clade with *Sweetia*, *Luetzelburgia*, and *Vatairea*; Fig. 2). The close relationship of *Exostyles*, *Harleyodendron*, *Luetzelburgia*, *Sweetia*, *Vatairea*, and *Vataireopsis* suggested by Pennington et al. (2001) had not been recognized in formal classifications but was suspected on the basis of overall morphology and wood anatomical similarities (e.g., de Lima, 1990; Gasson, 1994). The inclusion of *Exostyles* and *Harleyodendron* in the vataireoid clade (Pennington et al., 2001) was particularly unexpected given the placement of these genera in the tribe Swartzieae because of their very nonpapilionoid floral morphology. As Pennington et al. (2001) point out, members of the vataireoid clade lack the ability to nodulate with symbiotic rhizobia, thus revealing the existence of more than one well-defined subgroup of woody papilionoid legumes that are not known to form root nodules (Sprent, 2001).

Other well-supported clades within papilionoids with unresolved relationships to the major subclades include one comprising *Holocalyx* and *Uribea* and another represented by *Sophora* sect. *Calia* (Fig. 2). *Holocalyx* and *Uribea*, both monotypic genera of trees from Central America and tropical South America, have been treated in Sophoreae although *Holocalyx* has been recently transferred to Swartzieae (Polhill, 1994). The placement of *Calia arizonica* as sister to *C. secundiflora* and distinctly separate from both *Styphnolobium* (of the *Cladrastis* clade) and *Sophora* L. s.s. (of the genistoid s.l. clade) lends further credence to Sousa and Rudd's (1993) distinction of both the North American *Calia* Berland. and predominantly Mesoamerican *Styphnolobium* Schott as separate from *Sophora* s.s.

Our analyses also provide support for the baphioid clade of Pennington et al. (2001) as the sister group to the mirbelioids plus Hypocalypteae (Fig. 4). Represented here by two species of *Baphia*, the baphioid clade corresponds closely with the *Baphia* group of Polhill (1981b), a small group (c. 60 spp.) of largely African and southern Asian Sophoreae that includes *Baphia*, *Airyantha*, *Baphiastrum*, *Bowringia*, *Dalhousiea*, and *Leucomphalos*, plus tropical African *Baphiopsis* of Swartzieae (Polhill, 1994).

The 50-kb inversion clade—A 50-kb inversion within the large, single-copy region of the plastid genome is a structural rearrangement that Doyle et al. (1996) suggested is a synapomorphy for a clade that includes most papilionoids (Fig. 2). Other molecular phylogenetic studies have noted that certain genera of Sophoreae, Swartzieae, and Dipterygeae known to lack the 50-kb inversion have unresolved relationships within the papilionoid clade (e.g., Doyle et al., 1997; Pennington et al., 2001). Consistent with the *trnL* sequence analysis of Pennington et al. (2001), *matK* sequences also resolve a clade consistent with the distribution of this inversion, albeit with weak support (Fig. 2). The precise placement of the MRCA of this inversion clade in our phylogeny is rendered uncertain because it is not yet known whether *Calia* has this rearrangement in its plastid genome. However, every taxon known to possess the 50-kb inversion, based on the taxa sampled by Doyle et al. (1996) and R. T. Pennington (Royal Botanic Garden, Edinburgh, Scotland, unpublished data), is nested in the clade that is sister to the *Cladrastis* clade (Fig. 2). Furthermore, the greater nodal support and presence of quinolizidine alkaloids in *Calia* (Kite and Pennington, 2003) argue for the more inclusive clade, i.e., MRCA of *Calia secundiflora* and *Vicia faba*, being marked by the presence of this inversion. To characterize this node more precisely, greater sampling of *matK* sequences of genera putatively branching from near the presently detected 50-kb inversion node (Fig. 2) is required, as are experiments designed to detect the presence or absence of the 50-kb inversion in these taxa (cf. Doyle et al., 1996).

The canavanine-accumulating clade—The *matK* phylogeny supports a large clade, distinguished by the ability to produce the nonprotein amino acid canavanine and defined by the MRCA of the mirbelioid clade and the IRLC. This clade is poorly supported by bootstrap proportions, but moderately supported by Bayesian posterior probabilities and its presence in the parsimony strict consensus (Fig. 4). A hypothesis for the single origin of canavanine biosynthesis was put forth by Bell (1981), who showed that canavanine, a close analog of arginine, was essentially mutually exclusive of alkaloid accumulation and restricted to 16 closely related papilionoid tribes (notwithstanding a dubious report from *Laburnum anagyroides*, Genisteae; Wink and Mohamed, 2003), all of which group here in what we term the ''canavanine-accumulating clade'' (Fig. 4). However, at that time, the close relationship of the tribes Bossiaeeae, Mirbelieae, and Hypocalypteae (formerly synonymized under Liparieae; Polhill, 1981c) to other tribes accumulating canavanine was suspect. This was because the original circumscriptions of these tribes suggested they were polymorphic for degrees of floral fusion and canavanine or alkaloid accumulation and thus were possibly intermediate between the canavanine- and alkaloid-accumulating tribes (see Fig. 4 or Polhill, 1981a, where this primary division of papilionoid tribes is centered on ''Tephrosieae''). Notably, the *matK* results presented in this analysis place the mirbelioid clade as the sister group to the rest of the canavanine-accumulating clade. Once the taxonomy of the constituents of the current mirbelioid clade and its sister group was resolved (i.e., Hypocalypteae segregated from Liparieae, Bossiaeeae stripped of the *Templetonia* group; see Crisp et al., 2000), it has become clear that the pattern of canavanine accumulation in papilionoid legumes is mutually exclusive of alkaloid accumulation and thus appears to have evolved only once along the branch subtending the MRCA of the mirbelioid clade and IRLC.

Pennington et al. (2001) and Kajita et al. (2001) did not resolve a group corresponding to this canavanine-accumulating clade in the detail shown in the present analysis. Most importantly, the mirbelioid clade was not resolved as a lineage within this canavanine-accumulating clade while the ''baphioid clade'' of Pennington et al. (2001), apomorphically diagnosed by unifoliolate leaves, was not resolved as the sole sister group to the canavanine-accumulating clade, as they are in this *matK* analysis (Fig. 4). Pennington et al. (2001) do

show a strongly supported baphioid clade nested within what would otherwise be the canavanine-accumulating clade, while resolution with *rbcL* was so poor that the canavanine accumulators are represented by at least three clades nested among several alkaloid-accumulating lineages (e.g., genistoids in part, baphioids, the genus *Amorpha*; fig. 3 of Kajita et al., 2001). In sum, the detection in this analysis of the sister relationship of the baphioid clade to that of the canavanine-accumulating clade is additional evidence for how informative the *matK* locus is for phylogenetic analysis at these hierarchical levels.

Correspondence of the matK phylogeny with Polhill's tribal classification and its implications for legume evolution—The taxonomic implications of the results of this and other recent, especially molecular, phylogenetic studies of legumes are significant. Clearly, if a higher-level classification of the entire family is to reflect phylogenetic relationships (Fig. 6), important changes will be necessary. For example, of the papilionoid tribes with two genera or more circumscribed by Polhill (1994) only Amorpheae and Dipterygeae, and perhaps Bossiaeeae and Psoraleeae, are monophyletic. Sophoreae and Swartzieae are clearly polyphyletic, based on molecular and morphological data (Herendeen, 1995; Doyle et al., 1997; Pennington et al., 2001), and dispersed throughout several papilionoid subclades, both outside and inside the 50-kb inversion clade (e.g., the baphioids, *Maackia*, *Bolusanthus*, and *Sophora* s.s. in the genistoids, or the vataireoids). The nonmonophyly of such groups of genera (Figs. 2, 4), especially those classified in Swartzieae, reveals that many morphological features once considered distinctive (e.g., the swartzioid valvate calyx that bursts open at anthesis) are more prone to evolve independently than previously thought. Thus, Swartzieae can no longer be viewed as a ''transitional'' lineage (cf. Polhill, 1994) between Caesalpinioideae and Papilionoideae if its component lineages are thoroughly interdigitated among those of Papilionoideae or portrayed as sufficiently distinct from either Papilionoideae or Caesalpinioideae to warrant separate subfamily status (e.g., Tucker, 2003b). Likewise, the large tropical tribe Millettieae, once considered the paraphyletic stem clade from which many of the temperate and tropical papilionoid groups in the Old World were derived (Geesink, 1984), comprises three lineages that belong to distantly related groups, the genistoids, millettioids, and IRLC (Hu et al., 2002). Lastly, the morphologically diverse and cosmopolitan tribe Galegeae is polyphyletic within the IRLC, which comprises the bulk of temperate legume diversity (Wojciechowski et al., 2000). Comparable examples of incongruence between current classification schemes for the other subfamilies and the results from recent molecular and morphological studies have been presented elsewhere (e.g., Caesalpinioideae: Bruneau et al., 2001; Mimosoideae: Luckow et al., 2003).

In contrast to the described examples in which one traditionally recognized tribe is now known to comprise several distinct lineages, multiple traditionally recognized tribes or subtribes are now known to form a single well-marked clade. The pantropical dalbergioid clade includes most genera of Dalbergieae and all of Aeschynomeneae, as well as the subtribe Bryinae of tribe Desmodieae. The genera of Dalbergieae are polyphyletic within this dalbergioid clade (Lavin et al., 2001). This composition of heretofore-eclectic genera is made even more notable when its sister group, the primarily North American Amorpheae, is included (McMahon and Hufford, 2004). Similarly, the primarily north temperate Loteae plus North American Robinieae form the strongly supported sister clade to the largely temperate and herbaceous IRLC. The pantropical *Sesbania*, previously of Robinieae, is weakly but consistently resolved as sister to Loteae. The finding of a close relationship of the South American genus *Anarthrophyllum* to *Lupinus* and related Genisteae and that of *Hypocalytus* to the Australasian Mirbelieae plus Bossiaeeae, rather than to the South African tribes such as Crotalarieae, is consistent with the distribution of quinolizidine alkaloids in these taxa (Van Wyk et al., 1993; van Wyk and Schutte, 1995).

While the new informal classification system proposed here for papilionoid legumes bears some resemblance to Polhill's tribal classification, many genera have relationships that are novel and in some cases not predicted by previous classification schemes. This incongruence may in part be caused by the conventional perception of tribes (and subfamilies) as occupying ''basal'' or ''primitive'' vs. ''derived'' or ''advanced'' positions within the legume phylogenetic tree (e.g., Pennington et al., 2000, 2001; Kajita et al., 2001; Herendeen et al., 2003a; Luckow et al., 2003; Tucker, 2003a; Wojciechowski, 2003, to list a few recent examples where such terminology is still used). Terms like ''basal'' have been used in the legume literature to interpret morphologies, in addition to relationships, from an evolutionary perspective. Traditional legume classifications such as Polhill's (1981a, 1994) invoked certain suites of morphologies that served a priori as a measure of divergence of an extant taxon from the MRCA of legumes.

In contrast, the comprehensive and well-resolved *matK* phylogeny reveals relationships in terms of sister clades and supports none of the previous ideas that invoked terms like ''basal'' and ''derived.'' For example, floral ontogenetic studies once suggested that radial or actinomorphic floral symmetry and suppression of floral organs, observed in taxa such as *Ceratonia* and *Gleditsia*, reflected minimal divergence from the MRCA of all legumes (e.g., Tucker, 1992; Tucker and Douglas, 1994), while the perfect, bilaterally symmetric flower of *Cercis* suggested it was advanced and thus phylogenetically nested within the papilionoid line (Tucker and Douglas, 1994). Like recent studies using different phylogenetic data (Pennington et al., 2000, 2001; Tucker, 2002; Herendeen et al., 2003a), the *matK* phylogeny reveals no tendency of genera with radial floral symmetry to occupy a basal position or form a paraphyletic grade in which is nested a clade(s) of taxa with bilateral floral symmetry. For example, all molecular analyses are unequivocal in supporting *Cercis* in Cercideae as the sister clade to the rest of legumes (Bruneau et al., 2001; Kajita et al., 2001) with *Ceratonia* and *Gleditsia* part of the *Umtiza* clade (Herendeen et al., 2003b) nested inside the caesalpinioid crown clade from which mimosoids were derived (Fig. 1). The implications of results such as these for hypotheses of floral evolution in caesalpinioids or early branching lineages of papilionoids have been pointed out by others (e.g., Pennington et al., 2000; Bruneau et al., 2001) and need not be discussed further here. Similarly, the correlation of floral connation, especially of the staminal filaments and keel petals, and accumulation of nonprotein amino acids in seeds marked the putative advanced tribes (e.g., Polhill, 1981a; Doyle et al., 1997). The *matK* phylogeny suggests that the alternative states, flowers with free petals and filaments and accumulating alkaloids in seeds, are not a priori an indicator of little divergence from the MRCA of legumes. The genistoids produce most of the distinctive legume alkaloids, and genera with free stamens can be nested within a fused stamen clade (e.g., *Adesmia*). It is

indeed likely that terms such ''basal,'' ''primitive,'' ''derived,'' and ''advanced'' carry some, if not many, unwarranted implications from previous classification schemes and will have to be abandoned in order to approach legume phylogeny and taxonomy from an open, more contemporary perspective.

LITERATURE CITED

ALFARO, M. E., S. ZOLLER, AND F. LUTZONI. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov Chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 20: 255–266.

ALLAN, G. J., E. A. ZIMMER, W. L. WAGNER, AND D. D. SOKOLOFF. 2003. Molecular phylogenetic analyses of tribe Loteae (Leguminosae): implications for classification and biogeography. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, Higher level systematics, 371–393. Royal Botanic Garden, Kew, UK.

ANGIOSPERM PHYLOGENY GROUP. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

BELL, E. A. 1981. Non-protein amino acids in the Leguminosae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 2, 489–499. Royal Botanic Gardens, Kew, UK.

BENTHAM, G. 1837. Commentationes de Leguminosarum Generibus. Sollinger, Vienna.

BOYER, S. K., AND J. E. MULLET. 1988. Pea chloroplast $tRNA^{Lys}$ (UUU) gene: transcription and analysis of an intron-containing gene. *Photosynthesis Research* 17: 7–22.

BRUNEAU, A., F. FOREST, P. S. HERENDEEN, B. B. KLITGAARD, AND G. P. LEWIS. 2001. Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Systematic Botany* 26: 487–514.

CHAPPILL, J. A. 1995. Cladistic analysis of the Leguminosae: the development of an explicit phylogenetic hypothesis. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, Phylogeny, 1–9. Royal Botanic Gardens, Kew, UK.

CHASE, M. W., ET AL. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanic Garden* 80: 528–580.

CREPET, W. L., AND P. S. HERENDEEN. 1992. Papilionoid flowers from the early Eocene of southeastern North America. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, The fossil record, 43–55. Royal Botanic Gardens, Kew, UK.

CREPET, W. L., AND D. W. TAYLOR. 1985. The diversification of the Leguminosae: first fossil evidence of the Mimosoideae and Papilionoideae. *Science* 228: 1087–1089.

CREPET, W. L., AND D. W. TAYLOR. 1986. Primitive mimosoid flowers from the Paleocene-Eocene and the systematic and evolutionary implications. *American Journal of Botany* 73: 548–563.

CRISP, M. D., AND L. G. COOK. 2003. Phylogeny and embryo sac evolution in the endemic Australasian Papilionoid tribes Mirbelieae and Bossiaeeae. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, Higher level systematics, 253–268. Royal Botanic Garden, Kew, UK.

CRISP, M. D., S. GILMORE, AND B.-E. VAN WYK. 2000. Molecular phylogeny of the genistoid tribes of papilionoid legumes. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 249–276. Royal Botanic Garden, Kew, UK.

CRISP, M. D., AND P. H. WESTON. 1987. Cladistics and legume systematics, with an analysis of the Bossiaeeae, Brongniartieae and Mirbelieae. *In* C. H. Stirton [ed.], Advances in legume systematics, part 3, 65–130. Royal Botanic Gardens, Kew, UK.

DE LIMA, H. C. 1990. Tribo Dalbergieae (Leguminosae Papilionoideae)—morfologia do frutos, sementes e plântulase sua aplicaçao na sistemática. *Arquivos do Jardim Botânico de Rio de Janeiro* 30: 1–42.

DE QUEIROZ, K., AND J. GAUTHIER. 1994. Toward a phylogenetic system of biological nomenclature. *Trends in Ecology and Evolution* 9: 27–31.

DOYLE, J. J. 1995. DNA data and legume phylogeny: a progress report. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, Phylogeny, 11–30. Royal Botanic Gardens, Kew, UK.

DOYLE, J. J., J. A. CHAPPILL, C. D. BAILEY, AND T. KAJITA. 2000. Towards a comprehensive phylogeny of legumes: evidence from *rbcL* sequences and non-molecular data. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 1–20. Royal Botanic Gardens, Kew, UK.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation method for small quantities of fresh tissues. *Phytochemical Bulletin* 19: 11–15.

DOYLE, J. J., J. L. DOYLE, J. A. BALLENGER, E. E. DICKSON, T. KAJITA, AND H. OHASHI. 1997. A phylogeny of the chloroplast gene *rbcL* in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541–554.

DOYLE, J. J., J. L. DOYLE, J. A. BALLENGER, AND J. D. PALMER. 1996. The distribution and phylogenetic significance of a 50-kb chloroplast DNA inversion in the flowering plant family Leguminosae. *Molecular Phylogenetics and Evolution* 5: 429–438.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

GASSON, P. 1994. Wood anatomy of the Sophoreae and related Caesalpinioideae and Papilionoideae. *In* I. K. Ferguson and S. C. Tucker [eds.], Advances in legume systematics, part 6, Structural botany, 165–203. Royal Botanic Gardens, Kew, UK.

GASSON, P. 2000. Does wood anatomy support tribal and generic classification in papilionoid Leguminosae? *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 201–215. Royal Botanic Gardens, Kew.

GEESINK, R. 1981. Tephrosieae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 245–260. Royal Botanic Gardens, Kew, UK.

GEESINK, R. 1984. Scala Millettiearum. E. J. Brill/Leiden University Press, Leiden, The Netherlands.

GOLDBLATT, P. 1981. Cytology and the phylogeny of Leguminosae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 2, 427–463. Royal Botanic Garden, Kew, UK.

HERENDEEN, P. S. 1995. Phylogenetic relationships of the tribe Swartzieae. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, phylogeny, 123–131. Royal Botanic Gardens, Kew, UK.

HERENDEEN, P. S., A. BRUNEAU, AND G. P. LEWIS. 2003a. Phylogenetic relationships in caesalpinioid legumes: a preliminary analysis based on morphological and molecular data. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, Higher level systematics, 37–62. Royal Botanic Garden, Kew, UK.

HERENDEEN, P. S., W. L. CREPET, AND D. L. DILCHER. 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. *In* P. S. Herendeen and D. L. Dilcher [eds.], Advances in legume systematics, part 4, The fossil record, 303–316. Royal Botanic Gardens, Kew, UK.

HERENDEEN, P. S., G. P. LEWIS, AND A. BRUNEAU. 2003b. Floral morphology in caesalpinioid legumes: testing the monophyly of the ''*Umtiza* clade.'' *International Journal of Plant Sciences* 164: S393–S407.

HU, J.-M. 2000. Phylogenetic relationships of the tribe Millettieae and allies—the current status. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 299–310. Royal Botanic Garden, Kew, UK.

HU, J.-M., M. LAVIN, M. F. WOJCIECHOWSKI, AND M. J. SANDERSON. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on *matK* sequences, and implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418–430.

HU, J.-M., M. LAVIN, M. F. WOJCIECHOWSKI, AND M. J. SANDERSON. 2002. Phylogenetic analysis of nuclear ribosomal ITS/5.8 S sequences in the tribe Millettieae (Fabaceae): *Poecilanthe-Cyclolobium*, the core Millettieae, and the *Callerya* group. *Systematic Botany* 27: 722–733.

HUELSENBECK, J. P., B. LARGET, R. E. MILLER, AND F. RONQUIST. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology* 51: 673–688.

HUELSENBECK, J. P., AND F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.

HUGHES, C. E., G. P. LEWIS, A. DAZA YAMONA, AND C. REYNEL. 2004. *Maraniona*, a new dalbergioid legume genus (Leguminosae, Papilionoideae) from Peru. *Systematic Botany* 29: 366–374.

HUTCHINSON, J. 1964. The genera of flowering plants (Angiospermae), vol. I. Clarendon Press, Oxford, UK.

IRELAND, H., R. T. PENNINGTON, AND J. PRESTON. 2000. Molecular systematics of the Swartzieae. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 217–231. Royal Botanic Gardens, Kew, UK.

JOHNSON, L. A., AND D. E. SOLTIS. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Annals of the Missouri Botanical Garden* 82: 149–175.

KAJITA, T., H. OHASHI, Y. TATEISHI, C. D. BAILEY, AND J. J. DOYLE. 2001. *rbcL* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and Allies. *Systematic Botany* 26: 515–536.

KÄSS, E., AND M. WINK. 1995. Molecular phylogeny of the Papilionoideae (Family Leguminosae): *rbcL* gene sequences versus chemical taxonomy. *Botanica Acta* 108: 149–162.

KÄSS, E., AND M. WINK. 1996. Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcL* sequences. *Biochemical Systematics and Ecology* 24: 365–378.

KÄSS, E., AND M. WINK. 1997. Phylogenetic relationships in the Papilionoideae (Family Leguminosae) based on nucleotide sequences of cpDNA (*rbcL*) and ncDNA (ITS1 and 2). *Molecular Phylogenetics and Evolution* 8: 65–88.

KINGHORN, A. D., AND M. F. BALANDRIN. 1984. Quinolizidine alkaloids in the Leguminosae: structural types, analysis, chemotaxonomy and biological activities. *In* S. W. Pelletier [ed.], Alkaloids: chemical and biological perspectives, 105–148. John Wiley and Sons, New York, New York, USA.

KITE, G. C., AND R. T. PENNINGTON. 2003. Quinolizidine alkaloid status of *Styphnolobium* and *Cladrastis* (Leguminosae). *Biochemical Systematics and Ecology* 31: 1409–1416.

LACKEY, J. A. 1981. Phaseoleae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 301–327. Royal Botanic Gardens, Kew, UK.

LAVIN, M., J. J. DOYLE, AND J. D. PALMER. 1990. Evolutionary significance of the loss of the chloroplast—DNA inverted repeat in the Leguminosae subfamily Papilionoideae. *Evolution* 44: 390–402.

LAVIN, M., E. ESHBAUGH, J.-M. HU, S. MATHEWS, AND R. A. SHARROCK. 1998. Monophyletic subgroups of the tribe Millettieae (Leguminosae) as revealed by phytochrome nucleotide sequence data. *American Journal of Botany* 85: 412–433.

LAVIN, M., P. S. HERENDEEN, AND M. F. WOJCIECHOWSKI. In press. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of the major family lineages immediately following an Early Tertiary emergence. *Systematic Biology*.

LAVIN, M., R. T. PENNINGTON, B. B. KLITGAARD, J. I. SPRENT, H. C. DE LIMA, AND P. E. GASSON. 2001. The dalbergioid legumes (Fabaceae): delimitation of a monophyletic pantropical clade. *American Journal of Botany* 88: 503–533.

LAVIN, M., AND M. SOUSA S. 1995. Phylogenetic systematics and biogeography of the tribe Robinieae. *Systematic Botany Monographs* 45: 1–165.

LAVIN, M., M. THULIN, J.-N. LABAT, AND R. T. PENNINGTON. 2000. Africa, the odd man out: molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. *Systematic Botany* 25: 449–467.

LAVIN, M., M. F. WOJCIECHOWSKI, P. GASSON, C. HUGHES, AND E. WHEELER. 2003. Phylogeny of robinioid legumes (Fabaceae) revisited: *Coursetia* and *Gliricidia* recircumscribed, and a biogeographical appraisal of the Caribbean endemics. *Systematic Botany* 28: 387–409.

LEE, J., AND T. HYMOWITZ. 2001. A molecular phylogenetic study of the subtribe Glycininae (Leguminosae) derived from chloroplast DNA *rps16* intron sequences. *American Journal of Botany* 88: 2064–2073.

LEWIS, G. P., B. D. SCHRIRE, B. A. MACKINDER, AND M. LOCK [EDS.]. In press. Legumes of the world. Royal Botanic Gardens, Kew, UK.

LISTON, A. 1995. Use of the polymerase chain reaction to survey for the loss of the inverted repeat in the legume chloroplast genome. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, Phylogeny, 31–40. Royal Botanic Gardens, Kew, UK.

LUCKOW, M., J. T. MILLER, D. J. MURPHY, AND T. LIVSHULTZ. 2003. A phylogenetic analysis of the Mimosoideae (Leguminosae) based on chloroplast DNA sequence data. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, Higher level systematics, 197–220. Royal Botanic Gardens, Kew, UK.

LUCKOW, M., P. J. WHITE, AND A. BRUNEAU. 2000. Relationships among the basal genera of Mimosoid legumes. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 165–180. Royal Botanic Gardens, Kew, UK.

MABBERLEY, D. J. 1997. The plant book, 2nd ed. Cambridge University Press, Cambridge, UK.

MANOS, P. S., AND K. P. STEELE. 1997. Phylogenetic analyses of ''higher'' Hamamelididae based on plastid sequence data. *American Journal of Botany* 84: 1407–1419.

MCKEY, D. 1994. Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. *In* J. I. Sprent and D. McKey [eds.], Advances in legume systematics 5, the nitrogen factor, 211–228. Royal Botanic Gardens, Kew.

MCMAHON, M., AND L. HUFFORD. 2004. Phylogeny of Amorpheae (Fabaceae: Papilionoideae). *American Journal of Botany*. 91: 1219–1230.

MILLER, J. T., J. W. GRIMES, D. J. MURPHY, R. J. BAYER, AND P. Y. LADIGES. 2003. A phylogenetic analysis of the Acacieae and Ingeae (Mimosoideae: Fabaceae) based on *trnK, matK, psbA-trnH,* and *trnL/trnF* sequence data. *Systematic Botany* 28: 558–566.

PALOMINO, G., P. MARTINEZ, C. BERNAL, AND M. SOUSA S. 1993. Differencias cromosomicas entre algunas especies de los generos *Sophora* L. y *Styphnolobium* Schott. *Annals of the Missouri Botanic Garden* 80: 284–290.

PENNINGTON, R. T., B. B. KLITGAARD, H. IRELAND, AND M. LAVIN. 2000. New insights into floral evolution of basal Papilionoideae from molecular phylogenies. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 233–248. Royal Botanic Gardens, Kew, UK.

PENNINGTON, R. T., M. LAVIN, H. IRELAND, B. KLITGAARD, J. PRESTON, AND J.-M. HU. 2001. Phylogenetic relationships of basal papilionoid legumes based upon sequences of the chloroplast *trnL* intron. *Systematic Botany* 26: 537–556.

PERSSON, C. 2001. Phylogenetic relationships in Polygalaceae based on plastid DNA sequences from the *trnL-F* region. *Taxon* 50: 763–779.

POLHILL, R. M. 1981a. Papilionoideae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 191–208. Royal Botanic Gardens, Kew, UK.

POLHILL, R. M. 1981b. Sophoreae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 213–230. Royal Botanic Gardens, Kew, UK.

POLHILL, R. M. 1981c. Liparieae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 398. Royal Botanic Gardens, Kew, UK.

POLHILL, R. M. 1994. Classification of the Leguminosae. *In* F. A. Bisby, J. Buckingham, and J. B. Harborne [eds.], Phytochemical dictionary of the Leguminosae, xxxv–xlviii. Chapman and Hall, New York, New York, USA.

POLHILL, R. M., P. H. RAVEN, AND C. H. STIRTON. 1981. Evolution and systematics of the Leguminosae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 1–26. Royal Botanic Gardens, Kew, UK.

POSADA, D., AND K. A. CRANDALL. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.

RAVEN, P. H., AND D. I. AXELROD. 1995. Origin and relationships of the California Flora. California Native Plant Society, Sacramento, California, USA.

RILEY-HULTING, E., A. DELGADO-SALINAS, AND M. LAVIN. 2004. Phylogenetic systematics of *Strophostyles* (Fabaceae): a North American temperate genus within a neotropical diversification. *Systematic Botany*. 29: 627–653.

RUNDEL, R. W. 1989. Ecological success in relation to plant form and function in the woody legumes. *In* C. H. Stirton and J. L. Zarucchi [eds.], Advances in legume biology, *Monographs in Systematic Botany from the Missouri Botanical Garden* 29: 377–398.

SANDERSON, M. J., AND M. F. WOJCIECHOWSKI. 1996. Diversification rates in a temperate legume clade: are there ''so many species'' of *Astragalus* (Fabaceae)? *American Journal of Botany* 83: 1488–1502.

SCHRIRE, B. D., G. P. LEWIS, AND M. LAVIN. In press. Global distribution patterns of the Leguminosae: insights from recent phylogenies. *In* I. Friis and H. Balsley [eds.], Plant diversity and complexity patterns—Local, regional and global dimensions. *Biologiske Skrifter*.

SCHUTTE, A. L., AND B.-E. VAN WYK. 1998. The tribal position of *Hypocalyptus* Thunberg (Fabaceae). *Novon* 8: 178–182.

SOLTIS, D. E., ET AL. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atp*B sequences. *Botanical Journal of the Linnean Society* 133: 381–461.

SOLTIS, D. E., P. S. SOLTIS, D. R. MORGAN, S. M. SWENSEN, B. C. MULLIN, J. M. DOWD, AND P. G. MARTIN. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences, USA* 92: 2647–2651.

SOUSA, M., AND V. E. RUDD. 1993. Revision del genero *Styphnolobium* (Leguminosae: Papilionoideae: Sophoreae). *Annals of the Missouri Botanical Garden* 80: 270–283.

SPRENT, J. I. 1994. Nitrogen acquisition systems in the Leguminosae. *In* J. I. Sprent and D. McKey [eds.], Advances in legume systematics, part 5, The nitrogen factor, 1–16. Royal Botanic Gardens, Kew, UK.

SPRENT, J. I. 2001. Nodulation in legumes. Royal Botanic Gardens, Kew, UK.

STEELE, K. P., E. TIZON, R. C. EVAN, C. S. CAMPBELL, AND M. F. WOJCIECHOWSKI. 2000. Sister group relationships of Fabaceae and Rosaceae: phylogenetic relationships of Eurosids I. *American Journal of Botany* 87: S160 (abstract).

STEELE, K. P., AND R. VILGALYS. 1994. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *matK*. *Systematic Botany* 19: 126–142.

STEELE, K. P., AND M. F. WOJCIECHOWSKI. 2003. Phylogenetic analyses of tribes Trifolieae and Vicieae, based on sequences of the plastid gene *matK* (Papilionoideae: Leguminosae). *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, higher level systematics, 355–370. Royal Botanic Garden, Kew, UK.

STIRTON, C. H. 1981. Psoraleeae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 2, 337–343. Royal Botanic Garden, Kew, UK.

SUZUKI, Y., G. V. GLAZKO, AND M. NEI. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences, USA* 99: 16 138–16 143.

SWOFFORD, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Massachusetts, USA.

TAKHTAJAN, A. 1969. Flowering plants. Origin and dispersal. Smithsonian Institution Press, Washington, D.C., USA.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876–4882.

THOMPSON, I. R., P. Y. LADIGES, AND J. H. ROSS. 2001. Phylogenetic studies of the tribe Brongniartieae (Fabaceae) using nuclear DNA (ITS-1) and morphological data. *Systematic Botany* 26: 557–570.

THULIN, M., M. LAVIN, R. PASQUET, AND A. DELGADO-SALINAS. In press. Phylogeny and biogeography of *Wajira* (Leguminosae): a monophyletic segregate of *Vigna* centered in the Horn of Africa region. *Systematic Botany*.

TUCKER, S. C. 1987a. Floral initiation and development in legumes. *In* C. H. Stirton [ed.], Advances in legume systematics, part 3, 183–239. Royal Botanic Garden, Kew, UK.

TUCKER, S. C. 1987b. Pseudoracemes in papilionoid legumes: their nature, development and variation. *Botanical Journal of the Linnean Society* 95: 181–206.

TUCKER, S. C. 1992. The developmental basis for sexual expression in *Ceratonia siliqua* (Leguminosae: Cassieae). *American Journal of Botany* 79: 318–327.

TUCKER, S. C. 2002. Floral ontogeny of *Cercis* (Leguminosae: Caesalpinioideae: Cercideae): does it show convergence with papilionoids? *International Journal of Plant Sciences* 163: 75–87.

TUCKER, S. C. 2003a. Floral development in legumes. *Plant Physiology* 131: 911–926.

TUCKER, S. C. 2003b. Floral ontogeny in *Swartzia* (Leguminosae: Papilionoideae: Swartzieae): distribution and role of the ring meristem. *American Journal of Botany* 90: 1271–1292.

TUCKER, S. C., AND A. W. DOUGLAS. 1994. Ontogenetic evidence and phylogenetic relationships among basal taxa of legumes. *In* I. K. Ferguson and S. C. Tucker [eds.], Advances in legume systematics, part 6, Structural botany, 11–32. Royal Botanic Gardens, Kew, UK.

TURNER, B. L. 1981. Thermopsideae. *In* R. M. Polhill and P. H. Raven [eds.], Advances in legume systematics, part 1, 403–407. Royal Botanic Gardens, Kew, UK.

VAN DER BANK, M., M. W. CHASE, B.-E. VAN WYK, M. F. FAY, F. H. VAN DER BANK, G. REEVES, AND A. HULME. 2002. Systematics of the tribe Podalyrieae (Fabaceae) based on DNA, morphological and chemical data. *Botanical Journal of the Linnean Society* 139: 159–170.

VAN WYK, B.-E. 2003. The value of chemosystematics in clarifying relationships in the genistoid tribes of papilionoid legumes. *Biochemical Systematics and Ecology* 31: 875–884.

VAN WYK, B.-E., R. GREINWALD, AND L. WITTE. 1993. The taxonomic significance of alkaloids in the South American genus *Anarthrophyllym*. *Biochemical Systematics and Ecology* 21: 705–709.

VAN WYK, B.-E., AND A. L. SCHUTTE. 1995. Phylogenetic relationships in the tribes Podalyrieae, Liparieae, and Crotalarieae. *In* M. D. Crisp and J. J. Doyle [eds.], Advances in legume systematics, part 7, Phylogeny, 283–308. Royal Botanic Gardens, Kew, UK.

WHITING, M. F., J. C. CARPENTER, Q. D. WHEELER, AND W. C. WHEELER. 1997. The strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* 46: 1–68.

WILCOX, T. P., D. J. ZWICKL, T. A. HEATH, AND D. M. HILLIS. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution* 25: 361–371.

WINK, M., AND G. I. A. MOHAMED. 2003. Evolution of chemical defense traits in the Leguminosae: mapping of distribution patterns of secondary metabolites on a molecular phylogeny inferred from nucleotide sequences of the *rbcL* gene. *Biochemical Systematics and Ecology* 31: 897–917.

WOJCIECHOWSKI, M. F. 2003. Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective. *In* B. B. Klitgaard and A. Bruneau [eds.], Advances in legume systematics, part 10, Higher level systematics, 5–35. Royal Botanic Garden, Kew, UK.

WOJCIECHOWSKI, M. F., M. J. SANDERSON, AND J.-M. HU. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. *Systematic Botany* 24: 409–437.

WOJCIECHOWSKI, M. F., M. J. SANDERSON, K. P. STEELE, AND A. LISTON. 2000. Molecular phylogeny of the ''temperate herbaceous tribes'' of papilionoid legumes: a supertree approach. *In* P. S. Herendeen and A. Bruneau [eds.], Advances in legume systematics, part 9, 277–298. Royal Botanic Garden, Kew, UK.