

## ***rbcl* and Legume Phylogeny, with Particular Reference to Phaseoleae, Millettieae, and Allies**

TADASHI KAJITA

Botanical Gardens, Graduate School of Science, The University of Tokyo, Hakusan 3-7-1, Bunkyo, Tokyo, 112-0001, Japan

HIROYOSHI OHASHI

Biological Institute, Faculty of Science, Tohoku University, Sendai, 980-8578, Japan

YOICHI TATEISHI

College of Education, University of the Ryukyus, Nishihara, Okinawa, 903-01, Japan

C. DONOVAN BAILEY<sup>1</sup> and JEFF J. DOYLE<sup>2</sup>

L. H. Bailey Hortorium, Department of Plant Biology, Cornell University, Ithaca, New York 14853  
(jjd5@cornell.edu)

<sup>1</sup>Present address: Department of Plant Sciences, University of Oxford, Oxford, OX1 3RB, UK

<sup>2</sup>Author for correspondence

*Communicating Editor: Aaron Liston*

**ABSTRACT.** A parsimony analysis was conducted on 319 *rbcl* sequences, comprising 242 from 194 genera of Leguminosae and 77 from other families. Results support earlier conclusions from *rbcl* and other molecular data that a monophyletic Leguminosae is part of a Fabales that includes Polygalaceae, Surianaceae, and the anomalous rosid genus *Quillaja*. Within legumes, results of previous analyses were also supported, such as the paraphyletic nature of Caesalpinoideae and monophyly of Mimosoideae and Papilionoideae. Most new data (74 sequences) were from Papilionoideae, particularly Phaseoleae, Millettieae, and allies. Although the overall topology for Papilionoideae was largely unresolved, several large clades were well-supported. The analysis contained a large sample of Phaseoleae and Millettieae, and not surprisingly showed both tribes to be polyphyletic, though with all taxa except *Wisteria* and allied Millettieae belonging to a single well-supported clade. Within this clade was a strongly supported group that included Phaseoleae subtribes Erythrinae, Glycininae, Phaseolinae, Kennediinae, and Cajaninae, with only the last two being monophyletic. Desmodieae and Psoraleeae were also part of this clade. The monophyletic Phaseoleae subtribes Ophrestinae and Diocleinae grouped with most Millettieae in a clade that included a group similar to the core Millettieae identified in other studies. All but one of the remaining Millettieae sampled formed an additional clade within the overall millettoid/phaseoloid group.

Of the various genes used for plant molecular systematic analyses at higher taxonomic levels, *rbcl* has been by far the most widely used, particularly for comprehensive analyses of angiosperms, whether alone (e.g., Chase et al. 1993; Källersjö et al. 1998) or with other genes (e.g., Qiu et al. 1999; Soltis et al. 1999). Although several limitations of *rbcl* for angiosperm phylogeny reconstruction have been known since the earliest studies (e.g., Chase et al. 1993), the gene continues to be used in part because comparable sampling of a readily alignable sequence does not exist elsewhere. The availability of thousands of *rbcl* sequences in public databases (over 8,000 as of late 2000), representing all major groups of plants, allows the affinities of taxa whose

phylogenetic relationships are unknown to be hypothesized simply and quickly.

The *rbcl* gene has played a role in the evolving understanding of legume phylogeny. The earliest comprehensive cladistic analyses of legume phylogeny with broad sampling were those of Chappill (1995), using a wide array of characters, and *rbcl* studies by two groups (Doyle 1995; Käss and Wink 1995). Both groups subsequently expanded these studies (Käss and Wink 1996, 1997a, 1997b; Doyle et al. 1997). Results from these studies were largely concordant with earlier molecular work, confirming for example the monophyly of groups with structural mutations of the chloroplast genome (e.g., Lavin et al. 1990; Doyle et al. 1996), and with long-

standing views concerning the monophyly (or lack thereof) of the three subfamilies. Major groups in *rbcl* topologies were in many cases unresolved or weakly supported, particularly near the base of the tree in the paraphyletic Caesalpinioideae (Doyle et al. 1997). However, several large clades were identified within Papilionoideae, some of which were previously unknown, and several of which were well-supported.

Phylogenetic analyses of the combined sequences from the two 1997 studies have not been published, and numerous new legume *rbcl* sequences have been generated since then. Moreover, none of the legume *rbcl* phylogenies included many outgroups. The sister group relationships of legumes are controversial, with molecular results in conflict with traditional views. The availability of a large number of *rbcl* sequences both from legumes and from putatively related taxa makes it possible to study the effect of extensive legume sampling on outgroup relationships and of outgroup sampling on topologies within.

With recent improvements in computer hardware and software, as well as in search strategies, it is now possible to perform more thorough parsimony searches of tree space for large data sets (e.g., Nixon 1999). The goal of this paper is to conduct such a parsimony analysis on the large number of available legume *rbcl* sequences and numerous outgroups.

#### MATERIALS AND METHODS

**Taxon Sampling.** The sample of approximately 250 Leguminosae sequences publicly available at the commencement of this project was biased toward some groups, particularly the papilionoid tribe Genisteae, which had been the focus of studies by the Wink laboratory (e.g., Käss and Wink 1997a, b). There was some overlap in genera and in some cases even species sampled between our group (Doyle et al. 1997) and the Wink group (Käss and Wink 1997b). Initial parsimony analyses were conducted in order to develop a data set that minimized redundancy and excessive sampling of genera such as *Lupinus* L. (Käss and Wink 1997a). Relatively few of our many Desmodieae sequences were included here, because relationships in this tribe will be discussed elsewhere. No more than two sequences were retained for any genus whose sequences were monophyletic in such analyses; for genera whose sequences did not form monophyletic groups (e.g., *Sophora* L.), all sequences were used.

For species having multiple representative sequences but which did not belong to multiply-sampled genera, all sequences were used unless they were identical.

The resulting data set of 242 legume sequences represented 194 genera (Table 1) and included 74 new sequences. Emphasis was on Papilionoideae, with sequences from 164 of the 451 genera and all 30 of the tribes recognized by Polhill (1994), whose classification is used throughout this section. For Caesalpinioideae, 24 of 151 genera were included, representing all four tribes (Caesalpinieae, Cassieae, Cercideae, Detarieae). This included five of the nine informal "groups" of Caesalpinieae, four of the five subtribes of Cassieae, and both subtribes of Cercideae. However, only five genera were included from the large (81 genera) Detarieae, representing four of the 10 informal "groups." Outside of Detarieae, sampling deficiencies were due mostly to the difficulty in obtaining usable material. For example, numerous attempts to obtain sequences from collections of *Duparquetia* Baill. (Cassieae: Duparquetiinae) and *Poeppigia* Presl (Caesalpinieae: *Poeppigia* group) were unsuccessful. Sampling was lowest for Mimosoideae, with only six genera represented. However, this subfamily has been assumed to be monophyletic.

Seventy-seven sequences from families other than Leguminosae were also included (Table 1). These were chosen to represent: 1) families shown by previous comprehensive *rbcl* analyses (e.g., Chase et al. 1993; Soltis et al. 1995; Källersjö et al. 1998) to belong to clades near legumes; 2) families hypothesized to be near legumes on the basis of morphology, chemistry, and other non-molecular data (Dickison 1981; Thorne 1992); and 3) families identified as close to legumes by the molecular, non-molecular, or combined analyses of Nandi et al. (1998). *Asarum* (Aristolochiaceae) was included as the outgroup to this assembly of largely "rosid" taxa. One new sequence was added, from *Byrsocarpus coccinea*, as a check on the position of Connaraceae, a key family from which only a single sequence (from *Connarus conchocarpus*) was publicly available.

**Phylogenetic Analysis.** The first 1,434 bases of the *rbcl* gene were aligned in Winclada (Nixon 1999b); the first 30 positions, corresponding to the forward amplification primer, were not used in analyses. Approximately 2% of 530 parsimony-informative sites were missing in the data set, primarily at the extreme 3' or 5' ends of sequences. Among legume taxa, only partial sequences were

available for *Dialium* (335 of 530 informative positions), *Hymenaea protera* (122/530), *Hymenolobium excelsum* (235/530), *Fordia cauliflora* (269/530), and *Strongylodon macrobotrys* (316/530). The data matrix is available at TreeBASE (<http://www.herbaria.harvard.edu/treebase/>) as study accession number S578.

Parsimony analyses were conducted using NONA (Goloboff 1994), with nucleotide characters treated as unordered and equally weighted. Searches were conducted using the "parsimony ratchet" strategy, which has been shown to be very effective with data sets in excess of 500 terminals (Nixon 1999a), sampling tree space more efficiently than conventional methods (e.g., many iterations of random taxon additions optimizing all characters using equal weights). A typical ratchet analysis begins with a conventional starting tree from randomly ordered taxa (a single random addition sequence) and then initiates an iterative analysis consisting of the following steps: 1) perturbation of the matrix by increasing the weights of, or eliminating, a random small subset of characters; 2) branch swapping to obtain one representative shortest tree; 3) resetting weights to original values; 4) branch swapping with equal weights using the perturbed tree as the starting tree. The cycle is repeated by starting with the tree that resulted from the previous iteration and perturbing the data to start step one over again. A large number of iterations are conducted in a single ratchet analysis, with all equally parsimonious trees being retained. The efficiency of this method is attributed to the fact that shortest trees found with perturbed characters are not most parsimonious solutions, but are close enough that they serve as excellent starting trees for unperturbed analyses. The starting tree and weighting scheme also quickly jumps between tree islands. The use of such trees is a major improvement over conventional random addition trees, which are far from parsimonious and require considerable searching to achieve near-optimality (Nixon 1999a).

Ratchets were implemented as described by Nixon (1999a) using Winclada (Nixon 1999b) to run NONA (Goloboff 1994). Following the guidelines presented by Nixon (1999a), the matrix was analyzed by perturbing 10–20% of the informative characters (weighting step). Individual ratchet runs used the following parameters: 200 iterations, 50–90 characters sampled, 10% of nodes constrained holding one tree per iteration, and default "ambigopoly=" (no swapping on ambiguously supported

nodes). Constraining a subset of nodes during the character weighted tree search greatly increases the speed of the ratchet (Nixon 1999a). Nodes were unconstrained during the equally weighted search. Considerations of memory, topologies obtained, and support values for individual clades led to a decision to run sufficient ratchet analyses (435 in the case of the final complete analysis) to accumulate a total of at least 5,000 unique equally parsimonious trees, from which a strict consensus tree was then constructed.

Branch support values for the strict consensus tree were estimated using one hundred strict consensus bootstrap (Davis et al. 1998) replicates in NONA (Goloboff 1994) spawned in Winclada (Nixon 1999b). For each bootstrap tree, ten random addition sequences using TBR (tree bisection and reconnection) and holding ten trees per replication were conducted (100 replications of mult\*10; h/10—no max\*). The bootstrap values were plotted onto the ratchet strict consensus tree in Winclada and indicate the percentage of the bootstrap trees that contained each consensus clade. Jackknife clade support (Farris et al. 1996) was also estimated using WinClada to spawn jackknife replicates in Nona. One hundred replicates were conducted using 10 random addition sequences (mult\*10) holding 10 shortest trees for each replication (hold/10).

## RESULTS

The complete *rbcl* data set included 1,404 aligned bases with 530 potentially informative characters among the 319 sequences analyzed. 5,700 equally most parsimonious trees were accumulated in 435 ratchets; each tree had a length of 5,997 steps (excluding uninformative characters), an ensemble consistency index of 0.16, and an ensemble retention index of 0.67

For most systematic purposes there is no need to identify all equally parsimonious trees (even when it is possible to do so) because if tree space is searched thoroughly and many tree islands are sampled, no changes in the strict consensus topology will occur as more trees are included from individual islands (Farris et al. 1996; Goloboff 1999; Nixon 1999a). The ratchet strategy is designed to identify many more islands than would be found in a comparable time using a conventional strategy (Nixon 1999a); thus the strict consensus from the 5,700 trees obtained during our searches is unlikely to collapse further if more trees had been saved.

We initially tested these contentions in analyses

TABLE 1. Taxa sampled. Voucher information (collection and herbarium abbreviation) and GenBank accession numbers are given for *rbcl* sequences reported here for the first time. For samples for which *rbcl* sequences were reported elsewhere, only the GenBank number is provided here. Non-legumes (outgroups) are listed in alphabetical order, by genus, with family given in parentheses following the accession number. Legumes are listed by subfamily and tribe, following Polhill (1994) with the exception of *Hovea*, classified as Brongniartieae following Crisp and Weston (1987).

**NON-LEGUMES:** *Acer saccharum* Marsh. L01881 (Sapindaceae); *Aesculus pavia* Castigl. U39277 (Sapindaceae); *Ailanthus altissima* Swingle L12566 (Simaroubaceae); *Alnus incana* (L.) Moench X56618 (Betulaceae); *Aporosa frutescens* Blume Z75674 (Euphorbiaceae); *Asarum canadense* L. L14290 (Aristolochiaceae); *Balanops vieillardii* Boill. AF089760 (Balanopaceae); *Bauera rubrioides* N. Andr. L11174 (Saxifragaceae); *Begonia metallica x sanguinea* Maddi L01888 (Begoniaceae); *Brassica oleracea* L. M88342 (Brassicaceae); *Byrsocarpus coccinea* Benth. AF308704 (Connaraceae: Herendeen 9-XII-97-7, US); *Casuarina litorea* Stickm. L01893 (Casuarinaceae); *Ceanothus sanguineus* Pursh U06795 (Rhamnaceae); *Celtis sinensis* var. *japonica* (Planchon) Nakai D86309 (Ulmaceae); *Celtis yunnanensis* C.K. Schneid. L12638 (Ulmaceae); *Chrysobalanus icaco* L. L11178 (Chrysobalanaceae); *Citrus paradisi* Macfad. AJ238407 (Rutaceae); *Clarkia xantiana* A. Gray L01896 (Onagraceae); *Cleome hassleriana* Chodat M95755 (Capparaceae); *Comesperma ericinum* DC. L29492 (Polygalaceae); *Connarus conchocarpus* F. Muell. L29493 (Connaraceae); *Coriaria myrtifolia* L. L01897 (Coriariaceae); *Corynocarpus laevigatus* J.R. Forst. & G. Forst. X69731 (Corynocarpaceae); *Crossosoma californicum* Nutt. L11179 (Crossosomataceae); *Cucumis sativus* L. L21937 (Cucurbitaceae); *Datisca cannabina* L. L21939 (Datisceae); *Drypetes roxburghii* (Wall.) Hurus. M95757 (Euphorbiaceae); *Elaeagnus angustifolia* L. U17038 (Elaeagnaceae); *Elaeocarpus grandis* F. Muell. L28951 (Elaeocarpaceae); *Eucriphia lucida* (Labill.) Baill. L01918 (Cunoniaceae); *Euonymus alatus* (Thunb.) Siebold L13184 (Celastraceae); *Euphorbia polychroma* A. Kern L13185 (Euphorbiaceae); *Fagus grandifolia* Ehrh. L13338 (Fagaceae); *Girouneria subaequalis* Planch. D86311 (Ulmaceae); *Gossypium robinsonii* F. Muell. L13186 (Malvaceae); *Guaiaacum sanctum* L. AJ131770 (Zygophyllaceae); *Guilfoylia monostylis* F. Muell. L29494 (Surianaceae); *Heteropyxis natalensis* Harv. U26326 (Heteropyxidaceae); *Humulus lupulus* L. AF061992 (Cannabaceae); *Hymenanchera alpina* (T. Kirk) W.R.B. Oliv. Z75692 (Violaceae); *Juglans nigra* L. U00437 (Juglandaceae); *Koelreuteria paniculata* Laxm. U39283 (Sapindaceae); *Krameria lanceolata* Torr. Y15032 (Krameriaceae); *Leitneria floridana* Chapm. AF062003 (Simaroubaceae); *Licania tomentosa* (Benth.) Fritsch L11193 (Chrysobalanaceae); *Maclura pomifera* (Raf.) C.K. Schneid. D86318 (Moraceae); *Magnolia tripetala* (L.) L. AJ131927 (Magnoliaceae); *Myrica cerifera* L. L01934 (Myricaceae); *Opilia Roxb.* sp. AJ131773 (Opiliaceae); *Oxalis dillenii* Jacq. L01938 (Oxalidaceae); *Photinia x fraseri* Dress L11200 (Rosaceae); *Pilea pumila* (L.) A. Gray U00438 (Urticaceae); *Platytheca verticillata* Baill. L01944 (Tremandraceae); *Polygala cruciata* L. L01945 (Polygalaceae); *Prunus domestica* L. L01947 (Rosaceae); *Punica granatum* L. L10223 (Punicaceae); *Qualea Aubl.* sp. U02730 (Vochysiaceae); *Quillaja saponaria* Molina QSU06822 (Rosaceae); *Rhamnus cathartica* L. L13189 (Rhamnaceae); *Rhoiptelea chiliantha* Diels & Hand.-Mazz. AF017687 (Rhoipteleaceae); *Rinorea crenata* S.F. Blake AJ237591 (Violaceae); *Santalum album* L. L26077 (Santalaceae); *Saxifraga mertensiana* Bong. U06216 (Saxifragaceae); *Schinus molle* L. U39270 (Anacardiaceae); *Securidaca diversifolia* (L.) S.F. Blake L01955 (Polygalaceae); *Shepherdia canadensis* (L.) Nutt. U17039 (Elaeagnaceae); *Simarouba glauca* DC. U38927 (Simaroubaceae); *Spiraea x vanhouttei* Zabel L11206 (Rosaceae); *Sterculia tragacantha* Lindl. AF022126 (Sterculiaceae); *Stylobasium spathulatum* Desf. U06828 (Surianaceae); *Suriana maritima* L. U07680 (Surianaceae); *Swietenia macrophylla* King U39080 (Meliaceae); *Toxicodendron radicans* (L.) Kuntze U39271 (Anacardiaceae); *Trema micrantha* (L.) Blume U03844 (Ulmaceae); *Viola sororia* Willd. L11674 (Violaceae); *Viscum album* L. L26078 (Viscaceae); *Zygophyllum simplex* L. Y15031 (Zygophyllaceae).

#### LEGUMINOSAE

**CAESALPINIOIDEAE: Caesalpinieae:** *Acrocarpus* Wight & Arn. sp. AF308699 (Manos 1416, DUKE); *Caesalpinia pulcherrima* (L.) Sw. U74190; *Caesalpinia pulcherrima* (L.) Sw. Z70153; *Delonix regia* (Bojer ex Hook.) Raf. Z70156; *Erythrophleum ivorense* A. Chev. U74205; *Gleditsia triacanthos* L. Z70129; *Gymnocladus dioica* (L.) K. Koch U74193; *Parkinsonia aculeata* L. Z70157; *Peltophorum peltatum* U74183; *Peltophorum* (Vogel) Benth. sp. U74184; *Tachigali paniculata* Aubl. U74240. **Cassieae:** *Apuleia leiocarpa* (Vogel) J.F. Macbr. U74249; *Cassia fistula* L. U74195; *Cassia senna* L. Z70155; *Ceratonia siliqua* L. U74203; *Chamaecrista fasciculata* (Michx.) Greene U74187; *Dialium* L. sp. U74259; *Petalostylis labicheoides* R. Br. AF308719 (Clemens s.n., BH); *Senna alata* (L.) Roxb. U74250; *Senna didymobotrya* Fresen. Z70154 (deposited as *Cassia didymobotrya*); *Zenia insignis* Chun AF308722 (Pacific Tropical Garden 82s19, HI). **Cercideae:** *Bauhinia candicans* Benth. Z70161; *Bauhinia purpurea* DC. ex Welp. Z70162; *Cercis canadensis* L. U74188; *Cercis siliquastrum* L. Z70164. **Detarieae:** *Brownea* Jacq. sp. U74186; *Hymenaea protera* G. O. Poinar L08477; *Macrolobium acaciifolium* (Benth.) Benth. U74191; *Peltogyne confertiflora* Benth. AF308718 (Bridgewater 793, RBGE); *Tamarindus indica* L. Z70160.

**MIMOSOIDEAE: Acacieae:** *Acacia farnesiana* (L.) Willd. Z70146. **Ingeae:** *Albizia julibrissin* Durazz. Z70147; *Albizia saman* (Jacq.) F. Muell. Z70149; *Paraserianthes lophantha* (Willd.) I.C. Nielsen Z70148; *Pithecellobium mexicanum* L. Z70150. **Mimoseae:** *Mimosa spegazzinii* Pirota Z70151. **Parkieae:** *Parkia roxburghii* G. Don U74209.

**PAPILIONOIDEAE: Abreae:** *Abrus precatorius* L. U74224. **Adesmieae:** *Adesmia exilis* Clos U74254. **Aeschynomeneae:** *Aeschynomene americana* L. AB045784 (H.Ohashi et al. f.n. 12, TUS); *Aeschynomene indica* L. AF308701 (Carulli 58, CHR); *Arachis hypogaea* L. U74247; *Zornia cantoniensis* Mohlenbr. U74235. **Amorpheae:** *Amorpha fruticosa* L. U74212. **Bossiaeeae:**

TABLE 1. Continued.

*Goodia lotifolia* Hort. U74258. **Brongniartieae:** *Brongniartia pacifica* McVaugh U74253; *Hovea elliptica* (Sm.) DC. Z95537. **Carmichaelieae:** *Carmichaelia* R.Br. sp. AF308705 (Berlin Botanical Garden). **Cicereae:** *Cicer arietinum* L. AF308707 (Doyle 1448, BH). **Crotalarieae:** *Aspalathus cephalotes* Thunb. Z70132; *Crotalaria capensis* Baker Z70133; *Lotononis galpinii* Dummer Z95538; *Rafnia* Thunb. sp. Z70136. **Dalbergieae:** *Andira inermis* L. U74199; *Dalbergia hupeana* Hance U74236; *Hymenolobium excelsum* Ducke AB045806 (Vasquez & Jaramarillo 14583, MEXU); *Machaerium lunatum* (L.f.) Ducke U74248; *Vatairea lundelii* Aubl. AB045826 (Calzada 14521, MEXU). **Desmodieae:** *Brya ebenus* (L.) DC. AB045788 (Lewis 1411, K); *Campylotropis griffithii* Schindl. U74228; *Cranocarpus martii* Benth. AB045796 (S. Mori & B. Boom 14133, K); *Desmodium podocarpum* DC. subsp. *oxyphyllum* DC. Ohashi U74257; *Kummerowia stipulacea* (Maxim.) Makino U74229; *Lespedeza cuneata* (Dum. Cours.) G. Don U74215; *Mucuna macrocarpa* Wall. AB045811 (Cultivated Toyko Bot. Gard., TI); *Mucuna nigricans* Steud. AB045812 (Tateishi 10610, TUS); *Phylacium majus* Collett & Hemsl AB045815 (Tateishi et al. 29–6, TUS). **Dipterygeae:** *Dipteryx odorata* (Aubl.) Willd. U74245. **Euchretea:** *Euchresta horsfieldii* (Lesch.) Benn. U74225. **Galegeae:** *Astragalus sparsus* Decne. Z95550; *Caragana arborescens* Lam. Z70168; *Glycyrrhiza glabra* L. AB045804 (Kajita 96090602, TUS); *Glycyrrhiza pallidiflora* Maxim. AB012129; *Halimodendron halodendron* Druce Z95536. **Genisteae:** *Adenocarpus complicatus* J. Gay ex Grenier & Godron Z70113; *Argyrocytistus battandieri* (Maire) Raynaud Z70092; *Argyrolobium harveyanum* Oliv. Z95546; *Argyrolobium marginatum* Bolus Z95547; *Argyrolobium uniflorum* Harvey Z95548; *Argyrolobium zanonii* (Turra) P.W. Ball Z95549; *Calicotome villosa* (Poir.) Link Z70089; *Chamaecytisus austriacus* (L.) Link Z70079; *Chamaecytisus proliferus* (L.f.) Link Z70080; *Chamaecytisus purpureus* (Scop.) Link Z70081; *Chamaespartium sagittale* (L.) P. Gibbs Z70103; *Chamaespartium tridentatum* (L.) P. Gibbs Z70104; *Cytisophyllum sessilifolium* Lang Z70090; *Cytisus arboreus* DC. Z70083; *Dichilus lebeckioides* DC. U74223; *Erinacea anthyllis* Link Z70105; *Genista cinerea* DC. Z70094; *Genista florida* Asso Z70096; *Genista januensis* Viv. GJANURBCL; *Genista tricuspidata* Desf. Z70100; *Laburnum anagyroides* Medik. Z70077; *Lupinus bogotensis* Benth. Z70060; *Lupinus densiflorus* Benth. Z70062; *Melolobium microphyllum* Eckl. & Zeyh. Z95539; *Petteria ramentacea* (Sieber) C. Presl. Z70091; *Retama monosperma* (L.) Boiss. Z70117; *Spartium junceum* L. Z70102; *Spartocytisus supranubius* Christ Z70088; *Teline monspessulana* (L.) K. Koch Z70107; *Ulex europaeus* L. Z70111. **Hedysareae:** *Hedysarum vicioides* Turcz. U74246. **Indigoferae:** *Indigofera australis* Willd. U74214; *Indigofera australis* Willd. AF308711 (Crisp 9095, CBG); *Phylloxylon perrieri* Drake U74256. **Liparieae:** *Hypocalyptus sophoroides* Druce AF308710 (Crisp SA6, CBG); *Liparia splendens* (Burman f.) Bos & de Wit Z70125. **Loteae:** *Coronilla emerus* L. Z70169; *Coronilla varia* L. U74222; *Lotus corniculatus* L. U74213. **Millettieae:** *Afgekia sericea* Craib. AB045785 (855144, WTG); *Aganope* Miq. sp. AF308702 (Turner 165, F); *Apurimacia boliviana* (Britton) M. Lavin AF308703 (Alayo 025, F); *Austrosteensia blackii* (F. Muell.) R. Geesink. U74242; *Chadsia versicolor* Bojer AB045794 (Shrire et al. 2530, BH); *Craibia brevicaudata* (Vatke) Dunn AB045795 (Polhill & Robertson 5296, K); *Dalbergiella nyasae* E.G. Baker AF308724 (Lavin s.n., K); *Dalbergiella welwitschii* (Bak.) Bak. f. AB045797 (Harris & Fay 1840, MEXU); *Deguelia hatschbachii* A.M.G. de Azevedo-Tozzi AB045798 (Sousa & Tozzi 13379, MEXU); *Derris laxiflora* Benth. U74234; *Dewevrea bilabiata* M.Michel AB045799 (Harris & Fay 1432, MEXU); *Fordia cauliflora* Hemsl. AB045802 (Ding & Shi 1701, TUS); *Hesperothamnus pentaphyllus* Harms. AB045805 (Ventura 15244, MEXU); *Leptoderris fasciculata* (Benth.) Dunn. AB045807 (Jonkind et al. 1424, MEXU); *Lonchocarpus eriocarinalis* Micheli AB045808 (Lavin 5235a, BH); *Lonchocarpus hermannii* M. Sousa AB045809 (Friedman 206–94, MEXU); *Millettia japonica* (Siebold & Zucc.) A. Gray U74233; *Millettia lenneoides* Vatke AF308713 (Schrire et al. 2538, K); *Millettia pulchra* (Benth.) Kurz AB045810 (Tateishi et al. 18372, TUS); *Millettia richardiana* (Baill.) Du Puy & Labat AF308714 (Du Puy et al. M573, K); *Muelleria frutescens* (Aubl.) Standl. AB045813 (Sousa 13460, MEXU); *Mundulea sericea* (Willd.) A. Cheval AB045814 (08035, K); *Piscidia piscipula* Sarg. AB045816 (Lavin & Luckow 5793a, TEX); *Platycyamus regnellii* Benth. AB045817 (Pennington s.n., E); *Poecilantlie itapuanum* Benth. AB045818 (Lewis & Andrade 2018, K); *Pongamiopsis amygdalina* (Baill.) R. Vig. AB045819 (Dupuy et al. M560, BH); *Tephrosia grandiflora* (L'Her. ex Aiton) Pers. Z95542; *Tephrosia heckmanniana* Harms U74211; *Willardia mexicana* (S. Watson) Rose AF308721 (Martinez 1020, F); *Wisteria sinensis* (Sims) Sweet Z95544; *Wisteria* Nutt. sp. U74216; *Xeroderris stuhlmannii* Roberty AF308727, AB045827 (Corloy 2162, K). **Mirbelieae:** *Chorizema cordatum* Lindl. U74218; *Daviesia rhizomata* M.D. Crisp AF308708 (Crisp 9002, CBG); *Isotropis cuneifolia* Benth. ex B.D. Jacks. AF308712 (Crisp 8917, CBG). **Phaseoleae:** *Adenodolichos rupestris* Verdc. AF308700 (Potter 870418–01, BH); *Amphicarpaea bracteata* (L.) Fernald AF181930; *Amphicarpaea edgeworthii* Benth. AB045786 (Y.T. s.n. 1990, TUS); *Apios taiwaniana* Hosokawa AB045787 (Kajita s.n., TUS); *Atylosia lineata* Wight & Arn. AF181931; *Butea minor* Buch.-Ham. ex Wall. AB045789 (Mikage et al. 9554109, TUS); *Cajanus cajan* (L.) Millsp. AB045791 (Tsuda s.n., TUS); *Cajanus cajan* (L.) Millsp. AB045790 (Doyle 1300, BH); *Cajanus cajan* (L.) Millsp. Z95535; *Calopogonium caeruleum* (Benth.) Sauvalle AF308723 (Doyle 1294, BH); *Calopogonium mucunoides* Desv. AB045792 (Kajita et al. 95120124, TUS); *Canavalia ensiformis* (L.) DC. U74238; *Canavalia rosea* (Sw.) DC. AB045793 (Kajita et al. 98111401, TUS); *Centrosema virginianum* (L.) Benth. AF308706 (Doyle 1423, BH); *Clitoria ternatea* L. U74237; *Cologania lemmonii* L. AF181932; *Dioclea virgata* (L. C. Rich.) Amshoff AF308709 (Doyle 1290, BH); *Dipogon lignosus* (L.) Verdc. AB045800 (Doyle 1297, BH); *Dumasia villosa* DC. AF181933; *Erythrina crista-galli* L. Z70170; *Erythrina speciosa* Andrews AB045801 (Cultivated Tokyo Bot. Gard., TI); *Galactia tashiroi* Maxim. AB045803 (Tsuda s.n., TUS); *Glycine tabacina* (Labill.) Benth. U74197; *Hardenbergia violacea* (Schneev.) Stearn U74241; *Kennedia rubicunda* Vent. AB045828 (Doyle 1318, BH); *Neonotonia wightii* (Arnott) J.A.

TABLE 1. Continued.

---

Lackey AF181934; *Neorautanenia mitis* (A. Rich.) Verdc. AF308715 (Belsky 505, BH); *Ophrestia hedysaroides* (Willd.) Verdc. AF308716 (Doyle 1139, BH); *Ophrestia radicata* (A. Rich.) Verdc. AF308726 (Gris 2778, BH); *Oxyrhynchus volubilis* Brandege AF308717 (Doyle 1296, BH); *Pachyrhizus erosus* (L.) Urb. AF181935; *Pseudeminia comosa* (Baker) Verdc. AF181936; *Pseudovigna argentea* (Willd.) Verdc. AF181937; *Psophocarpus scandens* (Endl.) Verdc. AB045821 (T. Kajita 96090301, TUS); *Psophocarpus scandens* (Endl.) Verdc. AB045820 (Doyle 1125, BH); *Pueraria montana* (Lour.) Merr. AB045822 (Tsuda s.n., TUS); *Rhodopis planisiliqua* (L.) Urb. AF308728 (Zanoni et al. 43304, MO); *Rhynchosia clioorum* S. Moore AF308720 (Doyle 1289, BH); *Rhynchosia minima* (L.) DC. AB045823 (Tsuda s.n., TUS); *Shuteria vestita* 1 Benth. AF308725 (Doyle 1144, BH); *Shuteria vestita* 2 Benth. AB045824 (Kajita s.n., TUS); *Spatholobus parviflorus* Kuntze AB045825 (Mikage et al. 9554025, TI); *Strongylodon macrobotrys* A. Gray AF308729 (Bruneau 974, BH); *Teramnus labialis* (L. f.) Spreng. AF181938; *Vigna radiata* (L.) R. Wilczek X89403. **Podalyrieae:** *Calpurnia aurea* Baker U74239; *Cyclopia genistoides* Sieber ex C. Presl Z70124; *Podalyria calyptrata* (Retz.) Willd. U74217; *Podalyria sericea* R. Br. Z70128; *Virgilia divaricata* Adamson Z70131. **Psoraleeae:** *Bituminaria bituminosa* L. U74211; *Otholobium sericeum* (Poir.) C.H. Stirt. U74219. **Robinieae:** *Robinia pseudoacacia* L. U74220; *Sesbania sesban* (L.) Fawc. & Rendle Z95541. **Sophoreae:** *Acosmium dasycarpum* (Vogel) Yakovlev U74255; *Angylocalyx braunii* Harms. U74200; *Baphia massaiensis* Taub. U74196; *Bolusanthus speciosus* (Bolus) Harms U74243; *Cadia purpurea* Forssk. U74192; *Castanospermum australe* A. Cunn. & C. Fraser Z70143; *Castanospermum australe* A. Cunn. & C. Fraser U74202; *Cladrastis sikokiana* Mak. U74232; *Cladrastis sinensis* Hemsl. Z95551; *Dussia tessmannii* Harms U74206; *Luetzelburgia* Harms sp. U74185; *Maackia amurensis* Rupr. & Maxim. Z70137; *Maackia floribunda* (Miq.) Takeda U74227; *Myrospermum sousanum* A. Delgado & M.C. Johnst. U74207; *Myroxylon balsamum* (L.) Harms U74208; *Ormosia emarginata* (Hook. & Arn.) Benth. U74231; *Pericopsis mooniana* (Thwaites) Thwaites U74210; *Salweenia wardii* Baker f. U74251; *Sophora bhutanica* H. Ohashi U74226; *Sophora davidii* Kom. ex Pavlov Z70138; *Sophora flavescens* Aiton Z70139 *Styphnolobium japonicum* (L.) Schott Z70142 (as *Sophora japonica* L.) U74230; *Sophora jaubertii* Spach ex Jaubert & Spach Z70140; *Sophora secundiflora* (Gomez-Ortega) Lagerh. ex DC. Z70141; *Xanthocercis zambesiaca* (Baker) Dumaz-le-Grand U74189. **Swartzieae:** *Aldina latifolia* Spruce ex Benth. U74252; *Ateleia herbert-smithii* Pittier U74201; *Cordyla africana* Lour. U74204; *Holocalyx balansae* Micheli U74244; *Suartzia* Schreb. sp. U74194. **Thermopsidaeae:** *Anagyris foetida* L. Z70122; *Baptisia tinctoria* (L.) R. Br. Z70120; *Piptanthus nepalensis* (Hook.) D. Don Z70123; *Thermopsis fabacea* (Pall.) DC. Z70121. **Trifolieae:** *Medicago sativa* L. X04975; *Ononis spinosa* L. Z70174. **Vicieae:** *Lathyrus tuberosus* L. Z70172; *Pisum sativum* L. X03853.

---

of a preliminary data set of 315 sequences to verify that the method, and our implementation of it, was effective for our data. First, a single ratchet (200 iterations sampling 55 characters, constraining 10% of the nodes, and holding one tree per iteration) was compared to 200 random addition sequences holding 1 tree per replication (mult\*200h/1 in NONA). Both methods have the potential to find most parsimonious trees in each replication and therefore each could collect as many as 200 equally most parsimonious trees. The ratchet run found 32 equally most parsimonious trees ( $L=5,981$ ). The 200 random addition sequences failed to find most parsimonious trees but found three trees one step longer. The random addition sequence trees were then swapped using max\* to 1,000 equal length trees in a further attempt to find most parsimonious trees. The max\* of the three equal length ( $L=5,982$ ) starting trees did find shortest trees ( $L=5,981$ ) and 1,000 were saved. The strict consensus of these trees was then compared to the strict consensus of the 32 trees collected using the ratchet (all comparisons were made after unsupported nodes were collapsed). Topological conflict among the 1,000 trees from the traditional (non-ratchet) analysis resulted

in 40 nodes collapsing in the strict consensus topology. In contrast, the strict consensus of the 32 trees from the single ratchet run had 49 collapsed nodes. The greater number of collapsed nodes confirmed the ability of the ratchet strategy to sample more tree islands with different (conflicting) topologies. Thus, in this implementation a single ratchet was more efficient than the much more time-consuming traditional search strategy, both in finding shortest trees and in identifying tree islands.

Next we conducted a comparison to decide whether it was worth attempting to swap within islands once a large number of ratchets had been completed. Swapping the 32 ratchet trees to 1,000 trees using max\* in NONA collapsed two more nodes in the strict consensus (compared to the 32 trees alone). We compared this approach to a strategy in which more ratchets were run but with less swapping within islands by adding the 32 trees from the first ratchet analysis to 69 shortest trees ( $L=5,981$ ) identified in a second random ratchet analysis. The strict consensus of the resulting 101 trees collapsed six additional nodes compared to the strict consensus from the first 32 trees. Because the time it took to do one ratchet was equivalent to

the time necessary to swap to 1,000 trees, we concluded that it was more efficient for our data set to perform many ratchets and limit searches within individual tree islands.

In the course of preliminary analyses and in the strict consensus tree (Fig. 1), some apparent anomalies were noted. The sequence from *Averrhoa carambola* L. (GenBank L14692) did not join the other available Oxalidaceae sequence (*Oxalis dillenii*; GenBank L01938) but instead was resolved as sister to *Crossosoma californicum* (GenBank L11179), and a sequence from *Morus alba* L. (L01933) was sister to *Prunus domestica* (L01947) within an otherwise monophyletic Rosaceae. These two sequences were not included in the final analysis. Two sequences from *Celtis* were included, one of which (*C. sinensis* var. *japonica* D86309) grouped as expected with Rosales. The other, *C. yunnanensis* L12638, was nested within Fagales (Fig. 1). The apparent switch between *Abrus precatorius* (Z70167) and *Glycyrrhiza glabra* (Z70171) sequences has been noted elsewhere (Doyle et al. 1997); these sequences were not included in the analysis and the taxa were represented by other sequences.

Regarding legumes, the strict consensus tree (Figs. 1–5) was in most respects similar to results from previous analyses of *rbcL* both in topology and in relative support for individual clades. The tree identified some significant sets of relationships both within and outside the Leguminosae, but there were several unresolved areas, and others where support was weak. It should be noted that the strict consensus bootstrap method used here is expected to yield support values that are lower than the “frequency-within-replicates” bootstrap approach used more commonly in phylogenetic analyses (Davis et al. 1998). Support values from the parsimony jackknife were almost always somewhat higher than those of the strict consensus bootstrap.

#### DISCUSSION

Sampling of legumes for *rbcL* has been relatively low in comprehensive angiosperm studies. The Chase et al. (1993) analysis included only one complete papilionoid sequence (*Medicago sativa*), and two incomplete sequences from the other two subfamilies (Caesalpinioideae: *Bauhinia* L.; Mimosoideae: *Albizia julibrissin*). Even the 2,538 sequence parsimony jackknife analysis (Källersjö et al. 1998) included only 76 legume sequences. This is understandable, because the largest published legume *rbcL* study included fewer than 100 sequences

(Doyle et al. 1997). At the same time, this and other legume *rbcL* studies (Käss and Wink 1995, 1996, 1997b) incorporated sequences from relatively few outgroups. For example, although the Doyle et al. (1997) study involved some exploration of outgroups, their final ingroup analysis used only two Polygalaceae sequences. Thus the present study tests both the effect of more thorough ingroup sampling on sister group relationships of legumes, and of outgroup sampling on relationships within the family.

**Familial Relationships of Legumes.** Dickison (1981) reviewed hypotheses of legume relationships within the angiosperms. Traditionally, Leguminosae have been considered rosid in their affinities. Dickison particularly highlighted Connaraceae as a likely sister group, and also mentioned other rosid families such as Chrysobalanaceae and Crossosomataceae. Krameriaceae, a family once considered a legume (see Cronquist 1981), was not considered a close relative. Sapindalean affinities have also been considered (Dickison 1981; Thorne 1992).

It was thus a considerable surprise to legume systematists when the Chase et al. (1993) *rbcL* analysis placed Polygalaceae as sister to legumes. Floral similarities between the two families have been considered superficial parallelisms. The close relationship of Polygalaceae and Leguminosae has persisted in molecular analyses (e.g., Savolainen et al. 2000). A series of studies involving rosid taxa, including studies aimed at resolving relationships within a “nitrogen-fixing” clade (e.g., Soltis et al. 1995) identified a Fabales comprising Leguminosae, Polygalaceae, Surianaceae, and the anomalous rosaceous genus *Quillaja* Molina. The 2,538 taxon *rbcL* study (Källersjö et al. 1998) also identified this group with high jackknife support. This same Fabales was also strongly supported in the Angiosperm Phylogeny Working Group analysis of *rbcL*, chloroplast *atpB*, and the nuclear ribosomal 18S gene subunit for over 400 taxa (APG 1998), a continuing analysis that was updated in Soltis et al. (1999). These comprehensive molecular analyses place Connaraceae as sister to Oxalidaceae in a strongly supported Oxalidales. In both the jackknife and three-gene analyses the Oxalidales and Fabales are part of the same large clade; although they are not particularly close there, they are closer than either is to Sapindales, which is shown to be more closely related to Malvales and Brassicales in the APG (1998) phylogeny.

However, recent studies that include *rbcL* data are





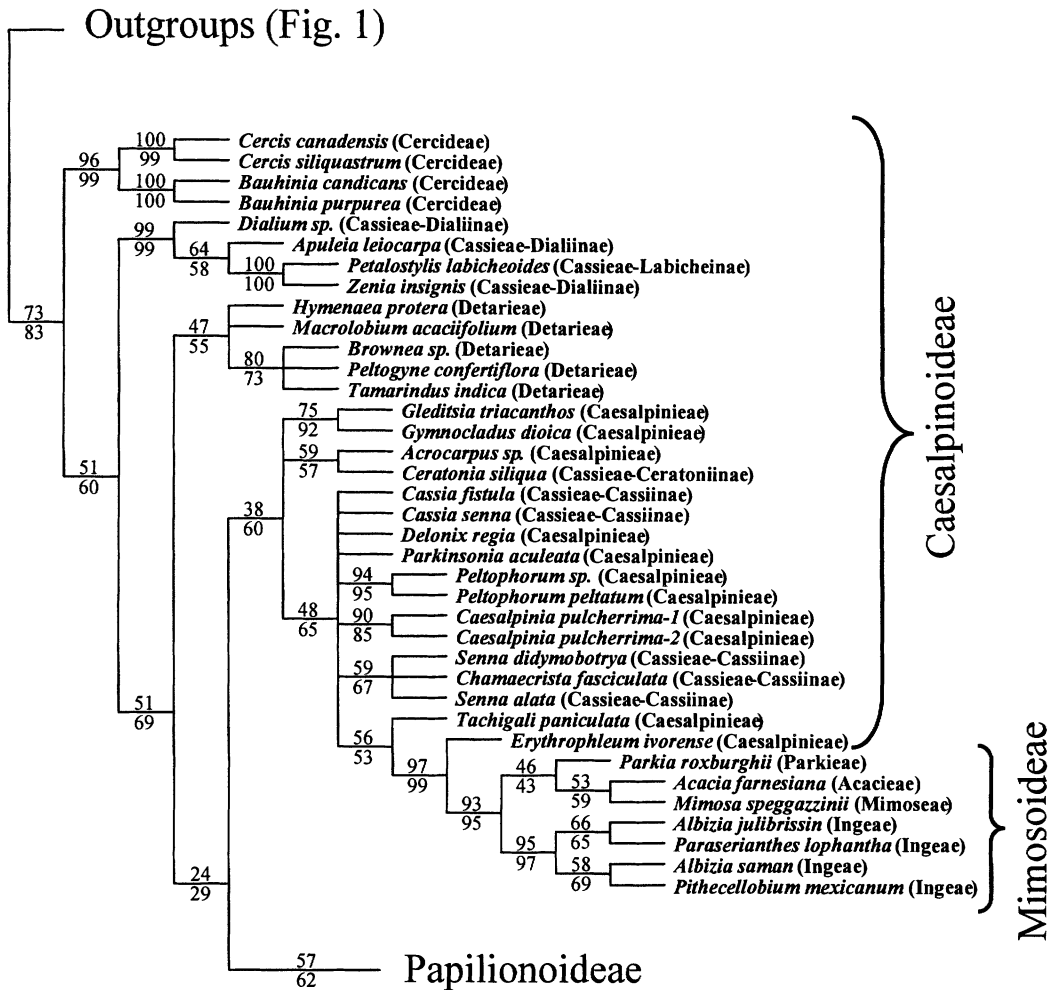


FIG. 2. Strict consensus *rbcL* topology for taxa of Caesalpinoideae and Mimosoideae. The position of Papilionoideae is indicated.

by no means unanimous in supporting these relationships. Nandi et al. (1998) published a study of 162 taxa that included separate analyses of *rbcL* and non-molecular data (morphology, development, chemistry, etc.), as well as a combined analysis. Their analysis of *rbcL* alone did indeed identify a clade that included legumes, Polygalaceae, and Surianaceae, but it also included Faganae (Fagaceae, Betulaceae, etc.) as sister to legumes. Their non-molecular analysis placed Leguminosae as sister to Rutales in a clade that included Sapindales, Euphorbiales, and Urticales; Surianaceae and Polygalaceae were separated from one another and from legumes. Their combined analysis results came to still another conclusion, placing legumes as sister

to Zygophyllaceae, with this clade sister to Krameriaceae. Surianaceae and Polygalaceae were sister taxa, forming part of a "nitrogen-fixing" clade closer to Connaraceae (in Cunoniales) and Chrysobalanaceae (in Malpighiales) than to legumes. However, a note added in proof stated that legumes were scored as lacking nodules in their non-molecular data set, which could explain the failure of legumes and other nitrogen-fixers to group together in some analyses (Nandi et al. 1998).

The *rbcL* strict consensus tree reported here gave results consistent with those of the largest *rbcL* studies (e.g., Källersjö et al. 1998) and the three-gene study (APG 1998) in identifying a strongly supported (87% strict consensus bootstrap, 95%

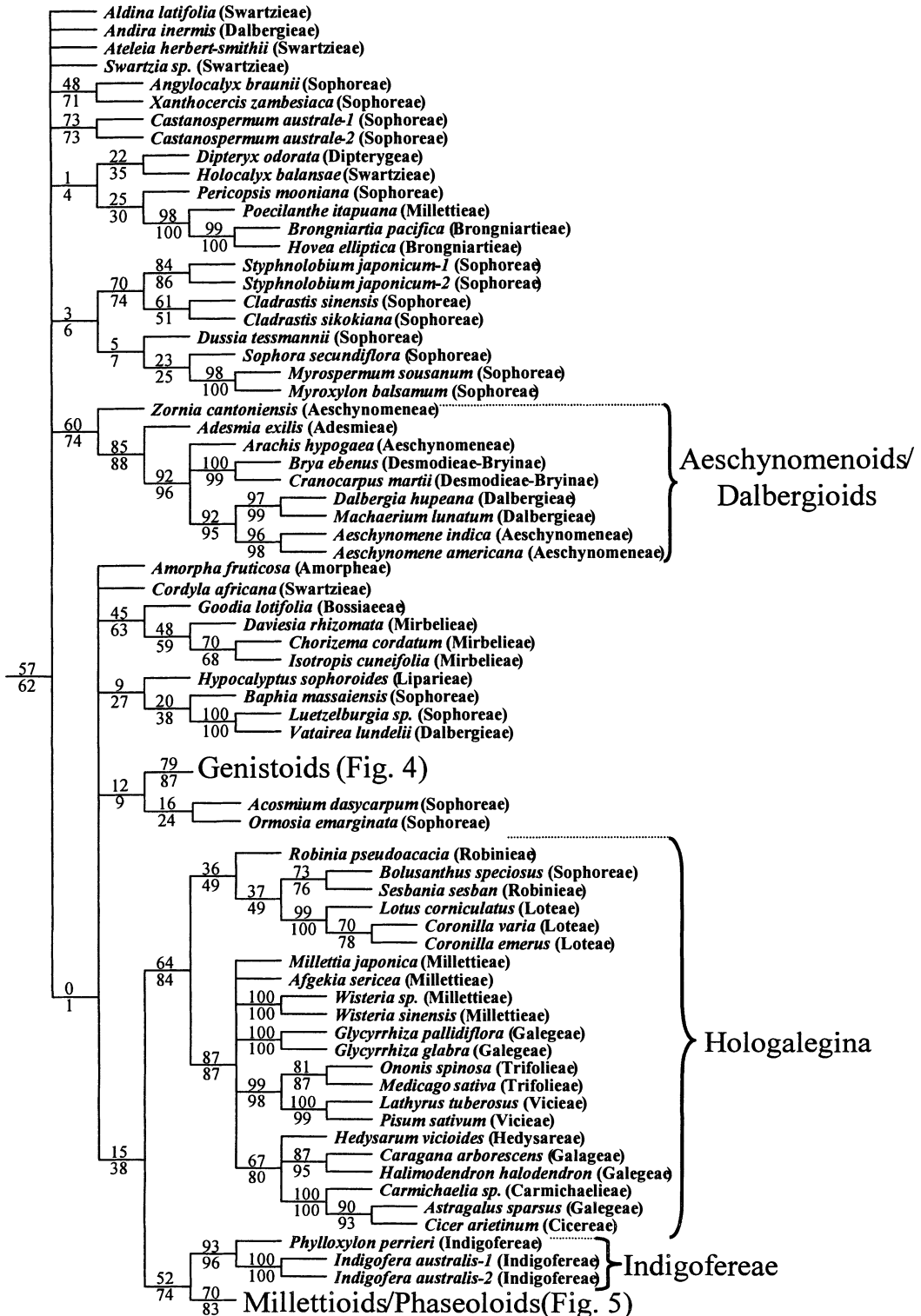


FIG. 3. Backbone of Papilionoideae in the *rbL* strict consensus tree. Large monophyletic groups are indicated.

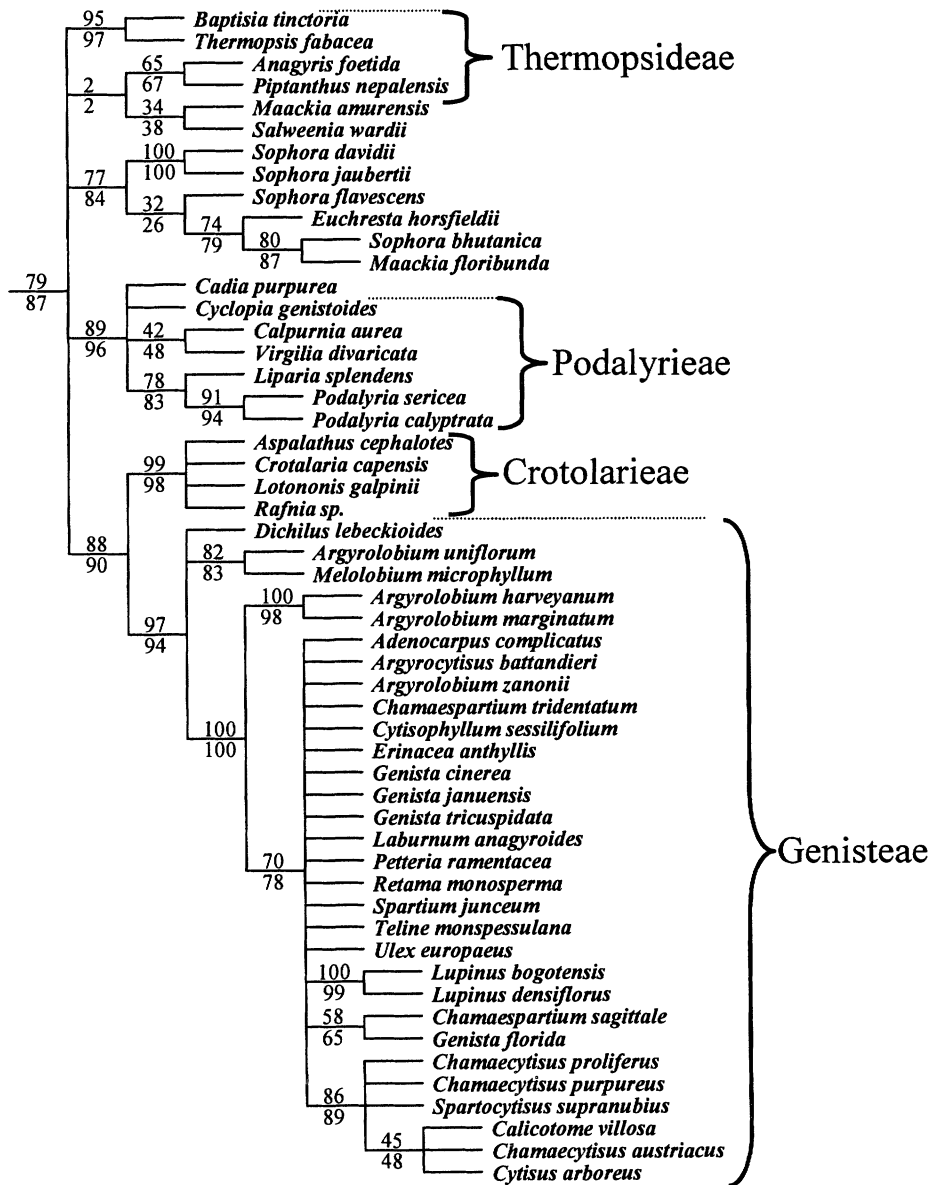


FIG. 4. Strict consensus *rbcL* topology of the genistoid group of tribes. With the exception of *Euchresta* (of the monogeneric tribe Euchresteeae), all taxa for which tribal classification is not given are classified as Sophoreae by Polhill (1994).

parsimony jackknife) Fabales comprising Leguminosae, Polygalaceae, Surianaceae, and *Quillaja* (Fig. 1). Within this clade *Quillaja* was weakly supported as sister to a trichotomy comprising the remaining taxa.

Other families suggested previously to have affinities with legumes were placed more distantly in the topology, and were all included in strongly sup-

ported groups (Fig. 1). These results were consistent with results from other recent molecular analyses (e.g., APG 1998). The previously available *Connarus conchocarpus* sequence and the new sequence from a second member of Connaraceae, *Byrsocarpus coccinea*, were strongly supported as sister taxa, and together were placed within Oxalidales (sensu APG 1998). This is of note, given the non-molecular sim-

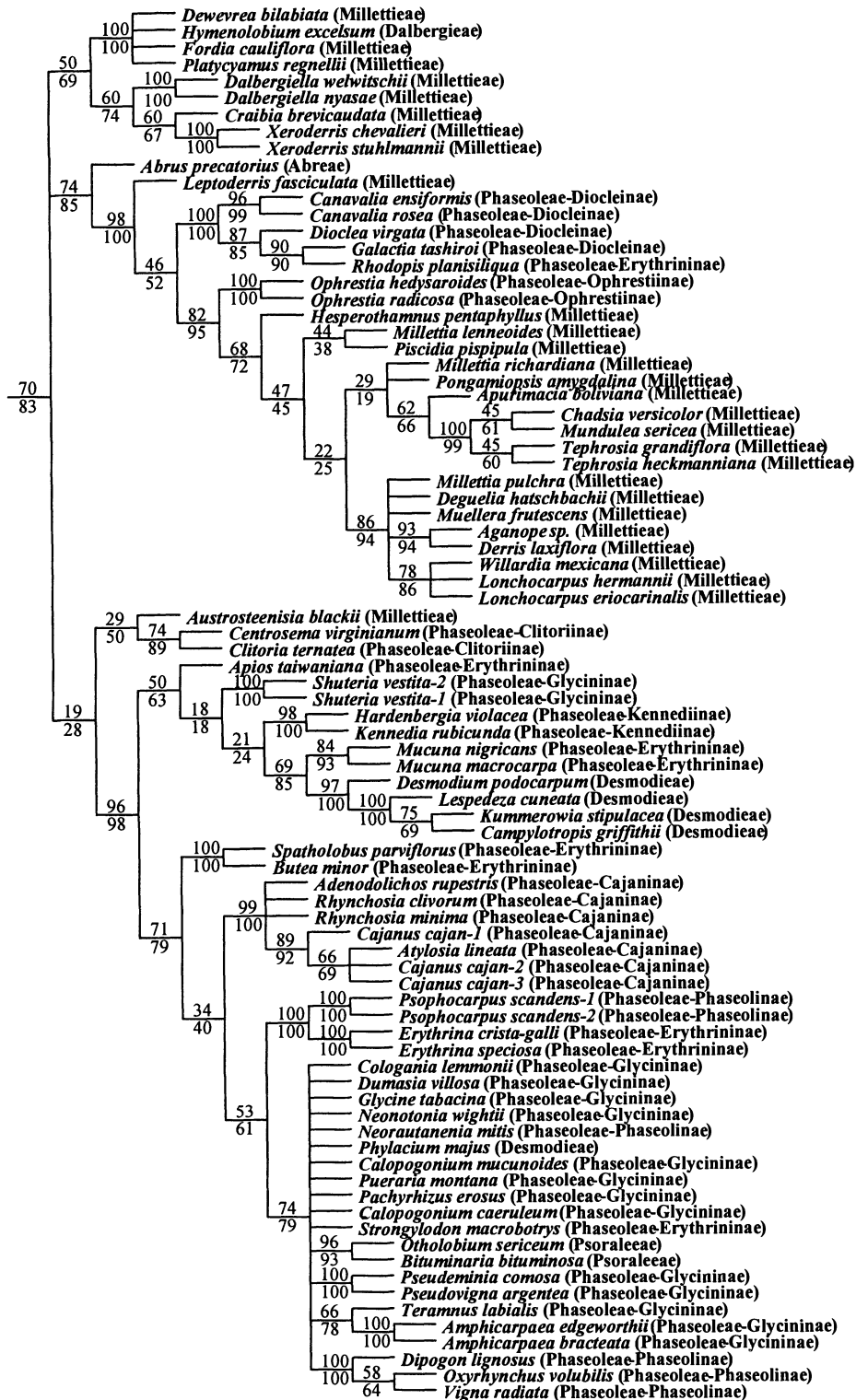


FIG. 5. Millettoid/phaseoloid groups in *rbcL* strict consensus tree. Subtribal designations are given for members of the large tribe Phaseoleae.

ilarities between Connaraceae and Leguminosae (e.g., Dickson 1981; Doyle et al. 2000).

Results using the 18S nuclear ribosomal RNA gene alone have consistently given unexpected results for legumes. For example, in the strict consensus tree of Soltis et al. (1997), the two papilionoid sequences (*Glycine* Willd. and *Pisum* L.) form a clade that is sister to *Clarkia* Pursh (Onagraceae), which in turn is part of a clade that includes Santalales, with mimosoid (*Albizia* Durazz.) and caesalpinoid (*Bauhinia* L.) sequences sister to this entire group. *Polygala* L. is near this clade, but the entire group falls well outside the "nitrogen-fixing" clade. Thus, it remains possible that characters of the nuclear genome, both sequences and non-molecular data such as the characters used by Nandi et al. (1998), could trace a different history than does the chloroplast genome. However, it appears that further studies of cpDNA should continue to support the relationships shown by *rbcl*, a prediction that is being borne out with *trnL* (A. Bruneau, pers. comm.).

***Leguminosae: Monophyly and Relationships Among Subfamilies and Within Caesalpinioideae.***

The overall topology for Leguminosae was consistent with previous *rbcl* analyses, as were relationships within Caesalpinioideae (Fig. 2). This is not surprising, because taxonomic sampling was nearly identical between this analysis and the study of Doyle et al. (1997). In several cases additional sequences were included for taxa already sampled; this was also true for Mimosoideae. New sequences for Caesalpinioideae were added from 1) individual members of Cassieae subtribes Dialiinae and Labicheinae, 2) *Peltogyne* Vog. (Detarieae), and 3) *Acrocarpus* Wight & Arn., the sole member of the *Acrocarpus* group of Caesalpinieae.

The entire family was supported as monophyletic (73% bootstrap/83% jackknife support). This is consistent with the numerous non-molecular synapomorphies that mark the family, most obviously the single carpel that gives rise to the legume fruit, but also the adaxial position of the median petal, endothecium rib spacing less than twice the length of the rib, and parietal placentation (Chappill 1995). Papilionoideae and Mimosoideae were both resolved as monophyletic, and were nested within a paraphyletic Caesalpinioideae. Mimosoideae was nested within the largest Caesalpinioideae clade, with Papilionoideae sister to this clade, though with very weak support.

Caesalpinioideae comprised basal elements in the family, though these relationships were weakly

supported. As in other *rbcl* analyses (e.g., Doyle et al. 1997), a monophyletic Cercideae was sister to the remainder of the family. The next-diverging clade in the Doyle et al. (1997) analysis was also found here, with additional taxa now joining it. All are members of Cassieae, from two subtribes (Labicheinae and Dialiinae). Detarieae of Polhill (1994) formed a clade, though with weak support; groups of genera recognized as tribes Detarieae and Amherstieae (Cowan and Polhill 1981a, 1981b) or as Macrolobieae and Detarieae by Breteler (1995) were not supported as monophyletic, but taxonomic sampling was very incomplete.

The Detarieae divergence was followed by the dichotomy between Papilionoideae and a clade that included the remaining Cassieae, all Caesalpinieae, and Mimosoideae. Within this clade neither of the two remaining caesalpinoid tribes was monophyletic. Cassieae, already paraphyletic due to the placement elsewhere of Labicheinae and Dialiinae, includes *Ceratonia* L. (Cassieae subtribe Ceratoniinae), which was weakly resolved as sister to *Acrocarpus* (Caesalpinieae). This clade was one branch of a large polytomy that also included the three genera of Cassiinae (*Cassia* L., *Senna* Mill., *Chamaecrista* Moench.). As in previous analyses (Doyle et al. 1997), Cassiinae itself was not supported as monophyletic. If this topology is due to more than simply a lack of resolution, it suggests that the very similar floral morphologies shared among the three genera could represent convergence, a possibility also noted by Tucker (1996) based on floral development.

Of the three informal generic groups of Caesalpinieae for which multiple genera were sampled, only the *Gleditsia* group (*Gleditsia* L. and *Gymnocladus* Lam.) was monophyletic. In contrast, neither the two genera sampled from the *Peltophorum* group (*Peltophorum* (Vogel) Benth. and *Delonix* Rafin.), nor the two from the *Caesalpinia* group (*Caesalpinia* L. and *Parkinsonia* L.) formed clades in the strict consensus tree.

Mimosoideae was strongly supported as sister to *Erythrophleum* Afzel ex G. Don (*Dimorphandra* group of Caesalpinieae), and this clade was more weakly supported as sister to *Tachigali* Aubl. (*Sclerolobium* group of Caesalpinieae). These relationships were described previously from *rbcl* data (Doyle et al. 1997), but that study included only one representative of Mimosoideae. A close relationship between the *Dimorphandra* group and Mimosoideae has been hypothesized for some time (see Elias 1981), and is supported by several morphological

and anatomical characters as synapomorphies in cladistic analyses (J. Chappill, pers. comm.). As noted previously (Doyle et al. 1997), this clade groups Mimosoideae, nearly all of which nodulate, with nodulating members of Caesalpinieae, where the syndrome is rare (de Faria et al. 1989; Sprent 2000). Within Mimosoideae, Ingeae formed a monophyletic group with strong support but *Albizia* appeared to be polyphyletic. *Mimosa* and *Acacia* formed a clade with moderate support, contrary to expectations based on morphology that Acacieae and Ingeae should form a natural group (Chappill 1995; Chappill and Maslin 1995; Grimes 1995).

**Papilionoideae: Overall Topology.** As in the Doyle et al. (1997) study, Papilionoideae was resolved as monophyletic in the strict consensus tree, though with relatively low (57%/62%) support (Figs. 2, 3). There is strong non-molecular support for papilionoid monophyly, however, from characters including clawed petals, hilar fissures, subhilar tracheid bars, and storied wood (Chappill 1995). As in the Doyle et al. (1997) analysis, the topology of the subfamily consisted of a number of major clades whose relationships to each other and to numerous individual genera and small clades were mostly unresolved or weakly supported (Fig. 3). However, several of these major clades were themselves relatively well-supported: a genistoid alliance including Genisteae, Thermopsidae, Crotalariaeae, Podalyrieae, and some Sophoreae (79%/87% support); an "aeschynomeneoid/dalbergioid" group primarily composed of Aeschynomeneae, Adesmieae, subtribe Bryinae of Desmodieae, and the *Dalbergia* group of Dalbergieae (60%/74%); a "Hologalegina" (64%/84%) with Robinieae, Loteae, and a "galeoid" clade comprising most of the temperate herbaceous (epulvinate series) tribes (87%/87%); and a "millettioid/phaseoloid" clade comprising Phaseoleae, Desmodieae subtribes Desmodiinae and Lespedezinae, Psoraleae, and most Millettieae (70%/83%). For most of these groups other than the millettioid/phaseoloid clade there is little new information to report here, so these clades and the remaining unresolved elements of the subfamily will be discussed only briefly, with references to more detailed studies using other data where these are available.

**Unresolved Papilionoid Genera and Small Clades.** Sophoreae and Swartzieae, long viewed as polyphyletic basal elements in the subfamily (e.g., Polhill 1981c), made up the bulk of the taxa unresolved along the backbone of the subfamily. Not surprisingly, neither tribe was supported as mono-

phyletic, even if more derived Sophoreae (e.g., *Sophora* s.s.) are not considered. Some individual groupings of Sophoreae genera (*Cladrastis* Rafin./*Styphnolobium* Schott; *Myroxylon* L.f./*Myrospermum* Jacq.) were strongly supported, but most were not, and there was little support for the informal generic groups listed by Polhill (1994). Pennington et al. (2001) have explored relationships in these "basal" groups in some detail using the chloroplast *trnL* region.

Three genera of the Australian tribe Mirbelieae formed a clade, although with low support. This clade was joined, also relatively weakly, by our sequence from *Goodia lotifolia*, a member of the Australian tribe Bossiaeeae. A second sequence of this genus from the Wink group (Z95553, from the same species) was found to be nested within the genistoid clade as sister to *Argyrolobium harveyanum* and *A. marginatum* in our preliminary analyses and was excluded from the final analysis on the assumption that it represented an error.

*Hovea* R. Br., a member of the *Templetonia* group of Bossiaeeae in the classification of Polhill (1994), was strongly supported as sister to *Brongniartia* H. B. & K. (Brongniartieae) in agreement with the results of Crisp and Weston (1987), who suggested the transfer of the *Templetonia* group to the otherwise American tribe Brongniartieae. This clade was joined with very high support by the sequence of *Poecilanthus* Benth., a genus whose affinities have been controversial, being either with Dalbergieae (Lavin 1987) or Millettieae (Geesink 1981; Polhill 1994). *Poecilanthus* has an alkaloid profile (Greinwald et al. 1995) that supports a close relationship with Brongniartieae (J. Chappill, pers. comm.). This relationship is also seen in analyses of the chloroplast *trnK/matK* region for 62 accessions from 57 genera, with primary focus on Millettieae (Hu et al. 2000). In that analysis, however, *Brongniartia* and *Poecilanthus* were part of a strongly supported clade that also included *Ormosia* Jacks., *Acosmium* Schott, and *Bolusanthus* Harms. Such a clade was not observed here; instead, *Bolusanthus* was placed with Robinieae and Loteae (Fig. 3; as in Doyle et al. 1997), whereas *Acosmium* and *Ormosia* formed a weak clade sister (though with negligible support) to the genistoid clade. A third chloroplast analysis, of 122 sequences of the *trnL* intron (Pennington et al. 2001), identified an expanded genistoid group that included (as separate elements) *Acosmium*, *Ormosia*, *Poecilanthus*, *Bolusanthus*, and a genistoid group comparable to that discussed below (*Brongniartia* was not included). Data from nrDNA ITS sequences

support a close relationship between *Poecilanth*e and Brongniartieae (Crisp et al. 2000).

Other tribes whose single representatives were included in this unresolved group were Dipterygeae and Amorpheae. Two genera sampled from the *Andira* group of Dalbergieae, *Andira* Juss. and *Vatairea* Aubl., were not sister to one another. Instead, *Vatairea* was strongly supported as sister to *Luetzelburgia* Harms, a genus classified in the *Myroxylon* group of Sophoreae (Polhill 1994). A relationship between these two genera has been suggested by de Lima (1982, 1990). *Luetzelburgia* apparently does not belong to the *Myroxylon* group of Sophoreae, two of whose genera (*Myroxylon* and *Myrospermum*) show the same close relationship to one another here as seen in previous analyses (Doyle et al. 1997). The relationships of the *Andira* group of Dalbergieae appear to be complex and vary in different analyses (Hu et al. 2000; Lavin et al. 2001; Pennington et al. 2001).

**Genistoids.** This clade was the focus of Käss and Wink's (1997b) study, and nearly all of the sequences included here are from their work. In the present study, as in theirs, this strongly supported clade included Crotalariaeae, Podalyrieae, Thermopsidaeae, Genisteeae, and also some members of the *Sophora* group of Sophoreae (Fig. 4). This composition is similar, but not identical to the "genistoid alliance" of Polhill (1981c, 1994), which also included Mirbelieae, Bossiaeeae, and Brongniartieae.

As shown previously by Käss and Wink (1997b), *Liparia splendens* was nested within Podalyrieae, supporting the union of Liparieae and Podalyrieae. This was recommended by Schutte and van Wyk (1998a); however, they also recommended that *Hypocalyptus* Thunb. and *Cadia* Forsk. be excluded from their unified Podalyrieae. *Hypocalyptus* was unresolved in the *rbcl* tree, but as had been shown in our previous analysis (Doyle et al. 1997), *Cadia* was part of a strongly supported clade of Podalyrieae, suggesting that its inclusion there should be reconsidered.

Podalyrieae s.l., Crotalariaeae, and Genisteeae were all strongly supported as monophyletic, as found previously by Käss and Wink (1997b). In their analysis, Thermopsidaeae comprised two pairs of genera that did not form a clade. In the present strict consensus tree, these two clades again were not supported as forming a monophyletic group. In contrast, nrDNA ITS trees include a strongly supported clade comprising *Thermopsis* R. Br., *Baptisia* Vent., *Anagyris* L., and *Piptanthus* Sweet (Crisp et al. 2000).

The remaining genistoid Sophoreae, including

*Maackia floribunda*, formed a strongly supported clade. *Euchresta* Benn., the single genus of Euchresteeae, was also part of this clade, sister to a group comprising *Sophora bhutanica* and *M. floribunda*. Affinities of *Euchresta* with *S. bhutanica* were suggested by Ohashi (1973), with *Maackia* Rupr. by Chen et al. (1992), with genistoid tribes as a group by Polhill (1981c, 1994), and most recently with *Maackia* and *Sophora* L. by Ohashi et al. (1999). The *rbcl* results reconcile these varying views by demonstrating that all of the taxa in question are closely related. Polhill (1981b) noted that jointed stamens occur in *Euchresta* and sporadically in the *Sophora* group, specifically mentioning *Sophora* s.l., *Calpurnia* E. Mey., and *Salweenia* E. G. Baker. All of these taxa are members of the genistoid clade, though they do not form a single group.

**Aeschynomenoide/Dalbergioide Group.** This group comprises Adesmieae, a paraphyletic Aeschynomeneae, the *Dalbergia* group of Dalbergieae, and subtribe Bryinae of Desmodieae (Fig. 3). A close cladistic relationship among the former three taxa was shown in our previous *rbcl* study (Doyle et al. 1997), which, in excluding *Andira* from this clade, also demonstrated the polyphyletic nature of Dalbergieae, as defined traditionally by the presence of indehiscent fruits with a hardened endocarp (Polhill 1981a). That result was found again here, with the addition that *Vatairea*, a second member of the *Andira* group of Dalbergieae (Polhill 1994), also fell outside this core aeschynomenoide/dalbergioide clade.

The implications of the inclusion of Bryinae in this clade rather than with Desmodieae s.s. are the subject of a separate study (TK and HO, unpublished), so only the history of its placement is mentioned here. Bailey et al. (1995, 1997) showed that *Brya* P. Br. and *Cranocarpus* Benth. retain the intron of the chloroplast gene *rpl2*, the loss of which is an apparent synapomorphy for a core Desmodieae. Based on this it was suggested that Bryinae might better be returned to Aeschynomeneae, from whence they had been transferred by Ohashi et al. (1981). Kajita et al. (1996) corroborated this hypothesis using *rbcl*, although sequences from *Cranocarpus* and *Brya* were not included in Doyle et al. (1997). The aeschynomenoide/dalbergioide group appears in the Pennington et al. (2001) *trnL* phylogeny, and is discussed in detail by Lavin et al. (2000) where it is referred to as "dalbergioide" after that polyphyletic tribe.

**"Hologalegina": the Galegoide Clade, Loteae, and Robinieae.** The largest papilionoid clade consisted of two subclades (Fig. 3). One was a group that

included Robinieae, the temperate herbaceous (TH) group of tribes (Loteae, Galegeae, Hedysareae, Cicereae, Trifolieae, Viciae, Carmichaelieae), and some elements of Millettieae, whereas the second comprised Indigofereae plus a millettoid/phaseoloid group, discussed below, that included Phaseoleae, Desmodieae s.s., Psoraleeae, Abreae, and most Millettieae. Essentially this same dichotomy has been observed previously with *rbcL* (Doyle et al. 1997; Käss and Wink 1997b), and *trnK/matK* (Hu et al. 2000).

Wojciechowski et al. (2000) refer to the group that includes Robinieae and the TH group as "Hologalegina" based on their *matK* and "supertree" analyses; their study discusses relationships among these taxa in detail. In the *rbcL* tree, this clade was divided into two groups (Fig. 3). The first of these, consisting of Robinieae, Loteae, and *Bolusanthus* (Sophoreae), was weakly supported. The placement of *Bolusanthus* in this clade disagrees markedly with results from phytochrome (Lavin et al. 1998), and also from the Hu et al. (2000), *trnK/matK* study, which places this genus strongly within a clade comprising *Brongniartia*, *Acosmium*, *Ormosia*, and *Poecilanthus* (genera discussed above). In the *rbcL* tree *Bolusanthus* is well-supported as sister to *Sesbania* Scop., a placement that contributes to the non-monophyly of Robinieae, whose other sampled genus (*Robinia* L.) is weakly supported as sister to the remainder of this clade. In the Hologalegina *matK* analysis of Wojciechowski et al. (2000) core Robinieae form a robust clade but *Sesbania* is unresolved relative to this group and Loteae. A sister group relationship of *Sesbania* and core Robinieae was proposed by Lavin and Sousa (1995).

The second clade of Hologalegina was a well-supported group that included the remaining tribes of the temperate herbaceous group. This included representatives of Carmichaelieae and Cicereae, tribes not included in previous *rbcL* analyses. The placement of *Carmichaelia* R. Br. and *Astragalus* L. in a strongly supported clade agrees with other data (e.g., Wojciechowski et al. 2000), but such studies place *Cicer* L. elsewhere, with Viciae and Trifolieae. One copy of the large inverted repeat (IR) is known to be absent from the chloroplast genomes of the majority of these taxa (Palmer et al. 1987; Lavin et al. 1990; Liston 1995) and is hypothesized to be missing from the remaining taxa. The group has been demonstrated to be monophyletic in previous *rbcL* studies (Doyle et al. 1997; Käss and Wink 1997b), and also in phylogenetic analyses of other chloroplast genes (Hu et al. 2000; Wojciechowski et

al. 2000) and the nuclear phytochrome gene family (Lavin et al. 1998), though support was low in some of these studies. The condition for the IR is unknown for *Afgekia* Craib (Millettieae), but the *rbcL* results strongly suggest that, like other Millettieae found in this clade (*Wisteria* Nutt., *Millettia japonica*), it lacks the IR. This would be consistent with morphological similarities among these taxa (Lavin et al. 1998).

The exclusion of Loteae from the IR-loss clade and the inclusion of some Millettieae taxa there is apparently inconsistent with morphological characters of these groups, notably the predominant herbaceous habit and leaf morphology that led to the recognition of the "temperate herbaceous" "epulvinate series" of tribes (Polhill 1981c). In equally weighted cladistic analyses of a "general" data set that includes morphological, anatomical, and chemical characters as well as structural characters such as the presence or absence of the IR, Loteae group with the remaining epulvinate tribes to the exclusion of *Wisteria* and presumably allied Millettieae (Chappill 1995 and pers. comm.), so the question of these relationships remains to be fully resolved.

**Millettoid/Phaseoloid Clade.** Sampling for *rbcL* in this study was most comprehensive for phaseoloid and millettoid taxa, due in part to the fact that Phaseoleae and Desmodieae are the taxonomic focus of several of the authors of this study, but also because of the long-standing supposition that some tribes and subtribes of this group are not monophyletic (e.g., Geesink 1981; Lackey 1981). Molecular data have borne out these suggestions for Phaseoleae and at least some of its subtribes (e.g., Doyle and Doyle 1993), and for Millettieae, even excluding those taxa that lack the IR (e.g., Doyle et al. 1997; Lavin et al. 1998; Hu et al. 2000).

The millettoid/phaseoloid clade was well supported by *rbcL* and included core Millettieae, Phaseoleae, core Desmodieae, Psoraleeae, and *Abrus* Adans. (Abreae) (Fig. 5); Indigofereae was sister to this group (Fig. 3). Major groupings within this overall clade included two groups dominated by Millettieae, one of which also contained several Phaseoleae and was sister to *Abrus*, and a strongly supported clade with most Phaseoleae but also including Desmodieae and Psoraleeae.

**MILLETTIEAE.** Relationships within Millettieae have been analyzed by Lavin et al. (1998) based on phytochrome sequences and by Hu et al. (2000) using *trnK/matK*. Readers are referred particularly to the latter paper for discussion of this polyphyletic



tribe. Geesink (1984) made a number of taxonomic recommendations for the tribe, which were used by Hu et al. (2000) but many of which have yet to be incorporated into public databases (e.g., International Plant Names Index, International Legume Database and Information Service, TROPICOS). Because many of his changes involved subsuming taxa into large genera such as *Millettia* Wight. & Arn., we use the earlier names here, but note Geesink's usages as well.

In the *rbcl* strict consensus tree, Millettieae appeared in four places, one with the IR-loss clade of Hologalegina (Fig. 3) and the remainder within the millettoid/phaseoloid clade (Fig. 5). Of the latter group, the position of one genus (*Austrosteenisia* R. Geesink) was essentially unresolved (weakly grouped with Phaseoleae subtribe Clitoriinae as sister to the clade that included most Phaseoleae, but with low support), whereas the remainder were members of two clades that were part of the basal trichotomy of the millettoid/phaseoloid group. The smaller and more weakly supported of these included *Dalbergiella* E. G. Baker and *Xeroderris* Robery (considered part of *Ostryocarpus* Hook f. by Geesink [1984]), which were also basal elements in the Phaseoleae/Millettieae clade of Hu et al. (2000). Other members of this *rbcl* grouping, however, assumed quite different placements in their *trnK/matK* analysis: *Platycyamus regnellii* was strongly supported as sister to core Phaseoleae, and *Fordia* Hemsl. (represented in their analysis by a different species, *F. splendidissima* (Blume ex Miq.) Buijssen) was strongly nested within their core Millettieae. Two other genera of this *rbcl* clade, *Dewevea* Micheli and *Hymenolobium* Benth., were not sampled by Hu et al. (2000); Geesink (1984) did not suggest any particular affinities for the former, whereas the latter is a member of Dalbergieae according to Geesink (1984), and is placed in the *Andira* group of that tribe by Polhill (1994).

The remaining Millettieae were all part of a well-supported clade sister to *Abrus* (Abreae). The position of *Abrus* remains somewhat controversial. Its placement with Aeschynomeneae in Chappill's (1995) analysis persists in analyses of the augmented version of her general data set (J. Chappill, pers. comm.). Disagreement between the two previous comprehensive legume *rbcl* analyses was noted by Doyle et al. (1997), who attributed the placement of *Abrus* by Käss and Wink (e.g., 1997b) within the IR-loss clade to a switch with *Glycyrrhiza* L. (Galegeae), which in the Käss and Wink (1997b) analysis was strongly supported as sister to *Tephro-*

*sia* Pers., the only member of core Millettieae they sampled. This explanation is supported by the placement of *Glycyrrhiza* within the IR-loss clade by Sanderson and Wojciechowski (1996) based on ITS sequence data, by Wojciechowski et al. (2000) based on *matK*, and by the placement of *Abrus* as sister to the clade containing core Millettieae in the *trnK/matK* trees of Hu et al. (2000).

Even within this clade, Millettieae remains paraphyletic. Sequences from most taxa comprised a well supported subclade, strongly supported as sister to members of Phaseoleae subtribe Ophrestinae. However, *Leptoderris fasciculata* was sister to this group plus Phaseoleae subtribe Diocleinae. In discussing the taxonomy of *Leptoderris* Dunn, Geesink (1984) did not suggest explicit relationships. However, despite mentioning an overall similarity to *Derris* Lour., he noted several differences, including some possible morphological plesiomorphies such as a pollen structure similar to the basal group of Millettieae (e.g., *Ostryocarpus*, *Craibia* Dunn).

The clade sister to Phaseoleae-Ophrestinae corresponds generally to Hu et al.'s (2000) core Millettieae, though differences in sampling complicate direct comparison of the two analyses. For example, the *rbcl* data set did not include any members of the *Philenoptera* clade recognized by Hu et al. (2000). Concerning the small genus *Hesperothamnus* Bradegee, sister to remaining core Millettieae in the *rbcl* analysis, Geesink (1984) found "no single morphological character different from *Millettia*" and concluded that it was simply the American representative of the otherwise Old World *Millettia* sect. *Millettia*. *Millettia* species occur throughout the *rbcl* core Millettieae clade as well as that of Hu et al. (2000), who detail the complicated taxonomy of this genus (and also *Derris*), expanding on earlier discussions by Lavin et al. (1998) based on phytochrome results. Among other genera in the *rbcl* core Millettieae clade, *Pongamiopsis* R. Viguier was originally described in *Millettia* and is part of the core *Millettia* group of Hu et al. (2000), and Geesink (1984) suggested that *Deguelia* Aubl. could also be included in an expanded *Millettia*.

Several robust clades occurred within this core Millettieae group, one of which included *Derris* and *Lonchocarpus* Kunth. Also in this group was *Muelleria* L. f., which Geesink (1984) suggested should be considered a derivative of *Lonchocarpus* sect. *Punctati*. The two taxa share pellucid-punctate leaflets, but *Muelleria* has pods specialized for aquatic dispersal. An additional small genus, *Willardia* Rose, was also lumped by Geesink in *Lonchocarpus* with

which it "does not differ in even one constant character" (Geesink 1984). The strong grouping with two species of *Lonchocarpus* in the *rbcl* tree is consistent with his conclusion.

PHASEOLEAE, DESMODIEAE, PSORALEEAE. Phaseoleae as currently defined was polyphyletic (Fig. 5). Subtribes Diocleinae and Ophrestinae were sisters to the core Millettieae, as noted above. Two morphologically quite distinct species of *Ophrestia* H. M. L. Forbes were sampled, and were supported as a clade at 100% bootstrap and jackknife support, and with strong support as sister to core Millettieae. This is consistent with chloroplast genome restriction mapping (cp RFLP) results (Doyle and Doyle 1993; Bruneau et al. 1995) but, as noted by Hu et al. (2000), contrasts with results of their *matK/trnK* analysis, in which the single species sampled (*O. radicata*) grouped with other Phaseoleae. Results for Diocleinae *rbcl* sequences also were consistent with previous cp RFLP studies (Doyle and Doyle 1993; Bruneau et al. 1995), which placed *Dioclea* H. B. & K. and allies in one group (strongly supported with *rbcl*) rather distant from two other genera, *Calopogonium* Desv. and *Pachyrhizus* Rich ex DC., which instead grouped with Glycininae and more recently have been classified there (Polhill 1994).

*Clitoria* L., of Phaseoleae subtribe Clitoriinae, was weakly supported as sister to the main Phaseoleae clade in our previous *rbcl* analysis (Doyle et al. 1997). The addition of a second genus of this small (four genera) subtribe, *Centrosema* (DC.) Benth., did not change that position, but did demonstrate the likely monophyly of the subtribe. Both this naturalness and its position distant from Glycininae and Phaseolinae agree with conclusions reached by Lackey (1981), who suggested that the subtribe might be basal within Phaseoleae. Clitoriinae grouped with *Austrosteenisia* (Millettieae) in a clade sister to the main Phaseoleae clade, but with little support.

The remaining Phaseoleae subtribes were all part of a very strongly supported clade (96%/98%; Fig. 5) that, in agreement with previous *rbcl* results (Doyle et al. 1997), also included core Desmodieae and Psoraleeae. Phaseoleae subtribes Cajaninae and Kennediinae were strongly supported as monophyletic, as was found in cp RFLP studies (Doyle and Doyle 1993; Bruneau et al. 1995). A core Phaseolinae that included three diverse members (*Vigna* Savi, *Dipogon* Lieb., *Oxyrhynchus* Brandegeae) of the group of genera having a 78 kb chloroplast DNA inversion (Bruneau et al. 1990) was also strongly supported. Of the three genera considered by Lack-

ey (1981) to be peripheral members of Phaseolinae and possibly requiring removal from the subtribe, two (*Psophocarpus* DC. and *Neorautanenia* Schinz) were sampled here and neither was part of this clade; the third genus, *Dysolobium* (Benth.) Prain, also did not join the core Phaseolinae and will be discussed elsewhere (TK and YT, unpublished data). *Psophocarpus* is known to lack the 78 kb inversion but was found to be sister to core Phaseolinae in cp RFLP analyses (Doyle and Doyle 1993; Bruneau et al. 1995). No close relationship with core Phaseolinae was supported by *rbcl* data; instead, *Psophocarpus* was sister to *Erythrina* L. (subtribe Erythrinae) with strong support. The poorly known *Neorautanenia* was not sampled in previous molecular studies, and its condition for the 78 kb inversion character is thus unknown. Its *rbcl* sequence was unresolved in the clade that contained Phaseolinae, Glycininae, Psoraleeae, and other taxa, suggesting that it may lack the 78 kb inversion.

Erythrinae has been considered an unnatural "subtribe of convenience" (Lackey 1981), and *rbcl* results were consistent with this conclusion. Only two genera, *Butea* Roxb. and *Spatholobus* Hassk., formed a clade; this grouping was also found by Bruneau et al. (1995) with cp RFLP data. Lackey (1981), citing Solereder (1908), noted that these genera share a four to five celled uniseriate hair type otherwise found in Millettieae. The grouping of *Erythrina* with *Psophocarpus* (Phaseolinae) has already been mentioned, and is intriguing given the long-standing controversy about the relationships of *Erythrina* (e.g., Lackey 1981). *Apios* Fabr. was part of a weakly supported clade that included Desmodieae, and *Mucuna* Adanson was sister to that tribe, as discussed below. *Strongylodon* Vogel, like *Neorautanenia*, was part of the Glycininae/Phaseolinae/Psoraleeae clade. Finally, *Rhodopis* Urban was strongly supported as sister to *Galactia* P. Browne in subtribe Diocleinae, an odd result given its placement, with *Neorudolphia* Britton (not sampled here), as sister to the core Millettieae genera *Lonchocarpus* and *Tephrosia* in cp RFLP studies (Bruneau et al. 1995). The same DNA sample was used for both analyses.

Glycininae was also not supported as monophyletic. *Shuteria* Wight. & Arn. was placed sister to Kennediinae plus core Desmodieae, as in previous studies (Doyle and Doyle 1993; Bruneau et al. 1995). This placement was corroborated with a sequence from a second accession of *S. vestita*. The remaining Glycininae were part of the clade that also included Phaseolinae, Psoraleeae, and one genus of Desmo-

dieae (*Phylacium* Benn.). Few relationships among Glycininae were supported in the strict consensus tree, similar to what was found in previous cp RFLP analyses (Doyle and Doyle 1993; Bruneau et al. 1995).

Relationships of Desmodieae will be discussed in detail elsewhere as part of an *rbcl* study with exhaustive sampling of its genera (TK and HO, unpublished data). The polyphyly of a Desmodieae that includes subtribe Bryinae, which has aeschynomeneoid/dalbergioid affinities, was noted above, where reference was made to the loss of the chloroplast *rpl2* intron as a diagnostic character for a core Desmodieae. The *rbcl* results were consistent with *rpl2* intron data in identifying a strongly supported core Desmodieae. Moreover, *Mucuna* (Phaseoleae-Erythrinae), like core Desmodieae taxa, lacks the *rpl2* intron (Bailey et al. 1997), and was strongly supported as sister to core Desmodieae in the *rbcl* trees. This is in contrast to *Phylacium*, which, although it is a member of Desmodieae, retains the intron and is part of the Glycininae/Phaseolinae/Psoraleae clade in the *rbcl* tree. We were unable to obtain a usable *rbcl* sequence from *Neocolletia* Hemsl., another member of Desmodieae subtribe Lespedezinae which also retains the *rpl2* intron (Bailey et al. 1997).

**The Value of *rbcl*: Past, Present, and Future.** As has been the case for angiosperms as a whole and within other angiosperm families, *rbcl* has played a very useful role as systematists have sought to achieve a basic understanding of broad groups of taxa. It has been clear nearly from the beginning of such studies that this single relatively slowly evolving gene, representing only one plant organelle, is incapable of resolving all issues by itself (e.g., Chase et al. 1993). However, the framework it has provided, sometimes only sketchily, is a source of hypotheses that can be tested by additional sampling, often with more rapidly evolving genes, multiple genes, or combinations of molecular and other data.

This has certainly been the case, and continues to be true, in legumes, where *rbcl* provided some of the first gene sequence data for testing broad hypotheses based on earlier molecular work or hypotheses based intuitively or objectively on non-molecular characters. Groups previously hypothesized to be monophyletic, such as an IR-loss clade (e.g., Lavin et al. 1990), were corroborated with *rbcl* (Doyle 1995; Doyle et al. 1997; Käss and Wink 1997b). Other relationships were revealed by *rbcl* that had not been identified previously, such as the

grouping of Aeschynomeneae/Adesmiae with core Dalbergiae (Doyle et al. 1997). In both of these examples, others have taken the story much further, most recently Wojciechowski et al. (2000) in the case of the IR-loss group and Lavin et al. (2001) on the aeschynomeneoid/dalbergioid taxa.

The *rbcl* gene, though not providing as much variation as more rapidly evolving sequences such as chloroplast intergenic spacers or the nrDNA ITS region, does have advantages in its lack of length variation in all but the 3' region. Thus, it is suitable for broad studies that include divergent taxa, particularly outgroups, and remains a useful tool in such studies.

Finally, combination of *rbcl* and other characters, such as was done by Nandi et al. (1998), should be a useful approach for higher level studies in legumes, using as a foundation the original "general" data set of Chappill (1995). The *rbcl* results are providing a guide for definitions of terminal taxa and character scoring in a data set that currently includes over 300 characters and 97 terminal taxa (J. Chappill, pers. comm.). Thus far, these characters have been used in preliminary analyses and have been mapped onto the *rbcl* tree (Doyle et al. 2000); the goal is a simultaneous analysis of the general and *rbcl* data sets.

**ACKNOWLEDGEMENTS.** TK received support from the JSPS Tropical Bio-resources Research Fund. JJD acknowledges long term support from the NSF Systematic Biology program, and thanks Jane Doyle for her many contributions to this project. We thank Dr. Mario S. Sousa, MEXU, and TUS for providing plant material. We are grateful to Toby Pennington, Jenny Chappill, Anne Bruneau, and Marty Wojciechowski for sharing unpublished results. We thank Helga Ochoterena for valuable consultations on phylogenetic analyses, and Kevin Nixon for making updated versions of his analytical programs available to us. Two reviewers and Aaron Liston provided helpful comments on an initial draft of the manuscript.

#### LITERATURE CITED

- ANGIOSPERM PHYLOGENY GROUP (APG). 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.
- BAILEY, C. D., J. J. DOYLE, T. KAJITA, T. NEMOTO, and H. OHASHI. 1995. Losses of the chloroplast *rpl2* intron and ORF184 in the legume tribe Desmodieae. *American Journal of Botany* 82(6, Supplement): 112 (abstract).
- \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_. 1997. The chloroplast *rpl2* intron and ORF184 as phylogenetic markers in the legume tribe Desmodieae. *Systematic Botany* 22: 133–138.

- BRUNEAU, A., J. J. DOYLE, and J. D. PALMER. 1990. A chloroplast DNA structural mutation as a subtribal character in the Phaseoleae (Leguminosae). *Systematic Botany* 15: 378–386.
- , J. L. DOYLE, and J. J. DOYLE. 1995. Phylogenetic evidence in Phaseoleae: evidence from chloroplast restriction site characters. Pp. 309–330 in *Advances in legume systematics, part 7: phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- BRETELER, F. J. 1995. The boundary between Amherstieae and Detarieae (Caesalpinioideae). Pp. 53–61 in *Advances in legume systematics, Part 7: Phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- CHAPPILL, J. A. 1995. Cladistic analysis of the Leguminosae: the development of an explicit hypothesis. Pp. 1–10 in *Advances in legume systematics, part 7: phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- and B. R. MASLIN. 1995. A phylogenetic assessment of tribe Acacieae. Pp. 77–99 in *Advances in legume systematics, part 7: phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- CHASE, M. W., et al. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden* 80: 528–580.
- CHEN, C., H. SUN, and M. MIZUO. 1992. On the genus *Euchresta* Benn. (Leguminosae) with 'Wallace's Line'. *Acta Phytotaxonomica Sinica* 30: 43–56.
- COWAN, R. S. and R. M. POLHILL. 1981a. Amherstieae. Pp. 135–142 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- and ———. 1981b. Detarieae. Pp. 117–134 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- CRISP, M. D., S. GILMORE, and B-E. VAN WYK. 2000. Molecular phylogeny of the genistoid tribes of papilionoid legumes. Pp. 249–276 in *Advances in legume systematics, part 9*, eds. P. S. Herendeen and A. Bruneau. Kew: Royal Botanic Gardens.
- and P. H. WESTON. 1987. Cladistics and legume systematics, with an analysis of the Bossiaeeae, Brongniartieae and Mirbelieae. Pp. 65–130 in *Advances in legume systematics, part 3*, ed. C. H. Stirton. Kew: Royal Botanic Gardens.
- CRONQUIST, A. 1981. *An integrated system of classification of flowering plants*. New York: Columbia University Press.
- DAVIS, J. I., M. P. SIMMONS, D. W. STEVENSON, and J. F. WENDEL. 1998. Data decisiveness, data quality, and incongruence in phylogenetic analysis: An example from the monocotyledons using mitochondrial *atpA* sequences. *Systematic Biology* 47: 282–310.
- DE FARIA, S. M., G. P. LEWIS, J. I. SPRENT, and J. M. SUTHERLAND. 1989. Occurrence of nodulation in the Leguminosae. *New Phytologist* 111: 607–619.
- DE LIMA, H. C. 1982. Revisão taxonômica do gênero *Vatairea* Aublet (Leguminosae-Faboideae). *Arquivos do Jardim Botânico do Rio de Janeiro* 26: 173–214.
- . 1990. Tribo Dalbergieae (Leguminosae-Papilionoideae)—morfologia do frutos, sementes e plântulas e sua aplicação na sistemática. *Arquivos do Jardim Botânico do Rio de Janeiro* 30: 1–42.
- DICKISON, W. C. 1981. The evolutionary relationships of the Leguminosae. Pp. 35–54 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- DOYLE, J. J. 1995. DNA data and legume phylogeny: a progress report. Pp. 11–30 in *Advances in legume systematics, Part 7: Phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- and J. L. DOYLE. 1993. Chloroplast DNA phylogeny of the papilionoid legume tribe Phaseoleae. *Systematic Botany* 18: 309–327.
- , J. A. CHAPPILL, C. D. BAILEY, and T. KAJITA. 2000. Towards a comprehensive phylogeny of legumes: evidence from *rbcL* sequences and non-molecular data. Pp. 1–20 in *Advances in legume systematics, part 9*, eds. P. S. Herendeen and A. Bruneau. Kew: Royal Botanic Gardens.
- , 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.
- , ———, and J. D. PALMER. 1995. Multiple independent losses of two genes and one intron from legume chloroplast genomes. *Systematic Botany* 20: 272–294.
- ELIAS, T. S. 1981. Mimosoideae. Pp. 143–151 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- FARRIS, S. J., V. A. ALBERT, M. KÄLLERSJÖ, D. LIPSCOMB, and A. G. KLUGE. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.
- GEESINK, R. 1981. Tephrosieae. Pp. 245–260 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- . 1984. *Scala Millettiarum*. Leiden Botanical Series 8: 1–131.
- GOLOBOFF, P. A. 1994. NONA: A tree searching program. Program and documentation. Available at ftp.unt.edu.ar/pub/parsimony.
- . 1999. Analyzing large data sets in reasonable time: solutions for composite optima. *Cladistics* 15: 415–428.
- GREINWALD, R., P. BACHMANN, G. P. LEWIS, L. WITTE, and F.-C. CZYGAN. 1995. Alkaloids of the genus *Poecilanthè* (Leguminosae: Papilionoideae). *Biochemical Systematics and Ecology* 23: 547–553.
- GRIMES, J. 1995. Generic relationships of Mimosoideae tribe Ingeae, with emphasis on the New World *Pithecellobium*-complex. Pp. 101–121 in *Advances in legume systematics, part 7: phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- GUINET, PH. 1981. Mimosoideae: the characters of their

- pollen grains. Pp. 835–857 in *Advances in legume systematics, part 2*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- HU, J.-M., M. LAVIN, M. F. WOJCIECHOWSKI, and M. J. SANDERSON. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast *trnK/matK* sequences and its implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418–430.
- KAJITA, T., H. OHASHI, T. NEMOTO, C. D. BAILEY, and J. J. DOYLE. 1996. Polyphyly of the legume tribe Desmodieae: evidence from the chloroplast genome. *American Journal of Botany* 83(6, Supplement): 165–166 (abstract).
- KÄLLERSJÖ, M., J. S. FARRIS, M. W. CHASE, B. BREMER, M. F. FAY, C. J. HUMPHRIES, G. PETERSEN, O. SEBERG, and K. BREMER. 1998. Simultaneous parsimony jackknife analysis of 2538 *rbcl* DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Plant Systematics and Evolution* 213: 259–287.
- KÄSS, E., and M. WINK. 1995. Molecular phylogeny of the Papilionoideae (family Leguminosae): *rbcl* gene sequences versus chemical taxonomy. *Botanica Acta* 108: 149–162.
- and ———. 1996. Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcl*-sequences. *Biochemical Systematics and Ecology* 24: 365–378.
- and ———. 1997a. Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the *rbcl* gene and ITS 1 + 2 regions of rDNA. *Plant Systematics and Evolution* 208: 139–167.
- and ———. 1997b. Phylogenetic relationships in the Papilionoideae (family Leguminosae) based on nucleotide sequences of cpDNA (*rbcl*) and ncDNA (ITS 1 and 2). *Molecular Phylogenetics and Evolution* 8: 65–88.
- LACKEY, J. A. Phaseoleae. 1981. Pp. 301–327 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- LAVIN, M. 1987. A cladistic analysis of the tribe Robinieae (Papilionoideae, Leguminosae). Pp. 31–64 in *Advances in legume systematics, part 3*, ed. C. H. Stirton. Kew: Royal Botanic Gardens.
- and M. SOUSA. 1995. Phylogenetic systematics and biogeography of the tribe Robinieae. *Systematic Botany Monographs* 45: 1–165.
- , 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.
- , 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.
- , 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 pantropical monophyletic clade. *American Journal of Botany* 88: 503–533.
- LISTON, A. 1995. Use of the polymerase chain reaction to survey for the loss of the inverted repeat in the legume chloroplast genome. Pp. 31–40 in *Advances in legume systematics, part 7: Phylogeny*, eds. M. D. Crisp and J. J. Doyle. Kew: Royal Botanic Gardens.
- NANDI, O. I., M. W. CHASE, and P. K. ENDRESS. 1998. A combined cladistic analysis of angiosperms using *rbcl* and non-molecular data sets. *Annals of the Missouri Botanical Garden* 85: 137–212.
- NIXON, K. C. 1999a. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics* 15: 407–414.
- . 1999b. Winclada (beta) version 0.9. Published by the author, Ithaca, NY. (Available at <http://www.cladistics.com>).
- OHASHI, H. 1973. Euchrestae, a new tribe of the family Leguminosae. *Journal of Japanese Botany* 48: 225–234.
- , R. M. POLHILL, and B. G. SCHUBERT. 1981. Desmodieae. Pp. 292–300 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- , K. TANAKA, and T. NEMOTO. 1999. Embryological studies in *Euchresta japonica* (Leguminosae). *Journal of Japanese Botany* 74: 84–95.
- PALMER, J. D., B. OSORIO, J. ALDRICH, and W. F. THOMPSON. 1987. Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. *Current Genetics* 11: 275–286.
- 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 intron *trnL*. *Systematic Botany* 26: 537–556.
- POLHILL, R. M. 1981a. Dalbergieae. Pp. 233–242 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- . 1981b. Euchrestae. P. 402 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- . 1981c. Papilionoideae. Pp. 191–208 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- . 1994. Classification of the Leguminosae. Pp. xxxv–lvii in *Phytochemical dictionary of the Leguminosae, Volume 1: Plants and their constituents*, eds. F. A. Bisby, J. Buckingham, and J. B. Harborne. London: Chapman and Hall.
- QIU, Y.-L., J. LEE, F. BERNASCONI-QUADRONI, D. E. SOLTIS, P. S. SOLTIS, P. S., M. ZANIS, E. A. ZIMMER, Z. CHEN, V. SAVOLAINEN, and M. W. CHASE. 1999. The earliest angiosperms: evidence from mitochondrial, plastid, and nuclear genomes. *Nature* 402: 404–406.
- SANDERSON, M. J. and M. F. WOJCIECHOWSKI. 1996. Diver-

- sification rates in a temperate legume clade: are there "so many species" of *Astragalus* (Fabaceae)? *American Journal of Botany* 83: 1488–1502.
- SAVOLAINEN, V., M. W. CHASE, S. B. HOOT, C. M. MORTON, D. E. SOLTIS, C. BAYER-CLEMENS, M. F. FAY, A. Y. DE BRUIJN, S. SULLIVAN, and Y.-L. QIU. 2000. Phylogenetics of flowering plants based on combined analysis of plastid *atpB* and *rbcl* gene sequences. *Systematic Biology* 49: 306–362.
- SCHUTTE, A. L. and B. E. VAN WYK. 1998a. Evolutionary relationships in the Podalyrieae and Liparieae (Fabaceae) based on morphological, cytological, and chemical evidence. *Plant Systematics and Evolution* 209: 1–31.
- and ———. 1998b. The tribal position of *Hypoclyptus* Thunberg (Fabaceae). *Novon* 8: 178–182.
- SOLEREDER, H. 1908. *Systematic anatomy of the Dicotyledons*. Oxford: Clarendon Press.
- 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.
- , ———, D. L. NICKRENT, L. A. JOHNSON, W. J. HAHN, S. B. HOOT, J. A. SWEERE, R. K. KUZOFF, K. A. KRON, M. W. CHASE, S. M. SWENSEN, E. A. ZIMMER, S.-M. CHAW, L. J. GILLESPIE, W. J. KRESS, and K. J. SYTSMA. 1997. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 84: 1–49.
- SOLTIS, P. S., D. E. SOLTIS, and M. W. CHASE. 1999. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature* 402: 402–404.
- SPRENT, J. I. 2000. Nodulation as a taxonomic tool. Pp. 21–43 in *Advances in legume systematics, part 9*, eds. P.S. Herendeen and A. Bruneau Kew: Royal Botanic Gardens.
- THORNE, R. F. 1992. Classification and geography of the flowering plants. *Botanical Review* 58: 255–348.
- TUCKER, S. C. 1996. Trends in evolution of floral ontogeny in *Cassia* sensu stricto, *Senna*, and *Chamaecrista* (Leguminosae: Caesalpinioideae: Cassieae: Cassiinae); a study in convergence. *American Journal of Botany* 83: 687–711.
- TURNER, B. L. 1981. Thermopsidaeae. Pp. 233–242 in *Advances in legume systematics, part 1*, eds. R. M. Polhill and P. H. Raven. Kew: Royal Botanic Gardens.
- 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. Pp. 277–298 in *Advances in legume systematics, part 9*, eds. P. Herendeen and A. Bruneau. Kew: Royal Botanic Gardens.