

Wojciechowski, M.F. (2003). Reconstructing the phylogeny of legumes (Leguminosae): an early 21st century perspective In: B.B. Klitgaard and A. Bruneau (editors). *Advances in Legume Systematics*, part 10, Higher Level Systematics, pp. 5–35. Royal Botanic Gardens, Kew.

RECONSTRUCTING THE PHYLOGENY OF LEGUMES (LEGUMINOSAE): AN EARLY 21ST CENTURY PERSPECTIVE

MARTIN F. WOJCIECHOWSKI

Department of Plant Biology, Arizona State University, Tempe, Arizona 85287 USA

Abstract

Elucidating the phylogenetic relationships of the legumes is essential for understanding the evolutionary history of events that underlie the origin and diversification of this family of ecologically and economically important flowering plants. In the ten years since the Third International Legume Conference (1992), the study of legume phylogeny using molecular data has advanced from a few tentative inferences based on relatively few, small datasets into an era of large, increasingly multiple gene analyses that provide greater resolution and confidence, as well as a few surprises. Reconstructing the phylogeny of the Leguminosae and its close relatives will further advance our knowledge of legume biology and facilitate comparative studies of plant structure and development, plant-animal interactions, plant-microbial symbiosis, and genome structure and dynamics.

Phylogenetic relationships of Leguminosae — what has been accomplished since ILC-3?

The Leguminosae (Fabaceae), with approximately 720 genera and more than 18,000 species worldwide (Lewis *et al.*, in press) is the third largest family of flowering plants (Mabberley, 1997). Although greater in terms of the diversity of forms and number of habitats in which they reside, the family is second only perhaps to Poaceae (the grasses) in its agricultural and economic importance, and includes species used for foods, oils, fibre, fuel, timber, medicinals, numerous chemicals, cultivated horticultural varieties, and soil enrichment. Ecologically, the family is important in a diversity of ecosystems, especially members of the subfamily Papilionoideae which are present and often dominant, in nearly every vegetation type on Earth, from tropical rain forests to deserts and alpine tundra, and play a vital role in global biogeochemistry; nearly all species root-nodulate with symbiotic bacteria to fix atmospheric nitrogen (Sprent and McKey, 1994; Sprent, 2001). Given the paramount ecological and economic importance of legumes, the evolution and systematics of the family are topics of long-standing interest to a large community of academic, agricultural and government researchers representing a broad spectrum of scientific disciplines.

The years since the Third International Legume Conference (ILC-3; Kew, 1992) have been especially productive ones for all areas of legume biology, from paleobotanical, anatomical, and phylogenetic studies, to studies dealing with crop diversity and nitrogen biology. The set of volumes in the “*Advances in Legume Systematics*” series published since the conference attest to this ongoing and ever increasing interest

* author for correspondence: mfwojciechowski@asu.edu

in legume systematics and comparative biology (Herendeen and Dilcher, 1992; Ferguson and Tucker, 1994; Sprent and McKey, 1994; Crisp and Doyle, 1995; Pickersgill and Lock, 1996; Herendeen and Bruneau, 2000). Furthermore, a cursory survey of the primary literature (Current Contents, on-line version) on legume systematics, phylogenetics, biogeography, and evolution (using search terms such as “phylogen#” and “systematic#”) reveals a minimum of ten articles published per year since 1993, and most in major journals (e.g., *American Journal of Botany*, *Plant Systematics and Evolution*, *Systematic Botany*, *International Journal of Plant Sciences*, *Molecular Phylogenetics and Evolution*, *Biochemical Systematics and Ecology*, *Theoretical and Applied Genetics*). Then, and more so now, the contributions of molecular data have played a prominent role in this progress, a role that is only likely to increase as we venture into the “post-genomics” era. In a larger context, this increase parallels the unprecedented advancements we have witnessed in the past ten years in our knowledge of the evolution and relationships of all green plants, beginning with the publication of the 500 taxon analysis of *rbdL* sequences by Chase *et al.* (1993), followed by several multi-gene analyses for most of the major lineages of vascular plants (e.g., Bowe *et al.*, 2000; Chaw *et al.*, 2000; Nickrent *et al.*, 2000; Soltis *et al.*, 2000; Pryer *et al.*, 2001). From these studies came the surprising result that Leguminosae had an unequivocal sister group relationship to Polygalaceae and Surianaceae. This period also saw the first attempts at resolving higher-level relationships within the family using cladistic analyses of morphological (Chappill, 1995) and molecular data (*rbdL* sequences, Käss and Wink, 1996; Doyle *et al.*, 1997). At the same time, numerous molecular studies were conducted on a diversity of genera and groups within the family.

Rather than focusing primarily on the results of intensive analyses of one molecular locus and their implications, this paper will attempt to highlight results from the large number of phylogenetic studies published since the Third International Legume Conference, especially those that have been published in the last 2 to 3 years. Many of these studies are now at the point where they have begun to explore additional, and some multiple, sources of molecular data for phylogenetic inference at all taxonomic levels. While most of these studies are reporting results from analyses of sequence data sets, the results from analyses of other kinds of molecular data (i.e., DNA fragment analysis) will be included only where relevant to the story. In this paper, discussion will be centred primarily on higher level phylogenetic relationships. Then, I conclude by presenting preliminary estimates for the ages of some of the major clades of legumes by reconstructing divergence times from molecular sequence data, using methods that incorporate variation in the rates of molecular evolution directly into phylogenetic inference procedures.

Despite a number of insights into higher level relationships of the family derived from analysis of the *rbdL* gene (Doyle, 1995; Doyle *et al.*, 1997), many issues in legume phylogeny remain unresolved. This is particularly true for the relationships among the larger clades. The need for more variable genes and non-coding sequences, alone or in combination with morphological data, to gain resolution was evident. To this end, many studies (including several in this volume) have explored a number of other chloroplast genes, and more recently a few nuclear genes as well, for phylogeny reconstruction within the family. Chloroplast sequences that are being used include the *matK* gene and flanking *trnK* intron (e.g., Hu *et al.*, 2000; Lavin *et al.*, 2000; Miller and Bayer, 2001), *trnL* intron and adjacent intergenic spacers (e.g., Bruneau *et al.*, 2000, 2001; Ireland *et al.*, 2000; Luckow *et al.*, 2000; Brouat *et al.*, 2001), *rpoC1/rpoC2* genes (Liston and Wheeler, 1994) and *rpoC1* intron (Downie *et al.*, 1998), *atpB-rbdL* intergenic spacer (Hurr *et al.*, 1999; Doi *et al.*, 2002), *psbA-trnH* intergenic spacer (Chandler *et al.*, 2001), *ndhF* and *rpl16* genes (Schnabel and Wendel, 1998; Davis *et al.*, 2002), and *psaA* genes (M. Wojciechowski and M. Sanderson, University of California, Davis, unpublished data). While most phylogenetic studies that have used nuclear regions have focused on the genes and the internal and external transcribed spacers of ribosomal DNA (e.g., Bena *et al.*, 1998; Aïnouche and Bayer, 1999; Wagstaff

et al., 1999; Wojciechowski *et al.*, 1999; Allan and Porter, 2000), several single or low copy nuclear genes, including the histone H3 intron (Doyle *et al.*, 1996b; Miller and Bayer, 2000), phytochrome genes (Lavin *et al.*, 1998), photohemagglutinin genes (Zink *et al.*, 1994), floral development gene *LEAFY* (Archambault and Bruneau, 2001), and gibberellic acid 20-oxidases (Steele *et al.*, 2000b; Yang *et al.*, 2001), are increasingly being investigated. In addition, analyses of chloroplast and nuclear DNA restriction fragment length polymorphisms (e.g., Watson *et al.*, 2000) and other kinds of DNA markers (e.g., RAPDs, AFLPs, ISSRs; Wolfe and Liston, 1998) continue to provide other systematically relevant molecular data on relationships in legumes especially at the species level, and below (e.g., Travis *et al.*, 1996; Gauthier *et al.*, 1997; Asmussen and Liston, 1998; Gillies and Abbott, 1998; Ramírez *et al.*, 1999; Robinson and Harris, 2000; van de Wouw *et al.*, 2001).

Relationships of Leguminosae to other Angiosperms

Within angiosperms, Leguminosae have traditionally been considered 'rosid' in their taxonomic relationships, with Connaraceae, Chrysobalanaceae, Crossosomataceae and Krameriaceae; and even Sapindaceae has been suggested as closely related (potential sister groups) by different authors at one time or another (reviewed by Dickison, 1981; Thorne, 1992). The first challenge to these hypotheses came from the now landmark study of chloroplast *rbdL* gene sequences by Chase and coworkers (1993) which placed Polygalaceae and Surianaceae as the sister groups to Leguminosae, while those traditionally suggested families were all resolved as more distantly related. This close relationship of Polygalaceae to Leguminosae suggested by the Chase *et al.* study has been substantiated by most subsequent analyses, such as the 2,538 taxa *rbdL* dataset (Källersjö *et al.*, 1998), combined chloroplast *atpB* and *rbdL* study (Savolainen *et al.*, 2000), and combined *atpB*, *rbdL*, and nuclear 18S rDNA analysis (APG, 1998; Soltis *et al.*, 2000). All these studies consistently show moderate to strong support for a monophyletic Fabales comprising a monophyletic Leguminosae sister to a clade of Surianaceae and Polygalaceae (Fig. 1). Only the results of Soltis *et al.* (1997) using the 18S nuclear rDNA gene alone, and that of Nandi *et al.* (1998), which included separate as well as combined analyses of *rbdL* and non-molecular (chemistry, development, morphology) data, came to several, sometimes significantly different, conclusions, most of which were not well supported. Additional studies on smaller sets of taxa, such as a *rbdL* analysis of the "nitrogen-fixing" clade (Soltis *et al.*, 1995) and a *matK* analysis of rosids (Steele *et al.*, 2000a), have shown the rosacean genus *Quillaja* to be a strongly supported member of Fabales, but weakly supported as sister to the clade containing Leguminosae, Polygalaceae, and Surianaceae. More recently, Persson (2001) found support (albeit weak) for a monophyletic Leguminosae-Quillajaceae-Surianaceae as the sister group to a strongly supported monophyletic Polygalaceae using *trnL-F* sequences. Clearly, resolving the sister group relationships of the family is an important issue. The three-gene analyses of Soltis *et al.* (2000) showed Fabales to be one of four well-supported subclades (along with Rosales, Cucurbitales, and Fagales) of a larger clade, nested within Eurosid I, a clade comprising all angiosperms that form symbiotic associations with nitrogen-fixing bacteria involving specialised root nodules. The strong support for this nitrogen-fixing clade reinforces earlier molecular evidence for the hypothesis of a single origin for the predisposition of nitrogen-fixing symbioses within angiosperms (Soltis *et al.*, 1995).

Phylogenetic relationships within Leguminosae: subfamilies and major clades

Attempts to unravel the phylogeny of the family through cladistic analyses of molecular data date back more than a decade. The analysis of higher level phylogenetic relationships in Leguminosae based on sequence data began with the chloroplast gene *rbdL* (Doyle, 1995; Doyle *et al.*, 1997; Käss and Wink, 1995, 1996, 1997a), a logical choice given the early, widespread use of this gene for a comprehensive hypothesis of

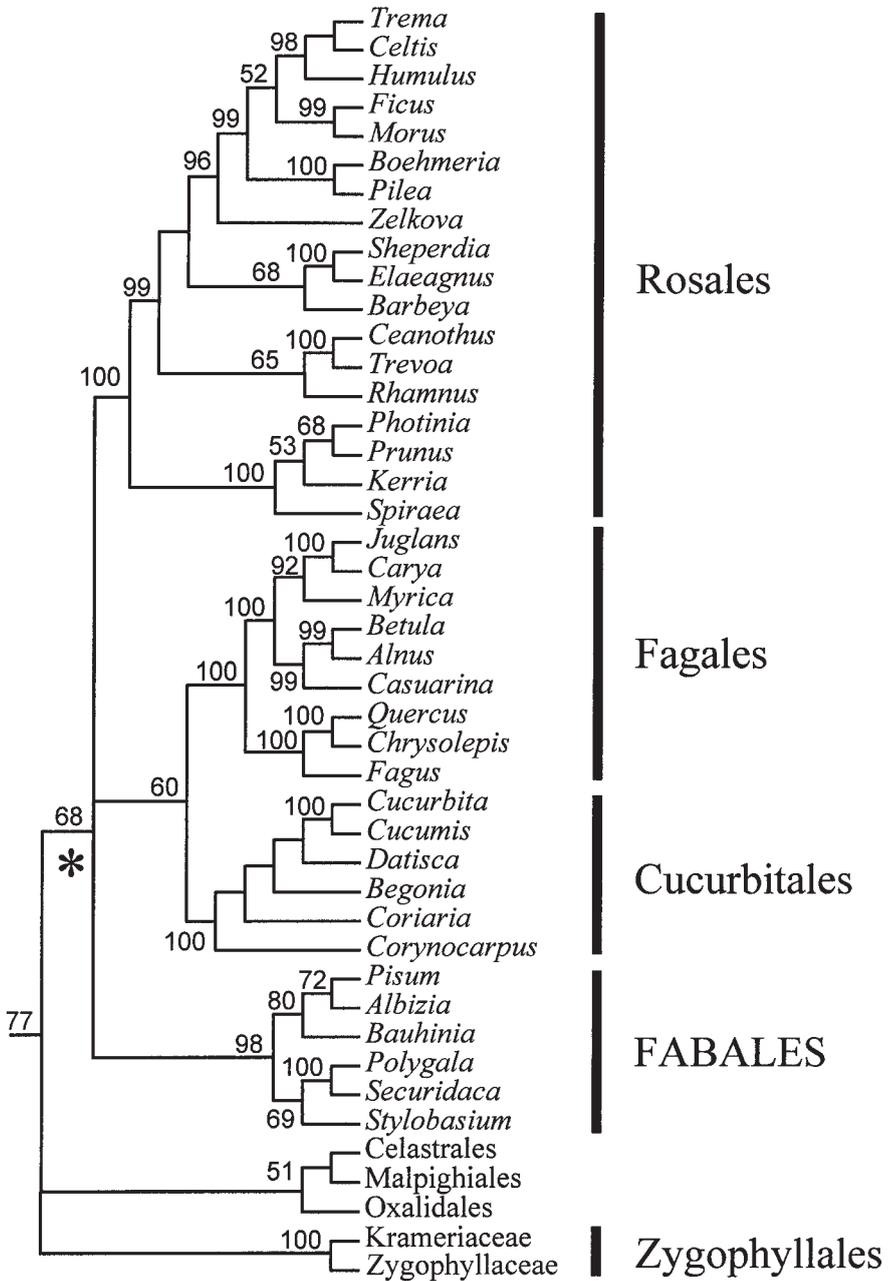


FIG. 1. Phylogenetic relationships of Rosales, Fagales, Cucurbitales, and Fabales portion of Eurosoid I clade of angiosperms, inferred from parsimony analyses of 18S rDNA, *rbcl*, and *atpB* sequences (modified from Soltis *et al.*, 2000, Fig. 7). Support values for nodes from jackknife consensus tree > 50% are indicated. *Denotes the clade of all angiosperms forming symbioses with nitrogen-fixing bacteria.

angiosperm relationships. Overall, the topology of Leguminosae that consistently emerges from analysis of molecular data is one in which the family is supported as monophyletic, although not always strongly so, as suggested by bootstrap and/or jackknife analyses. The traditionally defined (*sensu* Polhill *et al.*, 1981) subfamilies Mimosoideae and Papilionoideae have both been resolved as monophyletic, nested within a paraphyletic Caesalpinioideae, in all recent molecular studies (Käss and Wink, 1996; Doyle *et al.*, 1997, 2000; Kajita *et al.*, 2001). Subfamily Caesalpinioideae, with an estimated 161 genera and about 3,000 species (Lewis *et al.*, in press), comprise the basal elements in these phylogenies, although relationships among major groups were not always well supported. In all molecular studies to date, tribe Cercideae (e.g., *Cercis* and *Bauhinia*) is consistently resolved and strongly supported as the basal-most clade in the family, which notably conflicts with hypotheses based on floral ontogenetic studies that suggest such a position is occupied by caesalpinoid legumes with symmetrical flowers (Tucker and Douglas, 1994) or cladistic analyses of morphological characters that suggest a more derived position for Cercideae (Chappill, 1995; J. Chappill, University of Western Australia, unpublished data). While the flowers of *Cercis* resemble the strongly zygomorphic flowers that are a hallmark of subfamily Papilionoideae, recent ontogenetic analyses of numerous caesalpinoid and papilionoid flowers by Tucker (2002, and references therein) conclude this resemblance is a superficial one (convergence?). Alternatively, this resemblance may represent similar developmental manifestations of an underlying, fundamentally bilaterally symmetrical floral structure that is common to all legumes. Furthermore, not only do the molecular and ontogenetic studies differ with respect to Cercideae as the basal-most legume group, these studies provide contradicting hypotheses regarding the relationship of the Mimosoideae and Papilionoideae to a paraphyletic Caesalpinioideae as well (Fig. 1; Bruneau *et al.*, 2001). Relationships within *Cercis* have been the subject of recent studies by Hao *et al.* (2001) and Davis *et al.* (2002).

The next diverging clades identified in order of branching on the *rbL* phylogeny include members of tribe Cassieae (*Petalostylis* and subtribe Dialiinae) followed by tribe Detarieae. The remainder of the family is divided into two major clades, one comprising the remaining Cassieae, tribe Caesalpinieae, and Mimosoideae, and the clade comprising all members of Papilionoideae. Mimosoideae is strongly supported (>90% bootstrap and parsimony jackknife; Kajita *et al.*, 2001) as monophyletic and sister to *Erythrophleum* of the *Dimorphandra* group of Caesalpinieae. Remarkably, a monophyletic Papilionoideae receives only modest support (57% bootstrap, 62% jackknife) in the *rbL* analyses (Kajita *et al.*, 2001). This suggests that papilionoids are only weakly differentiated molecularly from their caesalpinoid sister groups, no one of which is singularly well supported as the sister clade to papilionoids.

Recent studies utilising chloroplast *trnL* intron sequences have sampled a much greater number and diversity of caesalpinoids (Bruneau *et al.*, 2000, 2001; Herendeen *et al.*, 2003), and have helped to clarify and provide further support for some relationships among the basal elements and major subclades in this subfamily (Fig. 2). As in the *rbL* studies, their combined analyses with anatomical, morphological and molecular data (Herendeen *et al.*, 2003), place the Cercideae as the basal-most clade, sister to rest of the family, with Detarieae *s.l.* (*sensu* Bruneau *et al.*, 2000, 2001) as the next diverging lineage, followed by Dialiinae *s.l.* (including subtribe Labicheninae). In the *trnL* analyses, the relative position of these latter two groups are reversed relative to that found in the *rbL* studies of Doyle *et al.* (2000) and Kajita *et al.* (2001); Käss and Wink (1996) did not sample Dialiinae *s.l.* Initially proposed by Polhill (1994), Detarieae *s.l.*, comprised of Detarieae and Amherstieae, is strongly supported as monophyletic in *trnL* analyses, although neither is individually monophyletic (Bruneau *et al.*, 2000). Interestingly, the presence of vestured pits in vessel elements, a feature that characterises most genera of Leguminosae, are noticeably lacking in the Cercideae and Dialiinae *s.l.* as they are in almost all of the rosid families suggested as close relatives of the Leguminosae

(Herendeen, 2000). Thus, based on the *trnL* results, vestured pits would have evolved twice in the family, once in *Detarieae s.l.*, and once in the clade that includes most of the other *Caesalpinioideae* (i.e., rest of tribe *Cassieae*, tribe *Caesalpinieae*), the *Mimosoideae*, and *Papilionoideae* (“clade A”, Figs. 2, 6; Bruneau *et al.*, 2001). Of the four tribes recognised in the latest classification of the *Caesalpinioideae* (Bruneau *et al.*, 2000, 2001), only *Cercideae* is supported as monophyletic. Also, unlike the results from *rbcL* analyses (e.g., Kajita *et al.*, 2001) the *Dimorphandra* group of *Caesalpinieae*, a diverse assemblage of genera, many of which share characteristics with the *Mimosoideae* (specifically the *Mimoseae*), is paraphyletic with respect to the *Mimosoideae* in the *trnL* analyses. Monophyly of the *Papilionoideae* is well supported by the *trnL* data (99% bootstrap), but with no well supported sister group from among *caesalpinioideae* legumes.

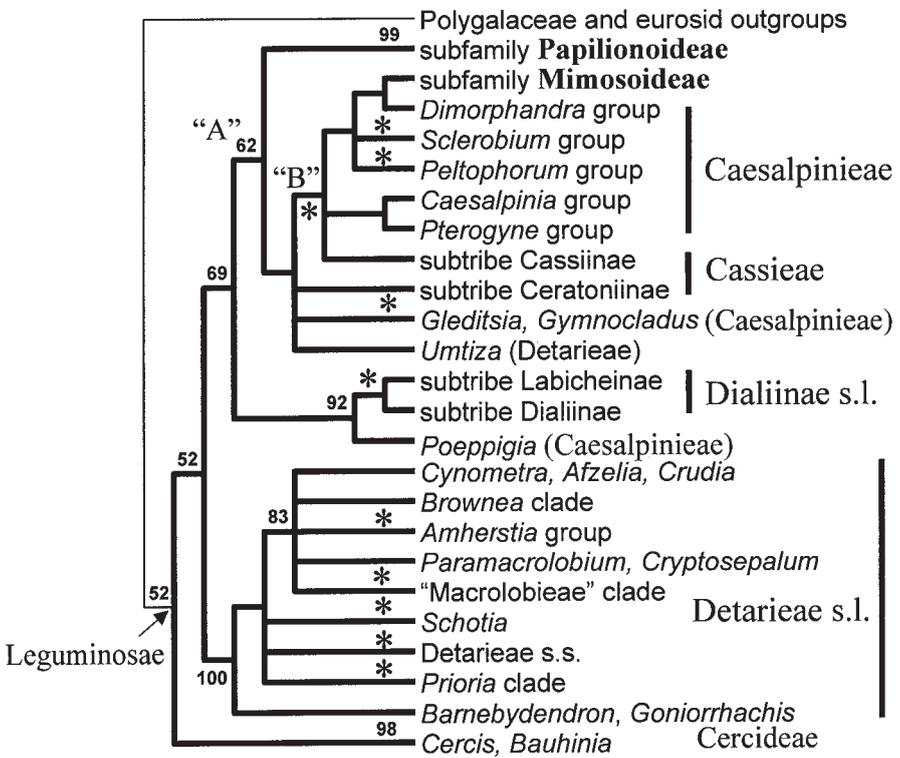


FIG. 2. Phylogenetic relationships at base of Leguminosae, based on parsimony analyses of *trnL* intron data (modified from Bruneau *et al.*, 2000, 2001). Some tribes and major clades indicated. Nodes labelled “A” and “B” refer to clades described in Bruneau *et al.* (2001, Fig. 6). Bootstrap support values indicated for major clades; *denotes clades supported by bootstrap levels greater than 50%. Support values for monophyletic Leguminosae based on analyses of other data sets: *rbcL*, 83% (jackknife; Kajita *et al.*, 2001); *matK*, 93% (bootstrap level; (M. Wojciechowski and M. Lavin, Montana State University, in preparation); *psaA*, 93% (bootstrap level, M. Wojciechowski and M. Sanderson, University of California, Davis, unpublished data).

Mimosoideae

The subfamily Mimosoideae, with an estimated 76 genera and some 3,000 species (Lewis *et al.*, in press), is the smallest of the subfamilies of legumes but probably the least understood from a phylogenetic perspective. While it has generally been accepted that the Mimosoideae was monophyletic, and derived from caesalpinoid ancestors (Polhill *et al.*, 1981; Chappill, 1995; Luckow *et al.*, 2000), the traditional “boundary” between the two subfamilies, with mimosoids distinguished by valvate aestivation of the petals and usually sepals, is not as distinct as once believed (e.g., Elias, 1981). Early analyses of molecular data (Käss and Wink, 1996; Doyle *et al.*, 1997) supported the monophyly of the subfamily, but these results were suspect because too few mimosoid taxa were included in these studies. More extensive sampling of both mimosoid and presumed closely related caesalpinoid taxa for *trnL* and *trnK/matK* sequences (Fig. 3), however, reveals there is no support for the monophyly of the Mimosoideae, with members of the *Dimorphandra* group of Caesalpinieae (*Dimorphandra*, *Erythrophleum*, *Mora*) nested within early diverging lineages of Mimoseae (*Dinizia*, *Piptadeniastrum*, *Pentaclethra*) or unresolved with respect to these and other caesalpinoid taxa (Luckow *et al.*, 2000; Bruneau *et al.*, 2001; Luckow *et al.*, 2003). The *Dimorphandra* group has long been hypothesised as an “apparent link” between these subfamilies (Elias, 1981). A close relationship between the *Dimorphandra* group and mimosoids is supported by several shared anatomical and morphological features such as bipinnate leaves, similar root nodules, and elongated spikes or paniculate inflorescences of small bisexual flowers (Polhill *et al.*, 1981; Chappill, 1995; Luckow *et al.*, 2000).

In addition to the suspect monophyly of the subfamily, the monophyly and relationships of its constituent tribes have been problematic as well (Luckow *et al.*, 2000, 2003). Within Mimosoideae, none of the traditionally recognised tribes, Acacieae, Ingeae, Mimoseae and Parkieae, are monophyletic (a fifth, Mimosygantheae, is monotypic) based on analyses of *trnL* and *trnK/matK* data (Luckow *et al.*, 2000, 2003); indeed, these authors suggest their continued recognition may not be feasible. Mimoseae form a basal grade, with Parkieae, Ingeae and Acacieae nested within it. For Parkieae, with *Parkia* more closely related to members of Ingeae and Mimoseae than to *Pentaclethra*, such a finding is not a surprise as several authors have suggested this on the basis of morphological evidence (e.g., Luckow and Hopkins, 1995). Only *Acacia* s.s. is strongly supported as monophyletic (Luckow *et al.*, 2003). Analyses of the nuclear histone H3-D region in addition to *trnK/matK* sequences (Miller and Bayer, 2000, 2001) are consistent with those of Luckow *et al.* (2000, 2003) in suggesting that neither the tribe Acacieae nor the larger of its two constituent genera, *Acacia* (c. 1,200 spp.), are monophyletic, the latter of which is also supported by RFLP analyses of chloroplast DNA (Robinson and Harris, 2000). Furthermore, all analyses agree that the other genus of Acacieae, the monotypic *Faidherbia*, is more closely related to Ingeae, a relationship that has been demonstrated by recent cladistic analyses of morphological data (Chappill and Maslin, 1995; Grimes, 1999) as well.

Papilionoideae: monophyly and overview of major clades

The Papilionoideae, the largest (with an estimated 483 genera and 12,000 species; Lewis *et al.*, in press) and most widely distributed of the three traditionally recognised subfamilies of Leguminosae, is readily distinguished from the other subfamilies by vegetative, floral, and fruiting characters (Polhill, 1981a; Chappill, 1995) including floral development (e.g., Tucker, 1987, 2002; Tucker and Douglas, 1994). While the subfamily has been the subject of intensive taxonomic research over the past few decades, higher level phylogenetic relationships within papilionoids have remained unclear. This may be due, in part, to a poor understanding of relationships among, and perhaps an over-emphasis on, the genera in the presumed basal tribes Sophoreae, Swartzieae, and to a lesser extent Dalbergieae (Polhill, 1994; Pennington *et al.*, 2001). Papilionoideae has been consistently resolved as monophyletic in analyses of *rbcL* data (Doyle *et al.*, 1997, 2000; Kajita *et al.*, 2001; Käss and Wink, 1995,

1996), but support for this conclusion has been low. Similarly, the overall topology (e.g., *rbtL* strict consensus tree; Doyle *et al.*, 2000; Kajita *et al.*, 2001) of the subfamily is largely unresolved and weakly supported, although there is an emergent pattern involving the resolution of at least four major papilionoid clades. Each of these clades has been the subject of recent investigations; the composition and structure of each has been further delimited and refined through more extensive sampling, and by virtue of the use of more informative molecular data. These include the “dalbergioids” (“aeschynomeneoid group”, Doyle *et al.*, 2000; Kajita *et al.*, 2001), which comprises tribes Adesmieae, Aeschynomeneae, the *Dalbergia* group of Dalbergieae, and subtribe Bryinae of Desmodieae (Lavin *et al.*, 2001a); the “genistoids”, tribes Genisteae, Thermopsidae, Crotalariae, Lipariae, Podalyriae, and members of Sophoreae *s.s.* (Käss and Wink, 1997a; Crisp *et al.*, 2000); a clade known as “Millettieae, Phaseoleae and allies” (“Phaseoloids”, Doyle *et al.*, 2000;

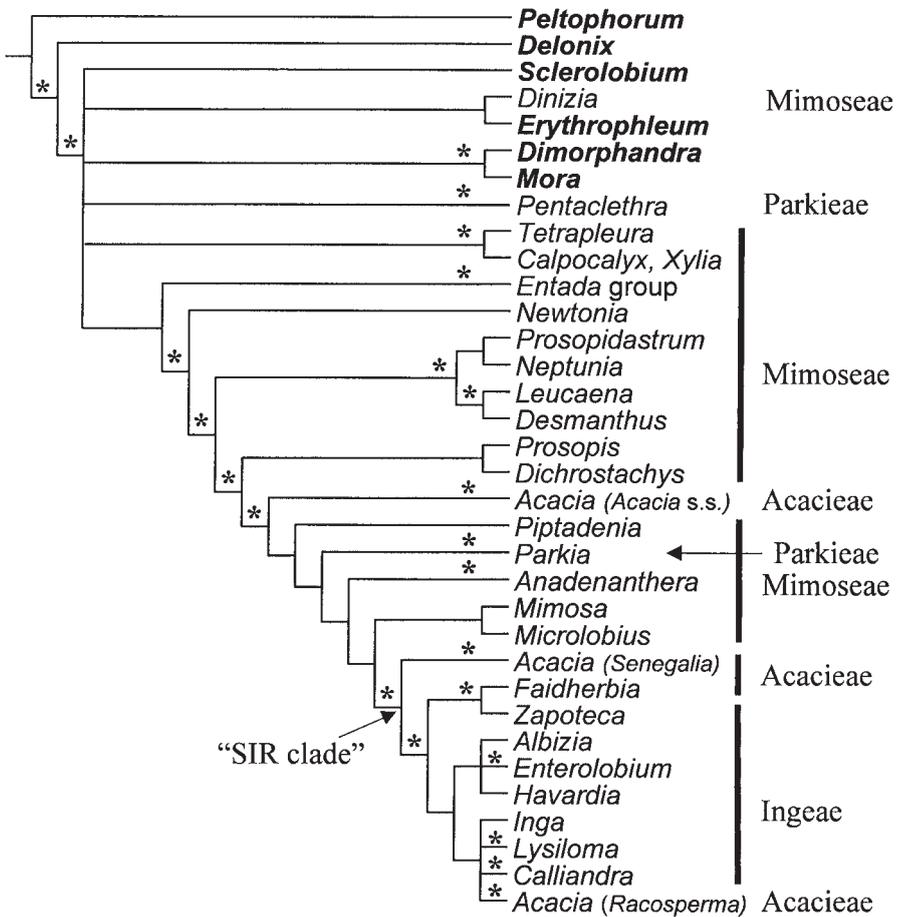


FIG. 3. Summary of phylogenetic relationships of mimosoids based on parsimony analyses of *trnL* intron and *trnK/matK* sequence data; *denotes clades supported by bootstrap levels greater than 50% (modified from Luckow *et al.*, 2000, 2003). Traditional Mimosoideae tribe names indicated to right, major groups of *Acacia* and genus *Faidherbia* (Acacieae) and “SIR” clade indicated; Caesalpinioideae genera names are shown in bold face.

“Millettoids/Phaseoloids”, Kajita *et al.*, 2001) that includes Millettieae, Desmodieae subtribes Desmodiinae and Lespedezinae, Abreae, Phaseoleae, and Psoraleae (Lavin *et al.*, 1998; Hu *et al.*, 2000; Kajita *et al.*, 2001); and “Hologalegina” (“galegoid clade”, Doyle *et al.*, 2000; Kajita *et al.*, 2001), a clade comprising the predominantly temperate herbaceous (epulvinate) tribes Galegeae, Carmichaelieae, Cicereae, Hedysareae, Trifolieae, Viciae, with Robinieae plus Loteae *s.l.* (including Coronilleae) forming a monophyletic subclade (Wojciechowski *et al.*, 2000). A potential fifth, large clade within papilionoids, composed primarily of the tribes Bossiaeeae and Mirbelieae, has been suggested by some studies (see below).

Cladistic analyses of *trnL* intron sequences (Ireland *et al.*, 2000; Pennington *et al.*, 2001) from those tribes long considered basal in the subfamily are also helping to clarify relationships among the earliest branching papilionoid groups. First, the *trnL* data (Fig. 4) strongly support the monophyly of the Papilionoideae, as currently circumscribed (Polhill, 1981a, 1994), thus clearly rejecting long standing suggestions based on morphological evidence that at least some genera of tribe Swartzieae might belong more appropriately in Caesalpinioideae (discussed in Polhill, 1994). Sophoreae, Swartzieae, and Dalbergieae (*sensu* Polhill, 1994) are all polyphyletic and nested at or near the base of the papilionoid clade, where the bulk of their constituent genera are mixed in a series of monophyletic groups whose relationships are still poorly resolved. However, several clades of significance are apparent, some of which were identified previously in *rbL* analyses despite the more limited sampling in those earlier studies (e.g., Doyle *et al.*, 1997). A clade of *Swartzia*, *Bobgunnia*, *Ateleia*, *Cyathostegia*, and *Bocoa* (“swartziod” clade) is resolved as sister to all other Papilionoideae (Ireland *et al.*, 2000). Although this association was never hypothesised previously on the basis of morphological evidence, the members of this clade all appear to possess an unusual ring meristem in the flower (Pennington *et al.*, 2000) and all possess nodulating ability (Sprent, 2000, 2001), clearly distinguishing them from other basal papilionoid groups. The next diverging lineages in the papilionoid clade include some individual groupings of Sophoreae genera (e.g., *Angylocalyx* and *Xanthocercis*, *Cladrastis* and *Styphnolobium*, *Myrospermum* and *Myroxylon*, *Castanospermum* and *Alexa*), Swartzieae genera (*Aldina*, *Cordyla*, and *Mildbraediendendron*), and a small, monophyletic tribe Dipterygeae (*Dipteryx*, *Pterodon*, *Taralea*) all of which are reasonably well-supported (Pennington *et al.*, 2001). Interestingly, the vast majority of genera in these latter groups have been surveyed and found to be non-nodulating (Sprent, 2000, 2001).

These studies have also provided some support for the hypothesis that the presence of the 50 kb inversion in the large single copy region of the chloroplast genome (Doyle *et al.*, 1996a) is a uniquely derived synapomorphy for a monophyletic group that includes most papilionoid groups (Fig. 4). As Pennington *et al.* (2001) point out, while many of the taxa in this “50 kb inversion” clade, which includes the four major papilionoid clades described above, remain to be screened for the presence of this inversion, all the taxa that have been demonstrated so far to lack the inversion (i.e., the “swartziod” clade, tribe Dipterygeae plus the several genera of Sophoreae and Swartzieae mentioned above; Doyle *et al.*, 1996a) fall outside of this clade, with the exception of *Aldina* (Swartzieae) which appears to possess the inversion (Doyle *et al.*, 1996a). Within the “50 kb inversion” clade several genera are known to lack nodulation activity, including *Exostyles*, *Harleyodendron*, *Lecointea*, *Luetzelburgia*, *Vatairea*, and *Zollernia* (Sprent 2000, 2001). This suggests that the ability to nodulate has been lost independently in different groups of woody papilionoids, both outside and within the “50 kb inversion” clade. Also among the major questions that remain to be resolved is the exact relationship among the major papilionoid clades, which will now be discussed in turn. The evidence obtained so far consistently support the Millettieae, Phaseoleae and allies clade as the sister group to Hologalegina, and suggests that together this larger clade occupies a more derived position in the papilionoid phylogeny than the other major clades.

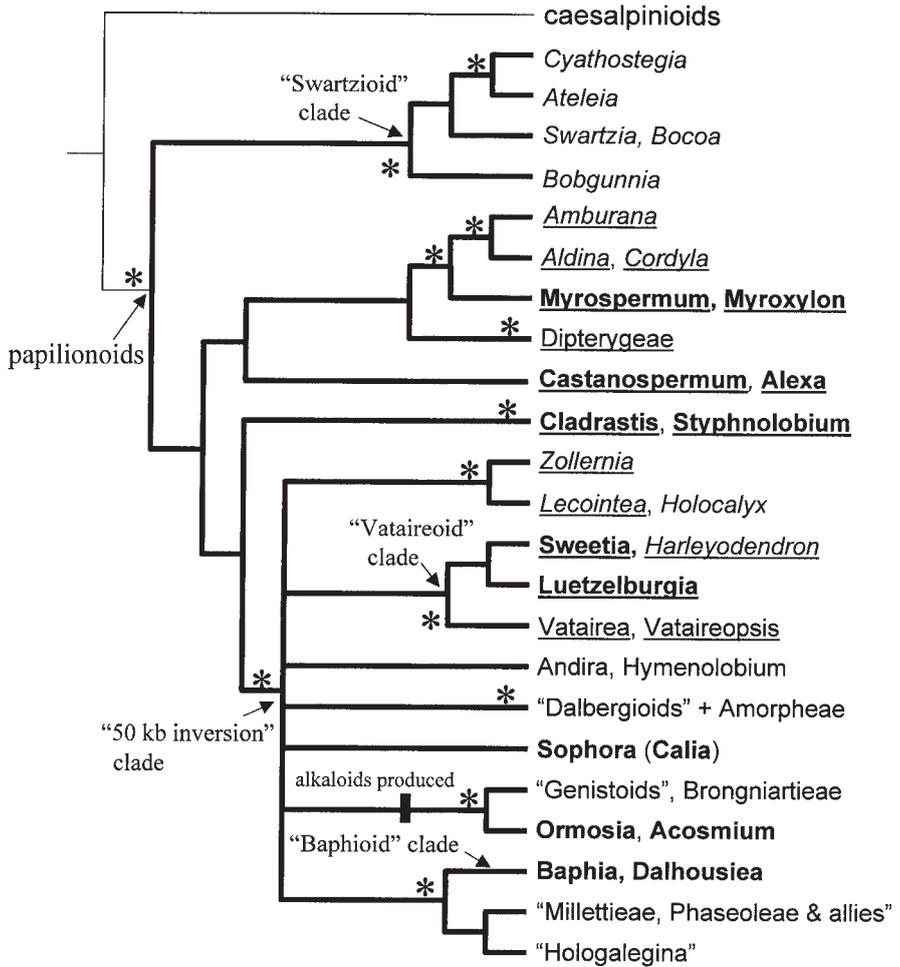


FIG. 4. Summary of phylogenetic relationships of basal papilionoid groups (bold lines), based on parsimony analyses of *trnL* sequence data (modified from Ireland *et al.*, 2000; Pennington *et al.*, 2000, 2001). Major clades identified. Names of Swartzieae genera shown in italics; Sophoreae genera in boldface. Genera with confirmed reports of negative nodulation ability are underlined (Sprent, 2000, 2001). *Denotes clades from *trnL* analyses with bootstrap support values greater than 50%. Bootstrap support for papilionoid clade based on analyses of *trnL* intron and other data sets: *trnL* data, 100% (Pennington *et al.*, 2001); *rbcl* data, 68% (Kajita *et al.*, 2001); *matK* data, 100% (M. Wojciechowski and M. Lavin, Montana State University, in preparation); *psaA* data, 99% (M. Wojciechowski and M. Sanderson, University of California, Davis, unpublished data).

Dalbergioid clade

The “dalbergioids” (Fig. 5) represent a newly described, monophyletic group comprising all genera previously referred to the tribes Adesmieae, Aeschynomeneae, subtribe *Bryinae* of Desmodieae, and Dalbergieae except the genera *Andira*, *Hymenolobium*, *Vatairea*, and *Vataireopsis* (Lavin *et al.*, 2001a). This unrecognised group (in previous classifications), as well as its three, well-supported constituent subclades, the *Adesmia* clade, *Dalbergia* clade, and *Pterocarpus* clade, were identified primarily on the basis of molecular data (chloroplast *trnK/matK* and *trnL* intron sequences) but barely distinguished, if at all, by non-molecular data (Lavin *et al.*, 2001a). The recognition of this pantropical group of 44 genera and c. 1,100 species clearly distinguishes the dalbergioids from the papilionoid groups that accumulate non-

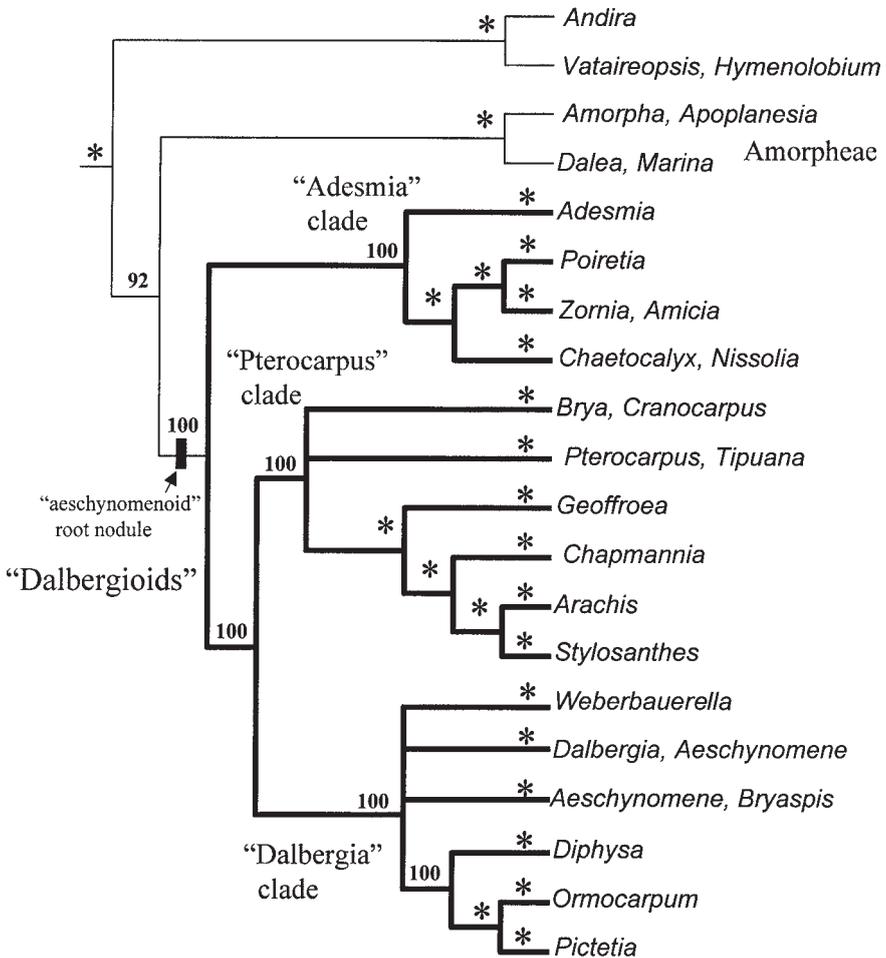


FIG. 5. Summary of phylogenetic relationships of the “dalbergioids” legumes (bold lines), based on combined parsimony analyses of *trnK/matK* and non-molecular data (modified from Lavin *et al.*, 2001a). Bootstrap support levels for major clades indicated; *denotes bootstrap support values for additional clades greater than 50%. Names of major clades and tribe Amorpheae indicated.

protein amino acids in their seeds, such as Millettieae and Hologalegina. The latter two groups have traditionally been considered closely related to and taxonomically confused with various members of the dalbergioids (Lavin *et al.*, 2001a). While the dalbergioids are apomorphically diagnosed by the aeschynomenoid root nodule (Sprent, 2001) and many molecular characters (Lavin *et al.*, 2001a), the shared inflorescence and fruiting similarity (e.g., Sousa and de Sousa, 1981) of certain dalbergioids and Millettieae is now attributed to independent evolution.

Interestingly, the analyses of Lavin *et al.* (2001a) indicate that members of tribe Amorpheae comprise the sister group to the dalbergioids (Fig. 5). Amorpheae are a group of eight genera and approximately 240 species confined primarily to the deserts and warm, arid regions of temperate and tropical North America, although *Dalea* extends from Canada to Argentina (Barneby, 1977, 1981). Preliminary results based on an analysis of *trnK*/*matK* sequences from all genera suggests Amorpheae comprises a monophyletic group (McMahon and Hufford, in press). Furthermore, both of the above conclusions are strongly supported in a larger *matK* analysis of more than 250 taxa as well (M. Wojciechowski and M. Lavin, Montana State University, in preparation).

Genistoid clade

The concept of the “genistoid alliance” (‘genistoids’) *sensu* Polhill (1981a) brought together for the first time a group of putatively related (albeit paraphyletic), predominantly southern hemisphere tribes that traditionally have been considered relatively isolated at the base of papilionoid legumes. The alliance as recognised consisted of four separate lineages; one includes the predominantly northern hemisphere Genisteae *s.s.*, Euchresteae, and Thermopsidae together with certain Sophoreae (*Sophora* group); a second comprising the mainly southern African Crotalariaeae, Lipariaeae, and Podalyrieae, and segregate tribe Hypocalypteae; a third comprising the endemic Australian Bossiaeeae and Mirbelieae; and the fourth, the neotropical Brongniartieae, variously linked to several of the other tribes but considered closest to the Australian lineage. Early molecular phylogenetic analyses by Käss and Wink (1996, 1997a, b) provided additional evidence that members of the alliance formed a monophyletic group with some members of Sophoreae (*Maackia* and *Sophora* spp.) near the base of papilionoids, while the Doyle *et al.* (1997) analysis indicates the genistoids are polyphyletic, forming three clades, the largest of which approximates the genistoid clade of Käss and Wink.

Using separate and combined nrDNA ITS and *rbcL* datasets, Crisp *et al.* (2000) have confirmed the polyphyly of the genistoids *sensu* Polhill (1981a). The authors suggest that this name be restricted to a well-supported “core genistoids” group (Fig. 6), centred in Africa and Eurasia, that comprises the majority of the tribes that made up Polhill’s genistoid alliance, including Genisteae, Crotalariaeae, Podalyrieae *s.l.* (plus Lipariaeae), Euchresteae, Thermopsidae, and members of the *Sophora* group of Sophoreae (some, but not all, species of *Sophora* and *Maackia*). This clade contains a large temperate and subtropical component with genera such as *Aspalathus* (250 spp.), *Crotalaria* (c. 600 spp.), *Lupinus* (c. 250 species), *Genista* (90 spp.), *Podalyria* (25 spp.), and *Thermopsis* (25 spp.). Additional molecular data (described below), supported by evidence from the distribution of quinolizidine alkaloids (e.g., Kinghorn and Balandrin, 1984), suggests the core genistoids clade is nested in a larger clade that includes several, largely Neotropical genera of Sophoreae such as *Acosmium*, *Bolusanthus*, *Bowdichia*, *Clathrotropis*, *Dicraeopetalum*, *Ormosia*, *Platycecyphium*, and *Salweenia*, tribe Brongniartieae, plus *Cyclobium* and *Poecilanth*e of Millettieae.

The exact composition and relationships of tribe Brongniartieae are still unresolved. *Poecilanth*e and *Cyclobium* are now recognised as belonging in a well-supported clade with Brongniartieae (*sensu* Crisp and Weston, 1987) based on nrDNA ITS (Crisp *et al.*, 2000; Hu *et al.*, 2002), *matK* (Hu *et al.*, 2000; M. Wojciechowski and M. Lavin, Montana State University, in preparation), *rbcL* (Kajita *et al.*, 2001), *trnL*

(Pennington *et al.*, 2001), and morphology (Doyle *et al.*, 2000). Unfortunately, since the most recent studies have not included representatives of all genera currently placed in the tribe plus *Poecilanthe* and *Cyclolobium* (Thompson *et al.*, 2001; Hu *et al.*, 2002), the relationships among these taxa remain unclear. Analyses of *matK* sequences (Hu *et al.*, 2000; M. Wojciechowski and M. Lavin, Montana State University, in preparation) suggest *Ormosia*, *Bolusanthus*, and *Acosmium* (Sophoreae) may be basally branching lineages in a larger “core genistoids+Brongniartieae+ Sophoreae” clade, while *trnL* analyses (Pennington *et al.*, 2001) indicate *Bowdichia* and *Clathrotropis* as potentially basal as well. However, sampling in those latter three studies was limited and may not reflect the correct placement of these taxa relative to

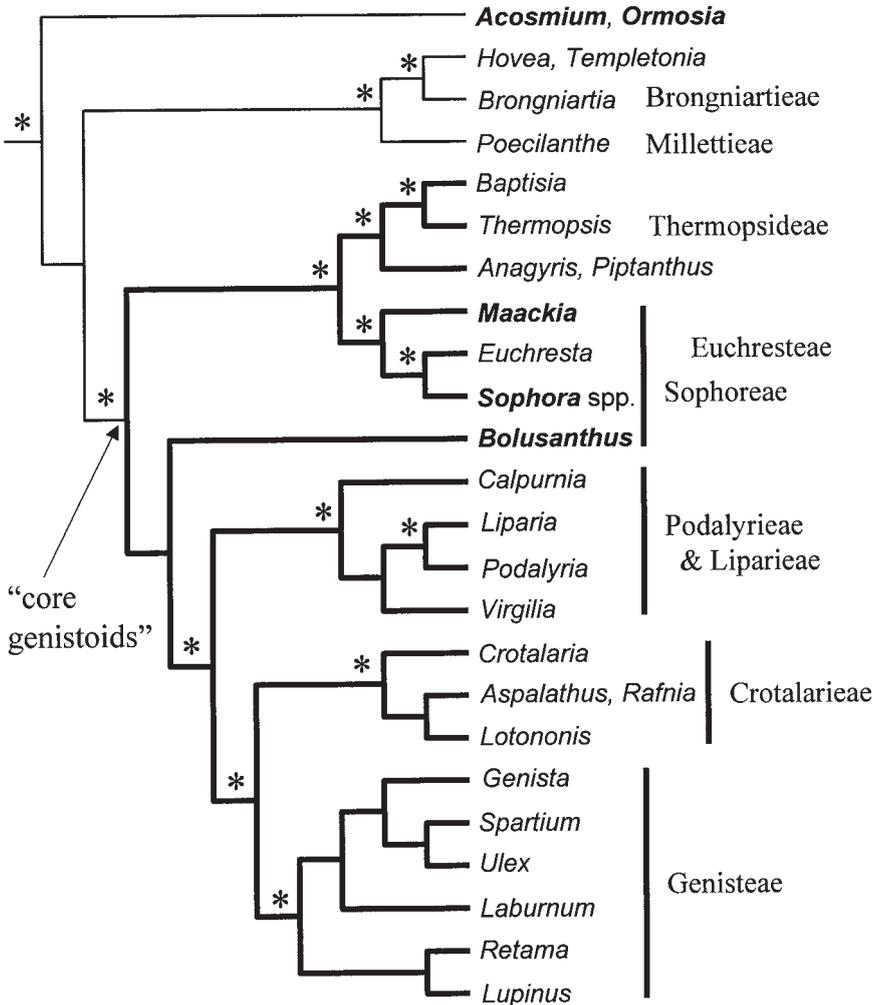


FIG. 6. Summary of phylogenetic relationships of “core genistoids” clade (bold lines), based on results from nrDNA ITS & *rbtL* (Crisp *et al.*, 2000), *matK* (Hu *et al.*, 2000; M. Wojciechowski and M. Lavin, Montana State University, in preparation), and *trnL* intron (Pennington *et al.*, 2001) analyses. *Denotes clades with bootstrap support levels greater than 50% (based mainly on Crisp *et al.*, 2000). Tribe names indicated to right; names of Sophoreae genera shown in boldface.

others in this clade. Furthermore, results from these studies with regard to the relationship of the Brongniartieae clade to that of the core genistoids, are largely equivocal; Crisp *et al.* (2000) and Doyle *et al.* (1997) find the Brongniartieae clade paraphyletic with respect to the core genistoids, while the most recent *rbL* analyses (Doyle *et al.*, 2000; Kajita *et al.*, 2001) essentially place Brongniartieae and core genistoids as unresolved clades at the base of papilionoids.

Several of these same studies have also consistently shown that the Australian tribes Bossiaeeae and Mirbelieae are not closely related to the core genistoids clade. Whether these tribes together form a monophyletic group, as suggested by Crisp and Weston (1987) and Crisp *et al.* (2000), is still unclear. However, evidence from *matK* sequences (M. Wojciechowski and M. Lavin, Montana State University, in preparation) suggests they form a monophyletic group with *Hypocalyptus* (Hypocalyptae; Schutte and van Wyk, 1998) as sister taxon, and that this clade, perhaps along with *Baphia* and close relatives in Sophoreae (Pennington *et al.*, 2001), comprise the immediate sister group to the larger clade that contains Millettieae, Phaseoleae, and their allies (Hu *et al.*, 2000; Kajita *et al.*, 2001) plus Hologalegina (Wojciechowski *et al.*, 2000).

Millettieae, Phaseoleae, and allies clade

The predominantly tropical tribe Millettieae, with over 40 genera and nearly 1,000 species, has figured prominently in hypotheses of the relationships of the “advanced” papilionoids — considered the “transitional link” from the more basal elements of Dalbergieae and Sophoreae, to putatively more derived tribes like Phaseoleae and Galegeae (Geesink, 1981, 1984; Polhill, 1981a). Early *rbL* analyses (Doyle *et al.*, 1997) suggested the polyphyly of Millettieae and Phaseoleae with their members falling into two main clades, one group dominated by Millettieae, the other comprising the bulk of Phaseoleae. In this and all subsequent analyses, Indigofereae emerges as the sister group to the larger clade containing most Millettieae, all of Phaseoleae, and including tribes Abreae, Desmodieae and Psoraleae.

More thorough molecular phylogenetic studies using nuclear phytochrome gene sequences by Lavin *et al.* (1998), *trnK/matK* sequences by Hu *et al.* (2000), and a combined analyses of *rbL*, *matK*, and nuclear ribosomal ITS data by Hu (2000), convincingly confirm a polyphyletic Millettieae (Fig. 7), with most of its taxa placed in two well-supported clades; one referred to as “core Millettieae” comprising the large genera *Millettia*, *Lonchocarpus*, *Derris* and *Tephrosia*, the other dominated by the majority of Phaseoleae. A third group of Millettieae comprises *Callerya*, *Wisteria*, and *Afgekia*, which are nested with *Glycyrrhiza* (Galegeae) at the base of the large clade comprising all temperate herbaceous legumes that lack one copy of the 25 kilobase inverted repeat in chloroplast DNA (IR-lacking clade or “IRLC” within Hologalegina, see below; Wojciechowski *et al.*, 2000). A fourth group of Millettieae taxa that includes *Cyclolobium* and *Poecilanthe* is distantly related to all other Millettieae and is nested within a clade containing taxa from Brongniartieae and Sophoreae amongst the basal lineages of papilionoids (Fig. 6). Relationships within each of these disparate groups of Millettieae taxa have been further resolved by more extensive sampling and analysis of nrDNA ITS sequences (Hu *et al.*, 2002).

The latest *rbL* analyses have demonstrated that Phaseoleae, the largest of the legume tribes in terms of number of genera (Lackey, 1977, 1981), as well as most of its larger subtribes, including Erythrinae, Glycininae and Phaseolinae are polyphyletic (Kajita *et al.*, 2001). The main Phaseoleae clade (Fig. 7), comprising subtribes Cajaninae, Erythrinae (at least the majority), Glycininae, Kennediinae and Phaseolinae, contains several agriculturally important genera such as *Glycine*, *Phaseolus* and *Vigna*. Tribes Desmodieae and Psoraleae are nested within this main Phaseoleae clade, but sampling to date from these latter tribes is too minimal to assess their monophyly. Phaseoleae subtribes Diocleinae and Ophrestinae appear to be part of the sister group to core Millettieae, while subtribe Clitoriinae groups with

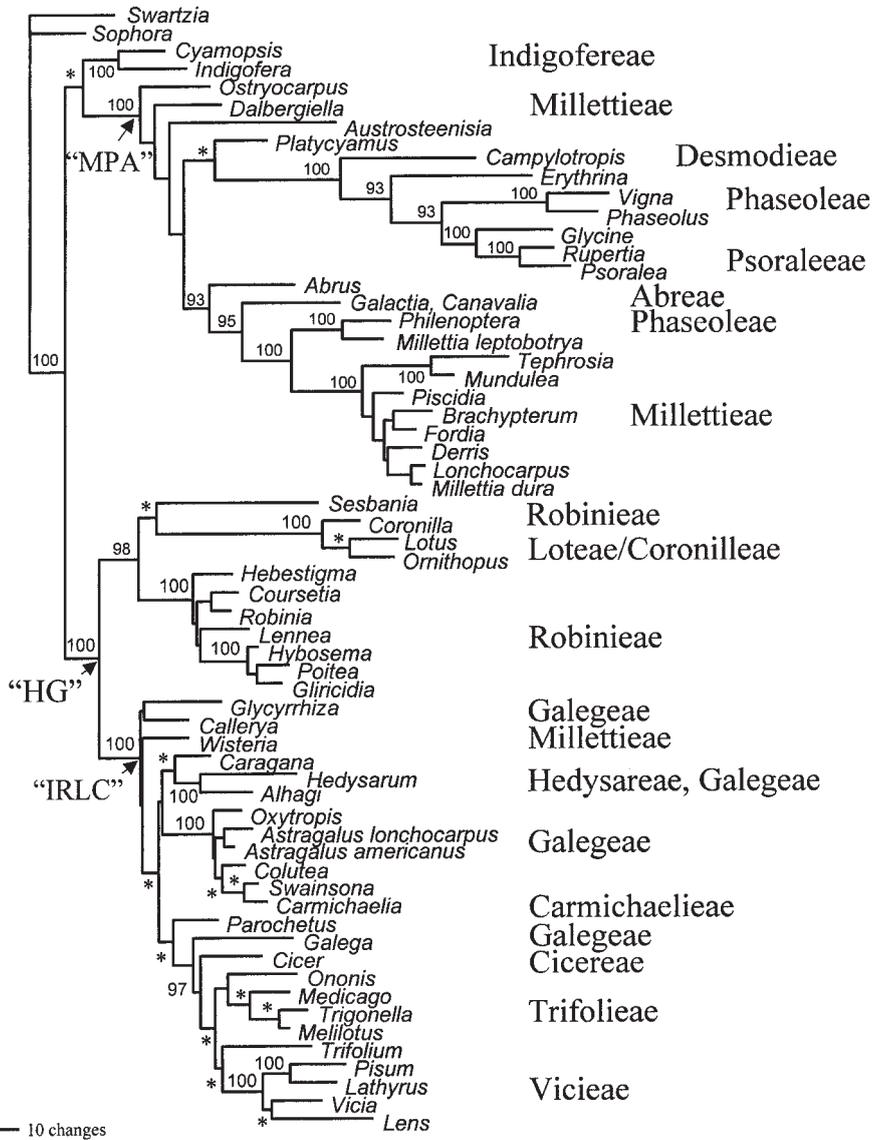


FIG. 7. Summary of phylogenetic relationships of “Millettieae, Phaseoleae & allies” (MPA) and “Hologalegina” (HG) clades. Tree shown is a phylogram based on maximum parsimony analyses of *matK* sequences (Hu *et al.*, 2000; M. Wojciechowski and M. Lavin, Montana State University, in preparation). “IRLC”, IR-lacking clade indicated. Bootstrap support levels shown above branches; * denotes clades with bootstrap levels greater than 70%. Note paraphyly of many of the indicated tribes.

Austrosteenisia (Millettieae) in a clade that appears sister to the main Phaseoleae clade (Kajita *et al.*, 2001). The results of Hu *et al.* (2000, 2002), are highly congruent with these latest findings based on *rbcL* (Kajita *et al.*, 2001) with regard to relationships in these two large clades containing the majority of Millettieae and all of Phaseoleae. Studies in progress on Desmodieae (T. Kajita, University of Tokyo, unpublished data), and Phaseoleae (M. Lavin, Montana State University and A. Delgado, UNAM, unpublished data) should help to further clarify phylogenetic relationships in the main Phaseoleae subclade of this major papilionoid radiation. Although Lee and Hymowitz (2001) resolve a monophyletic subtribe Glycininae, this may be an artifact of very limited taxon sampling of other Phaseoleae subtribes in their analysis.

Hologalegina clade

The largest of the major papilionoid clades is a group of largely temperate and herbaceous tribes (Carmichaelieae, Cicereae, Hedysareae, Trifolieae, Viciae and Loteae *s.l.*, i.e., including Coronilleae) plus taxa grouped by earlier authors (e.g. Polhill, 1981a) in the large, heterogeneous tribe Galegeae. This “temperate herbaceous group” has long been considered a natural group, characterised by a combination of morphological features shared by most of its members, with centres of diversity in Eurasia and the New World and likely derived from tropical ancestors (Polhill, 1981a). A number of earlier molecular studies (Lavin *et al.*, 1990; Liston, 1995; Liston and Wheeler, 1994; Sanderson and Wojciechowski, 1996; Doyle *et al.*, 1997; Käss and Wink, 1997a) lent further, though not always strong, support to the hypothesis that Galegeae is paraphyletic, but that together these tribes comprise a monophyletic lineage separate from other large temperate, herbaceous legume groups (e.g., core genistoids). The results from these and other studies (e.g., Lavin, 1995) also suggested this clade includes tribe Robinieae (*Robinia*) and at least some genera of Millettieae. Subsequent studies based on nuclear phytochrome genes (*PHY*, Lavin *et al.*, 1998), *trnK/matK* sequences (Hu *et al.*, 2000), and “supertree” analyses (Wojciechowski *et al.*, 2000) strongly support the conclusion that the majority of members of this temperate, herbaceous group of tribes (except Loteae, Coronilleae) plus *Callerya*, *Wisteria* and *Afgekia* of Millettieae, form a monophyletic group (the IRLC) that is the sister group to a strongly supported clade comprised of Robinieae (*sensu* Lavin and Sousa, 1995) plus Loteae *s.l.* Together these two clades (Fig. 7) comprise the monophyletic group referred to as “Hologalegina” (Wojciechowski *et al.*, 2000).

The composition and structure of major subclades within Hologalegina have recently been reviewed by Wojciechowski *et al.* (2000) and so will not be discussed in great detail here. The largest, the IRLC, containing about 93% of the species diversity represented by Hologalegina (c. 4,600 spp.), comprises three large, well-supported subclades. The largest of these, the “Astragalean” clade, is dominated by the genus *Astragalus* (2,500–3,000 spp.) and other, primarily north temperate (e.g., *Oxytropis*) and Australasian (e.g., *Swainsona*) genera from Galegeae (Wojciechowski *et al.*, 1999). The second, the “vicioid” clade, consists of the tribes Cicereae, Trifolieae and Viciae, which contain a number of agriculturally important genera including *Cicer*, *Lathyrus*, *Lens*, *Medicago*, *Melilotus*, *Pisum*, *Trifolium* and *Vicia*, plus *Galega* of Galegeae. Recent analyses based on *matK* sequences (Steele and Wojciechowski, 2003) are helping to clarify generic level relationships in this subclade, showing strong support for a monophyletic *Trifolium* and monophyletic Viciae as sister groups nested within a paraphyletic Trifolieae. Furthermore, these results are consistent with those of Bena (2001), based on nrDNA ITS and ETS sequences, in supporting a monophyletic *Medicago* (including the “medicagoid” *Trigonella* species) as the sister group to *Trigonella* plus *Melilotus*. The third subclade appears to include members of the tribe Hedysareae, *Alhagi* (Galegeae), and *Caragana* plus its close relatives from Galegeae. In spite of rather extensive taxonomic and molecular sampling across these three subclades, the data are not consistent in resolving relationships among them.

Traditional tribes Loteae (*s.l.*, including Coronilleae) plus Robinieae comprise the well-supported sister group to the IRLC within Hologalegina based upon molecular data (Doyle *et al.*, 2000; Wojciechowski *et al.*, 2000), while non-molecular data place Loteae nested close to Trifolieae and Viciae within the IRLC and Robinieae much closer to Indigofereae and members of Millettieae and Phaseoleae (Chappill, 1995; J. Chappill, unpublished data). Recent studies have more closely examined relationships within these two tribes. Using phylogenetic results from comprehensive sampling of non-molecular data and sequences from nrDNA ITS, *trnK/matK*, and the *trnL* intron, Lavin *et al.* (in press) clearly distinguish the exclusively New World “robinoid” legumes (and distinct subclades) from Robinieae (*sensu* Lavin and Sousa, 1995) so as to exclude the pantropical genus *Sesbania*, which accumulating evidence indicates is more closely related to Loteae *s.l.* Allan and Porter (2000) and Allan *et al.* (2003) show that while Loteae *s.l.*, as well as several of its constituent genera, is monophyletic based on nrDNA ITS sequences, traditional Loteae is paraphyletic with respect to Coronilleae. Furthermore, the largest and most widespread genus in the tribe, *Lotus* (“*s.l.*”), is not monophyletic, with Mediterranean species of *Lotus* more closely related to *Hammatolobium* (Coronilleae) than to New World members of the genus.

The sequence-based studies cited in Wojciechowski *et al.* (2000) provide overwhelming independent evidence for the conclusion that the structural mutation that resulted in the loss of one copy of the inverted repeat in the chloroplast genome of members of the IRLC is a unique synapomorphy for this clade (Lavin *et al.*, 1990). These and other studies (Doyle *et al.*, 2000; Kajita *et al.*, 2001) have clearly placed Hologalegina as the sister group to the clade comprised of taxa that have been assigned to the tribes Millettieae, Phaseoleae, and their primarily Old World, tropical relatives Indigofereae, Abreae, Desmodieae and Psoraleae. This result differs from earlier suggestions that the taxa which comprise Hologalegina were “derived from” a grade consisting of Millettieae and these latter tribes (Polhill, 1981a; Geesink, 1984). The recognition of Hologalegina and the Millettieae, Phaseoleae and allies clade, like that of the dalbergioids clade, should dispel any lingering doubts about the higher level relationships of tribes such as Aeschynomeneae (Polhill, 1994), Robinieae (e.g., Sousa and de Sousa, 1981) and Millettieae (e.g., Geesink, 1984), or certain of their constituent genera such as *Millettia/Tephrosia*, *Derris/Lonchocarpus* or *Poecilanthus*, groups whose composition and/or taxonomic affinities have been especially confused and fluid in recent history. Furthermore, the recognition of this larger clade of temperate and tropical papilionoids (i.e., Hologalegina plus Millettieae, Phaseoleae and allies clade, and Indigofereae) brings together the many disparate groups characterised by the accumulation of non-protein amino acids in seed, suggesting this character evolved (with few instances of loss or reversal) in the most recent common ancestor of this clade rather than multiple, independent times in legumes (Hu *et al.*, 2000).

Unresolved Papilionoid groups

The phylogenetic position of a number of genera and small groups nested among the major clades of papilionoids remain poorly resolved (if at all) in the majority of molecular analyses cited here. Many of these taxa include genera of Sophoreae, Swartzieae and Dalbergieae that are not members of the clades containing major radiations, and that form groups whose relationships are not resolved within the 50 kb inversion clade. For example, one of these groups (the “vataireoid” clade; Fig. 4), comprising *Vatairea* and *Vataireopsis* of Dalbergieae, *Luetzelburgia* and *Sweetia* of Sophoreae, and *Harleyodendron* and *Exostyles* of Swartzieae, is one of the best supported groups in the *trnL* analyses of Pennington *et al.* (2001). Likewise, most members of the largely African *Baphia* group of Sophoreae (*Baphia*, *Airyrantha*, *Dalhousiea*, *Leucomphalos*) plus *Baphiopsis* (Swartzieae) form a monophyletic group (“baphioid” clade) that is nested close to the large clade comprising Hologalegina plus Millettieae, Phaseoleae and allies, based on *trnL* analyses (Pennington *et al.*, 2001), or with Bossiaeeae plus Mirbelieae as the sister group to this huge clade based

on *matK* analyses (M. Wojciechowski and M. Lavin, Montana State University, in preparation). Both results, however, are weakly supported and in need of more thorough investigation.

Like the case with Sophoreae, these studies clearly illustrate the polyphyletic nature of *Sophora* itself, a diverse genus long considered central to any understanding of basal papilionoid relationships (Polhill, 1981b). *Sophora bhutanica* H. Ohashi, *S. davidii* (Franch.) Pavol., *S. flavescens* Aiton, *S. jaubertii* Spach, *S. macrocarpa* Sm., *S. microphylla* Aiton, *S. velutina* Lindl. (Käss and Wink, 1995; Crisp *et al.*, 2000; Doyle *et al.*, 1997; Pennington *et al.*, 2001), and *S. nuttalliana* B. L. Turner (M. Wojciechowski, unpublished data) nest within the core genistoids clade (Fig. 6). *Sophora affinis* Torr. & A. Gray (as *Styphnolobium affine* (Torr. & A. Gray) Walp.) and *S. japonica* L. (as *Styphnolobium japonicum* (L.) Schott) are nested in a clade with *Cladrastis* (Doyle *et al.*, 1997, 2000; Käss and Wink, 1997a; Kajita *et al.*, 2001; Pennington *et al.*, 2001) (Fig. 4) and the Californian chaparral endemic *Pickeringia* (M. Wojciechowski, unpublished data) outside the 50 kb inversion clade, closer to the base of papilionoids. The placement of *S. arizonica* S. Watson as the sister taxon to *S. secundiflora* (Ortega) Lagerh. ex DC. (*Calia secundiflora* (Ortega) Yakovlev) based on *matK* analysis (M. Wojciechowski, unpublished data), suggests a third (potential) group of *Sophora* species (Fig. 4), unresolved near the base of the 50 kb inversion clade (Kajita *et al.*, 2001; Käss and Wink, 1996, 1997a; Pennington *et al.*, 2001). These latter two results lend support to Sousa and Rudd's (1993) re-instatement of both *Styphnolobium* and *Calia* as genera distinct from *Sophora*.

Diversification of the Leguminosae: estimation of divergence times of major clades

The Leguminosae has been thought by earlier authors to have originated before the end of the Cretaceous, with the earliest reports consisting of fossil pollen and wood of what resemble Caesalpinioideae from the Maastrichtian of Canada, Siberia and Colombia (e.g., Raven and Polhill, 1981). Unfortunately, much of this evidence is limited and cannot be unequivocally assigned to the family (Herendeen *et al.*, 1992). The megafossil record for legumes from the Tertiary, however, is abundant and diverse, and indicates that all three subfamilies were present by the early Eocene (Crepet and Taylor, 1985, 1986; Crepet and Herendeen, 1992; Herendeen *et al.*, 1992). By the middle Eocene numerous taxa representing all three subfamilies were present at a number of localities, such as the diverse assemblage of caesalpinoid, mimosoid and papilionoid flowers, fruits and leaflets that have been described from the Buchanan clay pit site in southeastern North America (Herendeen, 1992). Indeed, the fossil record of this region (Mississippi Embayment), and other regions, documents the occurrence of a large number of tribes, including Acaciae, Caesalpinieae, Cassieae, Dalbergieae, Detarieae, Ingeae, Mimoseae, Sophoreae, Swartzieae, later Robinieae, by the middle Eocene (Herendeen *et al.*, 1992). Thus, the presence and abundance of most major lineages of woody legumes (except for the Cercideae) by the middle to upper Eocene suggests that extensive diversification had taken place by this time (Herendeen, 1992; Herendeen *et al.*, 1992).

The use of molecular sequence data for making inferences about the ages of lineages and clade diversification is increasingly revealing insights into the evolution of a diversity of taxa, such as metazoans (Wray *et al.*, 1996), Hawaiian silverswords (Baldwin and Sanderson, 1998), pandemic HIV-1 viruses (Korber *et al.*, 2000), and angiosperms (Sanderson and Doyle, 2001), including some legume groups (Wojciechowski *et al.*, 1999; Lavin *et al.*, 2001b; Richardson *et al.*, 2001). While comparisons of relative rates of molecular evolution (nucleotide substitution) have provided abundant evidence that rates vary widely between lineages, such characterisations provide neither an estimate of absolute rate of evolution nor any indication of how absolute rates change through

time (Sanderson, 1997, 2002). Moreover, because absolute rates of nucleotide substitution and divergence times are inextricably linked, one cannot be estimated without the other (Sanderson, 2002). Several general approaches for estimating divergence times of lineages in the absence of rate constancy (i.e., no molecular clock) have been proposed. Some involve the identification and subsequent pruning of terminal taxa or internal nodes that appear to deviate significantly from a tree-wide “average” rate of substitution (so-called “linearized” trees; Takezaki *et al.*, 1995), while some utilise “local” molecular clocks in which subtrees within a phylogeny are assigned different rates but the rate within each subtree is constant (e.g., Rambaut and Bromham, 1998). Others invoke highly parametric methods that model rate variation among branches explicitly (e.g., Thorne *et al.*, 1998). Recently, Sanderson has developed both an entirely non-parametric method (“non-parametric rate smoothing”: Sanderson, 1997) and a semi-parametric approach (“penalised likelihood”: Sanderson, 2002) for estimating rates of molecular evolution and reconstructing divergence times in the absence of rate constancy.

As an example of these approaches, I present here the results of a preliminary analysis in which the ages (in million years, MA) of some major clades of legumes are estimated based on molecular sequence data. The data set used for these analyses consists of complete chloroplast *matK* gene sequences (c. 1530 base pairs) sampled from 150 species of legumes culled from a larger data set of sequences representing more than 250 legume taxa (M. Wojciechowski and M. Lavin, Montana State University, in preparation), predominantly papilionoids. For this analysis, branch lengths (numbers of substitutions) were initially estimated on a single maximum parsimony tree, using sequences from *Polygala californica* Nutt. and *Quillaja saponaria* Molina as outgroups to root the topology; these taxa were then excluded from subsequent analysis. Zero-length branches, if any, were collapsed to hard polytomies. Ages were estimated for each node in the phylogeny, using the earliest unequivocal fossil evidence for Leguminosae (59.9 Ma; P. Herendeen, George Washington University, personal communication) to fix the age at the root of the tree (i.e., base of Leguminosae clade). Three methods of evolutionary rate analyses were employed, as implemented in the program ‘r8s’ (available at <http://ginger.ucdavis.edu/r8s>), and include: (1) Langley-Fitch (LF; Langley and Fitch, 1974), a maximum likelihood estimation of a tree-wide or globally constant rate of substitution (“clock-like” model); (2) non-parametric rate smoothing (NPRS; Sanderson, 1997) which estimates rates and times via a least-squares smoothing criterion that penalises rapid rate changes, where differences in the local rates of substitution about a node (subtending and immediate descendant branches) are minimised over the entire phylogeny; and (3) penalised likelihood (PL; Sanderson, 2002) which essentially uses a model in which every lineage has a separate rate that is penalised accordingly as rates vary across a phylogeny — a “cross validation criterion” is then used to identify an optimal level of rate smoothing. Because of the size of the data set analysed, the cross validation procedure to find the optimal level of smoothing was not performed; rather, smoothing values corresponding to a range of values from clock-like to one that departs significantly from clock-like rates of substitution (based on estimates of extremes of rate variation observed in the data) were chosen. For all analyses, three independent age estimates were performed on the data. For some NPRS analyses, minimum-age constraints were imposed on specific nodes: the papilionoid clade (PAPL), based on fossil flowers from early Eocene, 50 Ma (Crepet and Herendeen, 1992); the Robinieae clade (ROBB, Lavin *et al.*, 2001b), based on *Robinia*-like woods from late Eocene to Early Oligocene, 35 Ma (Wheeler, 2001); the *Ormocarpum* clade (ORMO, Lavin *et al.*, 2001b) based on the Caribbean or GAARlandia vicariance event, Late Eocene to Early Oligocene, 35 Ma (Rosen, 1985; Iturralde-Vinent and MacPhee, 1999); and the *Cladrastis-Styphnolobium* clade (CLAD), based on fossils from the Claiborne flora, 38 Ma (Herendeen, 1992); IRLC, based on *Vicia*- and *Astragalean*-like fossils from the Florissant Formation of Colorado, Upper Oligocene, 29 Ma (MacGinitie, 1953). The results of these analyses are presented in Table 1.

By assuming the origin (“root”) of the Leguminosae clade to be Late Paleocene, or c. 60 Ma, all three methods of rates analysis (LF, NPRS, and PL) estimate the age of the major legume lineages to be well into the Tertiary (c. 25–40 Ma), and many Late Eocene to Oligocene, in agreement with the fossil record (Table 1). In unconstrained analyses, all methods estimate the age of the papilionoids’ diversification to be older than that of the clade comprising mimosoids and its closest caesalpinoid relatives (Fig. 2, clade “B”; Bruneau *et al.*, 2001). Interestingly, the Millettieae, Phaseoleae and allies and Hologalegina clades, which contain the largest diversifications of herbaceous legumes in subtropical and temperate regions, as well as the Australian Bossiaeeae+Mirbelieae clade, were consistently older than other large papilionoid groups. Notably, the estimated ages of most nodes (with exception of caesalpinoids B/mimosoids) based on the LF and NPRS methods are very similar, and close to the range of estimates from the PL analyses. Furthermore, the addition of fossil-based minimum age constraints on selected nodes (NPRS w/constraints, Table 1) consistently estimated ages for all nodes older than those estimated by LF, NPRS, and PL alone, even those nodes whose minimum ages were constrained (e.g., PAPL and IRLC) in a NPRS analysis, suggesting that the estimates based on unconstrained analyses were conservative, underestimating the ages of all major clades.

TABLE 1. Estimated divergence times (Ma) for selected clades of legumes based on Langley-Fitch (LF), Non-parametric Rate Smoothing (NPRS) and Penalised Likelihood (PL) analyses. Analyses performed on 150 taxon *matK* gene sequence data set (M. Wojciechowski and M. Lavin, Montana State University, in preparation) using ‘r8s’ program (Sanderson, 1997, 2002) as described in text. All analyses performed with age of Leguminosae (ROOT) set to 59.9 Ma. *Minimum-age constraints of internal nodes, in Ma: PAPL = 50.0; ROBB = 35.0; ORMO = 35.0; IRLC = 29.0; CLAD = 38.0; see text for description of specific clades and fossil-constrained nodes. **Range of smoothing parameter. Note: node identified as “caesalpinoids B/mimosoids” corresponds to clade “B” of Bruneau *et al.* (2001); Astragalean clade is defined in Sanderson and Wojciechowski (1996); all other clades are described in text.

Node	LF	NPRS	NPRS w/constraints*	PL (1000-10)**
caesalpinoids				
B/mimosoids	23.9	40.7	53.5	31.3 – 40.8
Papilionoids	43.6	48.0	54.0	45.9 - 50.0
Core genistoids	23.7	27.3	32.6	28.6 - 29.9
Dalbergioids	32.3	29.2	46.4	34.1 – 34.2
Bossiaeeae-Mirbelieae	37.2	35.9	45.1	38.7 - 39.8
Millettieae, Phaseoleae, & allies	39.4	37.0	46.4	40.5 – 41.6
Hologalegina	36.3	34.1	45.4	37.3 – 38.1
IRLC	25.1	26.9	39.6	25.1 – 28.6
Astragalean clade	8.7	10.1	15.6	8.8 – 10.0

What do these estimated ages, based on rates of molecular evolution, for most of the large clades within the family suggest about the diversification of legumes? First, the two largest lineages, one comprising the majority of caesalpinioids and all mimosoids, the other that includes all papilionoids, originated by the early Tertiary and were diverse by mid Eocene (c. 35–40 Ma). Second, there was an early, rapid diversification of the major papilionoid lineages in both tropical *and* temperate regions, lineages that are now represented by the four large clades (and some minor ones) that make up almost 50% of extant legume diversity (c. 8,000 spp.). Third, most of the large temperate radiations, represented by the IRLC and genistoids clade, began diversifying by the Late Oligocene. These results contrast with the long standing view of legume diversification which suggested an early origin and long period of diversification of the main woody groups in the tropics followed by a much more recent radiation of temperate, herbaceous groups (Polhill, 1981a; Axelrod, 1992; Judd *et al.*, 1994).

Summary

In the ten years since ILC-3 perhaps the most significant developments in legume biology have been a tremendous increase in the amount of molecular data being gathered and used for phylogeny reconstruction in the family, and the emergence of the fields of comparative genomics and plant developmental mechanisms, using genetic, genomic and cell biological approaches. As they have in other plant groups, comparative molecular phylogenetic analyses, often coupled with non-molecular characters and information from the fossil record, have matured into mainstream methods in legume systematics studies, resolving some long-standing problems but also producing a number of surprises. Indeed, all biological sciences, biomedicine and bioinformatics, are relying increasingly on phylogenetic trees to provide a common framework for comparative analyses, and the number of phylogenetic analyses is increasing exponentially in response to this demand (e.g., Pagel, 1997). The years immediately following ILC-3 saw the initial attempts at construction of phylogenetic hypotheses for the family that combined morphological or molecular data with methods of cladistic analysis. It seems evident that, like the tremendous progress seen in the past five years on the phylogeny of the major groups of green plants (“Deep Green” initiative, <http://ucjeps.berkeley.edu/bryolab/greenplantpage.html>), and the phylogeny of the grass family (Barker *et al.*, 2001), the legume systematics community is now poised to rigorously develop comparable and comprehensive ephylogenies for the family based upon analyses of multiple data sets from a multitude of data sources. Such explicit hypotheses will have an even more profound impact on many areas of legume biology, including development and morphology, the origin and evolution of rhizobial- and arbuscular mycorrhizal-legume symbioses, legume/animal interactions such as pollination mechanisms, gene/genome evolution, as well as studies of evolutionary diversification and phylogeography.

It is also abundantly clear that the key to continued progress in legume biology is to focus on explicitly integrative studies — for example, in systematics there has been considerable progress in combining molecular and morphological data with a phylogenetic approach (e.g., Bruneau, 1996; Delgado-Salinas *et al.*, 1999; Pennington *et al.*, 2000; Lavin *et al.*, 2001a), while the combination of molecular phylogenies with paleontological/geological evidence and biogeography is illuminating the evolutionary histories of diverse legume groups (e.g., Lavin *et al.*, 2000, 2001b; Richardson *et al.*, 2001). Increasingly, a number of structural morphologists and developmental biologists (such as the participants at ILC-4) are viewing their work in a phylogenetic context, and are beginning to see an outline of the evolution of plant developmental programs, and the impact of genetic and metabolic regulatory pathways on plant development and morphology (e.g., Bharathan *et al.*, 1999; Reiser

et al., 2000; Hofer *et al.*, 2001; Möller and Cronk, 2001). Likewise, in studies of legume-animal and legume-microbe interactions the powerful tools of molecular and developmental biology combined with phylogenetic approaches are increasingly being used to test specific evolutionary questions about nodulation (e.g., Doyle, 1998), or the origin, specificity, and dynamics of rhizobial symbioses (e.g., Wernegreen and Riley, 1999; Debelle *et al.*, 2001; Loh *et al.*, 2001; Moulin *et al.*, 2001) in both model systems and related wild species. The emergence of new model legume molecular-genetic systems, such as *Medicago truncatula* (Cook, 1999) and *Lotus japonicus* (Sato *et al.*, 2002), and several, large-scale collaborative genomic sequencing/mapping and proteomics projects, of both legume species (e.g., *Medicago truncatula*, Bell *et al.*, 2001; Journet *et al.*, 2001; *Glycine max*, Shoemaker *et al.*, 2002) and their bacterial symbionts (e.g., *Sinorhizobium meliloti*, Galibert *et al.*, 2001), are now producing a wealth of new information and a set of new tools to address fundamental questions pertaining to many areas of legume biology, such as the dynamics of co-evolutionary relationships and the structure, function, and evolution of legume genomes. In addition to cataloguing the number and diversity of genes and other sequences, genome sequencing projects (including ESTs and physical/genetic map databases) will provide some of the first detailed and complete data on gene families, genomic organisation, and evolutionary changes in model systems, as well as new sources of data to mine for phylogenetic inference. A glimpse into the current status of the many legume genomics and bioinformatics initiatives and the power of comparative approaches to the analysis of this rapidly expanding data was provided at the First International Conference on Legume Genetics and Genomics, held in June 2002 in the USA. Phylogenetic biology will make an important contribution to these efforts by providing the comparative framework to understand the functions of new and novel genes and other sequences in both model and non-model species (“phylogenomics”, Eisen, 1998), as well as illuminate the evolution of individual genes as well as whole genomes within these organisms. Thus, one key to a greater understanding of both the developmental-genetic biology and evolutionary history of legumes is to be found in the kind of continued international collaborative work that was exemplified by those gathered to attend the Fourth International Legume Conference in Canberra, Australia.

Acknowledgements

I would like to thank the following people for their various contributions to the work presented here: Anne Bruneau, Michael Crisp, Jeff J. Doyle, Patrick Herendeen, Jer-Ming Hu, Tadashi Kajita, Matt Lavin, Aaron Liston, Melissa Luckow, Michele McMahon, Joseph Miller, R. Toby Pennington, Michael Sanderson and Kelly Steele. I would especially like to thank Matt Lavin, Anne Bruneau, Bente Klitgaard, Gwilym Lewis, Steffi Ickert-Bond and an anonymous reviewer for their many helpful comments and suggestions on this manuscript. Lastly, I would like to dedicate this paper to the memory of Rupert C. Barneby (1911–2000), whose long and remarkably productive career leaves a rich legacy of floristic and systematic work on the Leguminosae.

Literature cited

- Ainouche, A.-K. and Bayer, R.J. (1999). Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. *American Journal of Botany* 86: 590–607.
- Allan, G.J. and Porter, J.M. (2000). Tribal delimitation and phylogenetic relationships of tribes Loteae and Coronilleae (Faboideae: Fabaceae) with special

- reference to *Lotus*: evidence from nuclear ribosomal ITS sequences. *American Journal of Botany* 87: 1871–1881.
- Allan, G.J., Zimmer, E.A., Wagner, W.L. and Sokoloff, D.D. (2003). Molecular phylogenetic analyses of the tribe Loteae (Leguminosae): implications for classification and biogeography. In: B.B. Klitgaard and A. Bruneau (editors). *Advances in Legume Systematics*, part 10, Higher Level Systematics, pp. 371–393. Royal Botanic Gardens, Kew.
- Angiosperm Phylogeny Group (APG). (1998). An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* 85: 531–553.
- Archambault, A. and Bruneau, A. (2001). How useful is the *LEAFY* gene for phylogeny reconstruction in the Caesalpinioideae? Abstract, Botany 2001, American Society of Plant Taxonomists, Albuquerque, New Mexico.
- Asmussen, C.B. and Liston, A. (1998). Chloroplast DNA characters, phylogeny, and classification of *Lathyrus* (Fabaceae). *American Journal of Botany* 85: 387–401.
- Axelrod, D.I. (1992). Climatic pulses, a major factor in legume evolution. In: P.S. Herendeen and D.L. Dilcher (editors). *Advances in Legume Systematics*, part 4, The Fossil Record, pp. 259–280. Royal Botanic Gardens, Kew.
- Baldwin, B.G. and Sanderson, M.J. (1998). Age and rate of diversification of the Hawaiian silversword alliance. *Proceedings of the National Academy of Sciences USA* 95: 9402–9406.
- Barker, N.P., Clark, L.G., Davis, J.I., Duvall, M.R., Guala, G.F., Hsiao, C., Kellogg, E.A., Linder, H.P., Mason-Gamer, R.J., Mathews, S.Y., Simmons, M.P., Soreng, R.J. and Spangler, R.E. (2001). Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88: 373–457.
- Barneby, R.C. (1977). Daleae Imagines. *Memoirs of the New York Botanical Garden* 27.
- Barneby, R.C. (1981). Amorpheae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 344–346. Royal Botanic Gardens, Kew.
- Bell, C.J., Dixon, R.A., Farmer, A.D., Flores, R., Inman, J., Gonzales, R.A., Harrison, M.J., Paiva, N.L., Scott, A.D., Weller, J.W. and May, G.D. (2001). The *Medicago* genome initiative: a model legume database. *Nucleic Acids Research* 29: 114–117.
- Bena, G., Jubier, M.F., Olivieri, I. and Lejeune, B. (1998). Ribosomal external and internal transcribed spacers: Combined use in the phylogenetic analysis of *Medicago* (Leguminosae). *Journal of Molecular Evolution* 46: 299–306.
- Bena, G. (2001). Molecular phylogeny supports the morphologically based taxonomic transfer of the “medicagoid” *Trigonella* species to the genus *Medicago* L. *Plant Systematics and Evolution* 229: 217–236.
- Bharathan, G., Janssen, B.-J., Kellogg, E.A. and Sinha, N. (1999). Phylogenetic relationships and evolution of the *KNOTTED* class of plant homeodomain proteins. *Molecular Biology and Evolution* 16: 553–563.
- Bowe, L. M., Coat, G. and dePamphilis, C.W. (2000). Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales’ closest relative are conifers. *Proceedings of the National Academy of Sciences USA* 97: 4092–4097.
- Brouat, C., Gielly, L. and McKey, D. (2001). Phylogenetic relationships in the genus *Leonardoxa* (Leguminosae: Caesalpinioideae) inferred from chloroplast *trnL* intron and *trnL-trnF* intergenic spacer sequences. *American Journal of Botany* 88: 143–149.
- Bruneau, A. (1996). Phylogenetic and biogeographical patterns in *Erythrina* (Leguminosae: Phaseoleae) as inferred from morphological and chloroplast DNA characters. *Systematic Botany* 21: 587–605.
- Bruneau, A., Breteler, F.J., Wieringa, J.J., Gervais, G.Y.F. and Forest, F. (2000). Phylogenetic relationships in tribes Macrolobieae and Detarieae as inferred from chloroplast *trnL* intron sequences. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 121–149. Royal Botanic Gardens, Kew.

- Bruneau, A., Forest, F., Herendeen, P.S., Klitgaard, B.B. and Lewis, G.P. (2001). Phylogenetic relationships in the Caesalpinioideae (Leguminosae) as inferred from chloroplast *trnL* intron sequences. *Systematic Botany* 26: 487–514.
- Chandler, G.T., Bayer, R.J. and Crisp, M.D. (2001). A molecular phylogeny of the endemic Australian genus *Gastrolobium* (Fabaceae: Mirbelieae) and allied genera using chloroplast and nuclear markers. *American Journal of Botany* 88: 1675–1687.
- Chappill, J.A. (1995). Cladistic analysis of the Leguminosae: the development of an explicit phylogenetic hypothesis. In: M.D. Crisp and J.J. Doyle (editors). *Advances in Legume Systematics*, part 7, Phylogeny, pp. 1–9. Royal Botanic Gardens, Kew.
- Chappill, J.A. and Maslin, B.R. (1995). A phylogenetic assessment of the tribe Acaciaeae. In: M.D. Crisp and J.J. Doyle (editors). *Advances in Legume Systematics*, part 7, Phylogeny, pp. 77–99. Royal Botanic Gardens, Kew.
- Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedren, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn, G.H., Jr., Graham, S.W., Barrett, S.C.H., Dayanandan, S. and Albert, V.A. (1993). Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbL*. *Annals of the Missouri Botanical Garden* 80: 528–580.
- Chaw, S.-M., Parkinson, C.L., Cheng, Y., Vincent, T.M. and Palmer, J.D. (2000). Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings of the National Academy of Sciences USA* 97: 4086–4091.
- Cook, D.R. (1999). *Medicago truncatula* — a model in the making! *Current Opinion in Plant Biology* 2: 301–304.
- Crepet, W.L. and Taylor, D.W. (1985). The diversification of the Leguminosae: first fossil evidence of the Mimosoideae and Papilionoideae. *Science* 228: 1087–1089.
- Crepet, W.L. and Taylor, D.W. (1986). Primitive mimosoid flowers from the Paleocene-Eocene and the systematic and evolutionary implications. *American Journal of Botany* 73: 548–563.
- Crepet, W.L. and Herendeen, P.S. (1992). Papilionoid flowers from the early Eocene of southeastern North America. In: P.S. Herendeen and D.L. Dilcher (editors). *Advances in Legume Systematics*, part 4, The Fossil Record, pp. 43–55. Royal Botanic Gardens, Kew.
- Crisp, M. D. and Doyle, J. L. (1995). *Advances in Legume Systematics*, part 7, Phylogeny. Royal Botanic Gardens, Kew.
- Crisp, M.D., Gilmore, S. and van Wyk, B-E. (2000). Molecular phylogeny of the genistoid tribes of papilionoid legumes. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 249–276. Royal Botanic Gardens, Kew.
- Crisp, M.D. and Weston, P.H. (1987). Cladistics and legume systematics, with an analysis of the Bossiaeeae, Brongniartieae and Mirbelieae. In: C.H. Stirton (editor). *Advances in Legume Systematics*, part 3, pp. 65–130. Royal Botanic Gardens, Kew.
- Davis, C.C., Fritsch, P.W., Li, J. and Donoghue, M.D. (2002). Phylogeny and biogeography of *Cercis* (Fabaceae): evidence from nuclear ribosomal ITS and chloroplast *ndhF* sequence data. *Systematic Botany* 27: 289–302.
- Debellé, F., Moulin, L., Mangin, B., Denarie, J. and Boivin, C. (2001). *nod* genes and Nod signals and the evolution of the rhizobium legume symbiosis. *Acta Biochimica Polonica* 48: 359–365.
- Delgado-Salinas, A., Turley, T., Richman, A. and Lavin, M. (1999). Phylogenetic analysis of cultivated and wild species of *Phaseolus* (Fabaceae). *Systematic Botany* 24: 438–460.

- Dickson, W.C. (1981). The evolutionary relationships of the Leguminosae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 35–54. Royal Botanic Gardens, Kew.
- Doi, K., Kaga, A., Tomooka, N. and Vaughan, D.A. (2002). Molecular phylogeny of genus *Vicia* subgenus *Ceratotropis* based on rDNA ITS and *atpB-rbdL* intergenic spacer of cpDNA sequences. *Genetica* 114: 129–145.
- Downie, S.R., Katz-Downie, D.S., Rogers, E.J., Zujewski, H.L. and Small, E. (1998). Multiple independent losses of the plastid *rpoC1* intron in *Medicago* (Fabaceae) as inferred from phylogenetic analyses of nuclear ribosomal DNA internal transcribed spacer sequences. *Canadian Journal of Botany* 76: 791–803.
- Doyle, J.J. (1995). DNA data and legume phylogeny: A progress report. In: M.D. Crisp and J.J. Doyle (editors). *Advances in Legume Systematics*, part 7, Phylogeny, pp. 11–30. Royal Botanic Gardens, Kew.
- Doyle, J.J. (1998). Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. *Trends in Plant Science* 3: 473–478.
- Doyle, J.J., Chappill, J.A., Bailey, C.D. and Kajita, T. (2000). Towards a comprehensive phylogeny of legumes: evidence from *rbdL* sequences and non-molecular data. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 1–20. Royal Botanic Gardens, Kew.
- Doyle, J.J., Doyle, J.L., Ballenger, J.A., Dickson, E.E., Kajita, T. and Ohashi, H. (1997). A phylogeny of the chloroplast gene *rbdL* in the Leguminosae: taxonomic correlations and insights into the evolution of nodulation. *American Journal of Botany* 84: 541–554.
- Doyle, J.J., Doyle, J.L., Ballenger, J.A. and Palmer, J.D. (1996a). 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.
- Doyle, J.J., Kanazin, V. and Shoemaker, R.C. (1996b). Phylogenetic utility of histone H3 intron sequences in the perennial relatives of soybean (*Glycine*: Leguminosae). *Molecular Phylogenetics and Evolution* 6: 438–447.
- Eisen, J. (1998). Phylogenomics: improving functional predictions for uncharacterised genes by evolutionary analysis. *Genome Research* 8: 163–167.
- Elias, T.S. (1981). Mimosoideae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 143–152. Royal Botanic Gardens, Kew.
- Ferguson, I.K. and Tucker, S.C. (1994). *Advances in Legume Systematics*, part 6, Structural Botany. Royal Botanic Gardens, Kew.
- Galibert, F., Finan, R.M., Long, S.R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R.W., Dréano, S., Federspiel, N.A., Fisher, R.F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R.W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M.C., Pohl, T.M., Portetelle, D., Purnelle, B., Ramsperger, W., Surzycki, R., Thébault, P., Vandenbol, M., Vorhölter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K-C. and Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293: 668–672.
- Gauthier, P., Lumaret, R. and Bedecarrats, A. (1997). Chloroplast DNA variation in the genus *Lotus* (Fabaceae) and further evidence regarding the maternal parentage of *Lotus corniculatus* L. *Theoretical and Applied Genetics* 95: 629–636.
- Geesink, R. (1981). Tephrosieae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 245–260. Royal Botanic Gardens, Kew.
- Geesink, R. (1984). *Scala Millettiearum*. Leiden University Press, Leiden.
- Gillies, A.C.M. and Abbott, R.J. (1998). Evaluation of random amplified polymorphic DNA for species identification and phylogenetic analysis in *Stylosanthes* (Fabaceae). *Plant Systematics and Evolution* 211: 201–216.

- Grimes, J. (1999). Inflorescence morphology, heterochrony, and phylogeny in the mimosoid tribes Ingeae and Acacieae (Leguminosae: Mimosoideae). *Botanical Review* 65: 317–347.
- Hao, G., Zhang, D.X., Guo, L.X., Zhang, M.Y., Deng, Y.F. and Wen, X.Y. (2001). A phylogenetic and biogeographic study of *Cercis* (Leguminosae). *Acta Botanica Sinica* 43: 1275–1278.
- Herendeen, P.S. (1992). The fossil history of the Leguminosae from the Eocene of southeastern North America. In: P.S. Herendeen and D. L. Dilcher (editors). *Advances in Legume Systematics, part 4, The Fossil Record*, pp. 85–160. Royal Botanic Gardens, Kew.
- Herendeen, P.S. (2000). Structural evolution in the Caesalpinioideae (Leguminosae). In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics, part 9*, pp. 45–64. Royal Botanic Gardens, Kew.
- Herendeen, P.S. and Bruneau, A. (2000). *Advances in Legume Systematics, part 9*. Royal Botanic Gardens, Kew.
- Herendeen, P.S., Bruneau, A. and Lewis, G.P. (2003). Phylogenetic relationships in caesalpinoid legumes: a preliminary analysis based on morphological and molecular data. In: B.B. Klitgaard and A. Bruneau (editors). *Advances in Legume Systematics, part 10, Higher Level Systematics*, pp. 37–62. Royal Botanic Gardens, Kew.
- Herendeen, P. S., Crepet, W.L. and Dilcher, D.L. (1992). The fossil history of the Leguminosae: phylogenetic and biogeographic implications. In: P.S. Herendeen and D.L. Dilcher (editors). *Advances in Legume Systematics, part 4, The Fossil Record*, pp. 303–316. Royal Botanic Gardens, Kew.
- Herendeen, P.S. and Dilcher, D.L. (1992). *Advances in Legume Systematics, part 4, The Fossil Record*. Royal Botanic Gardens, Kew.
- Hofer, J., Gourlay, C., Michael, A. and Ellis, T.H.N. (2001). Expression of a class I *knotted1*-like homeobox gene is down-regulated in pea compound leaf primordia. *Plant Molecular Biology* 45: 387–398.
- Hu, J.-M. (2000). Phylogenetic relationships of the tribe Millettieae and allies — the current status. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics, part 9*, pp. 299–310. Royal Botanic Gardens, Kew.
- Hu, J.-M., Lavin, M., Wojciechowski, M.F. and Sanderson, M.J. (2000). Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on *trnK/matK* sequences, and implications for evolutionary patterns in Papilionoideae. *American Journal of Botany* 87: 418–430.
- Hu, J.-M., Lavin, M., Wojciechowski, M.F. and Sanderson, M.J. (2002). Phylogenetic analysis of nuclear ribosomal ITS/5.8 S sequences in the tribe Millettieae (Fabaceae): *Poecilanthus-Cyclolobium*, the core Millettieae, and the *Callerya* group. *Systematic Botany* 27: 722–733.
- Hurr, K.A., Lockhart, P.J., Heenan, P.B. and Penny, D. (1999). Evidence for the recent dispersal of *Sophora* (Leguminosae) around the Southern Oceans: molecular data. *Journal of Biogeography* 26: 565–577.
- Ireland, H., Pennington, R.T. and Preston, J. (2000). Molecular systematics of the Swartzieae. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics, part 9*, pp. 217–231. Royal Botanic Gardens, Kew.
- Iturralde-Vinent, M.A. and MacPhee, R.D.E. (1999). Paleogeography of the Caribbean region: implications for Cenozoic Biogeography. *Bulletin of the American Museum of Natural History* 238: 1–95.
- Journet, E.P., Carreau, V., Gouzy, J., Thoquet, P., Rosenberg, C., Barker, D., Huguet, T., Denarie, J. and Gamas, P. (2001). The model legume *Medicago truncatula*: recent advances and perspectives in genomics. *Ocl-Oleagineux Corps Gras Lipides* 8: 478–484.
- Judd, W.S., Sanders, R.W. and Donoghue, M.J. (1994). Angiosperm family pairs: preliminary phylogenetic analyses. *Harvard Papers in Botany* 5: 1–51.

- Kajita, T., Ohashi, H., Tateishi, Y., Bailey, C.D. and Doyle, J.J. (2001). *rbcl* and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and Allies. *Systematic Botany* 26: 515–536.
- Källersjö, M., Farris, J.S., Chase, M.W., Bremer, B., Fay, M.F., Humphries, C. J., Petersen, G., Seberg, O. and Bremer, K. (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 Wink, M. (1995). Molecular phylogeny of the Papilionoideae (Family Leguminosae): *rbcl* gene sequences versus chemical taxonomy. *Botanica Acta* 108: 149–162.
- Käss, E. and Wink, M. (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 Wink, M. (1997a). 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.
- Käss, E. and Wink, M. (1997b). 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.
- Kinghorn, A.D. and Balandrin, M.F. (1984). Quinolizidine alkaloids in the Leguminosae: structural types, analysis, chemotaxonomy and biological activities. In: S.W. Pelletier (editor). *Alkaloids: chemical and biological perspectives*, pp. 105–148. John Wiley and Sons, New York.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S. and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* 288: 1789–1796.
- Lackey, J.A. (1977). A revised classification of the tribe Phaseoleae (Leguminosae, Papilionoideae), and its relation to canavanine distribution. *Botanical Journal of the Linnean Society* 74: 163–178.
- Lackey, J.A. (1981). Phaseoleae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 301–327. Royal Botanic Gardens, Kew.
- Langley, C.H. and Fitch, W. (1974). An estimation of the constancy of the rate of molecular evolution. *Journal of Molecular Evolution* 3: 161–177.
- Lavin, M. (1995). Tribe Robinieae and allies; model groups for assessing early Tertiary northern latitude diversification of tropical legumes. In: M.D. Crisp and J.J. Doyle (editors). *Advances in Legume Systematics*, part 7, Phylogeny, pp. 141–160. Royal Botanic Gardens, Kew.
- Lavin, M., Doyle, J.J. and Palmer, J.D. (1990). Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. *Evolution* 44: 390–402.
- Lavin, M., Eshbaugh, E., Hu, J.-M., Mathews, S. and Sharrock, R.A. (1998). Monophyletic subgroups of the tribe Millettieae (Leguminosae) as revealed by phytochrome nucleotide sequence data. *American Journal of Botany* 85: 412–433.
- Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., De Lima, H.C. and Gasson, P.E. (2001a). The dalbergioid legumes (Fabaceae): delimitation of a monophyletic pantropical clade. *American Journal of Botany* 88: 503–533.
- Lavin, M. and Sousa S., M. (1995). Phylogenetic systematics and biogeography of the tribe Robinieae. *Systematic Botany Monographs* 45: 1–165.
- Lavin, M., Thulin, M., Labat, J.-N. and Pennington, R.T. (2000). Africa, the odd man out: molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. *Systematic Botany* 25: 449–467.
- Lavin, M., Wojciechowski, M.F., Richman, A., Rotella, J., Sanderson, M.J. and Beyra Matos, A. (2001b). Identifying Tertiary radiations of Fabaceae in the Greater

- Antilles: alternatives to cladistic vicariance analysis. *International Journal of Plant Sciences* 162: S53–S76.
- Lavin, M., Wojciechowski, M.F., Gasson, P., Hughes, C. and Wheeler, E. (in press) Phylogeny of robinoid legumes (Fabaceae) revisited: *Coursetia* and *Glinicidia* recircumscribed, and a biogeographical appraisal of the Caribbean endemics. *Systematic Botany*.
- Lee, J. and Hymowitz, T. (2001). A molecular phylogenetic study of the subtribe Glycininae (Leguminosae) derived from the chloroplast DNA *rps16* intron sequences. *American Journal of Botany* 88: 2064–2073.
- Lewis, G., Schrire, B., Mackinder, B. and Lock, M. (editors) (in press). Legumes of the world. Royal Botanic Gardens, Kew.
- 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 (editors). *Advances in Legume Systematics*, part 7, Phylogeny, pp. 31–40. Royal Botanic Gardens, Kew.
- Liston, A. and Wheeler, J.A. (1994). The phylogenetic position of the genus *Astragalus* (Fabaceae): evidence from the chloroplast genes *rpoC1* and *rpoC2*. *Biochemical Systematics and Ecology* 22: 377–388.
- Loh, J.T., Yuen-Tsai, J.P.Y., Stacey, M.G., Lohar, D., Welborn, A. and Stacey, G. (2001). Population density-dependent regulation of the *Bradyrhizobium japonicum* nodulation genes. *Molecular Microbiology* 42: 37–46.
- Luckow, M. and Hopkins, H.C.F. (1995). A cladistic analysis of *Parkia* (Leguminosae: Mimosoideae). *American Journal of Botany* 82: 1300–1320.
- Luckow, M., Miller, J.T., Murphy, D.J. and Livshultz, T. (2003). A phylogenetic analysis of the Mimosoideae (Leguminosae) based on chloroplast DNA sequence data. In: B.B. Klitgaard and A. Bruneau (editors). *Advances in Legume Systematics*, part 10, Higher Level Systematics, pp. 197–220. Royal Botanic Gardens, Kew.
- Luckow, M., White, P.J. and Bruneau, A. (2000). Relationships among the basal genera of mimosoid legumes. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 165–180. Royal Botanic Gardens, Kew.
- Mabberley, D.J. (1997). *The Plant Book*. 2nd edition. Cambridge University Press, Cambridge.
- MacGinitie, H.D. (1953). Fossil plants of the Florissant beds, Colorado. *Carnegie Institute of Washington Publication* 599.
- McMahon, M. and Hufford, L. (in press). Phylogeny of Amorphaeae (Fabaceae: Papilionoideae). *American Journal of Botany*.
- Miller, J.T. and Bayer, R.J. (2000). Molecular phylogenetics of *Acacia* (Fabaceae: Mimosoideae) based on the chloroplast *trnK/matK* and nuclear Histone H3-D DNA sequences. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 181–200. Royal Botanic Gardens, Kew.
- Miller, J.T. and Bayer, R.J. (2001). Molecular phylogenetics of *Acacia* (Fabaceae: Mimosoideae) based on the chloroplast *matK* coding sequence and flanking *trnK* intron spacer regions. *American Journal of Botany* 88: 697–705.
- Möller, M. and Cronk, Q.C.B. (2001). Evolution of morphological novelty: a phylogenetic analysis of growth patterns in *Streptocarpus* (Gesneriaceae). *Evolution* 55: 918–929.
- Moulin, L., Munive, A., Dreyfus, B. and Boivin-Masson, C. (2001). Nodulation of legumes by members of the β -subclass of Proteobacteria. *Nature* 411: 948–950.
- Nandi, O.I., Chase, M.W. and Endress, P.K. (1998). A combined cladistic analysis of angiosperms using *rbcL* and non-molecular data sets. *Annals of the Missouri Botanical Garden* 85: 137–212.
- Nickrent, D.L., Parkinson, C.L., Palmer, J.D. and Duff, R.J. (2000). Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. *Molecular Biology and Evolution* 17: 1885–1895.

- Pagel, M. (1997). Inferring evolutionary processes from phylogenies. *Zoologica Scripta* 26: 331–348.
- Pennington, R.T., Klitgaard, B.B., Ireland, H. and Lavin, M. (2000). New insights into floral evolution of basal Papilionoideae from molecular phylogenies. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 233–248. Royal Botanic Gardens, Kew.
- Pennington, R.T., Lavin, M., Ireland, H., Klitgaard, B., Preston, J. and Hu, J.-M. (2001). Phylogenetic relationships of basal papilionoid legumes based upon sequence 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.
- Pickersgill, B. and Lock, J.M. (1996). *Advances in Legume Systematics*, part 8, Legumes of Economic Importance. Royal Botanic Gardens, Kew.
- Polhill, R.M. (1981a). Papilionoideae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 191–208. Royal Botanic Gardens, Kew.
- Polhill, R.M. (1981b). Sophoreae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 213–230. Royal Botanic Gardens, Kew.
- Polhill, R.M. (1994). Classification of the Leguminosae. In: F.A. Bisby, J. Buckingham and J.B. Harborne (editors). *Phytochemical dictionary of the Leguminosae*, pp. xxxv – xlviii. Chapman and Hall, New York, NY.
- Polhill, R.M., Raven, P.H. and Stirton, C.H. (1981). Evolution and systematics of the Leguminosae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 1–26. Royal Botanic Gardens, Kew.
- Pryer, K.M., Schneider, H., Smith, A.R., Cranfill, R., Wolf, P.G., Hunt, J.S. and Sipes, S.D. (2001). Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* 409: 618–622.
- Rambaut, A. and Bromham, L. (1998). Estimating divergence data from molecular sequences. *Molecular Biology and Evolution* 15: 442–448.
- Ramírez, L., de la Vega, A., Razkin, N., Luna, V. and Harris, P.J.C. (1999). Analysis of the relationship between species of the genus *Prosopis* revealed by the use of molecular markers. *Agronomie* 19: 31–43.
- Raven, P.H. and Polhill, R.M. (1981). Biogeography of the Leguminosae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics*, part 1, pp. 27–34. Royal Botanic Gardens, Kew.
- Reiser, L., Sánchez-Baracaldo, P. and Hake, S. (2000). Knots in the family tree: evolutionary relationships and functions of *knox* homeobox genes. *Plant Molecular Biology* 42: 151–166.
- Richardson, J.E., Pennington, R.T., Pennington, T.D. and Hollingsworth, P.M. (2001). Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* 293: 2242–2245.
- Robinson, J. and Harris, S.A. (2000). A plastid DNA phylogeny of the genus *Acacia* Miller (Acaciaeae, Leguminosae). *Botanical Journal of the Linnean Society* 132: 195–222.
- Rosen, D.E. (1985). Geological hierarchies and the biogeographic congruence in the Caribbean. *Annals of the Missouri Botanical Garden* 72: 636–659.
- Sanderson, M.J. (1997). A non-parametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* 14: 1218–1231.
- Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalised likelihood approach. *Molecular Biology and Evolution* 19: 101–109.
- Sanderson, M.J. and Doyle, J.A. (2001). Sources of error and confidence intervals in estimating the age of angiosperms from *rbdL* and 18S rDNA data. *American Journal of Botany* 88: 1499–1516.
- Sanderson, M.J. and Wojciechowski, M.F. (1996). Diversification rates in a temperate legume clade: Are there “so many species” of *Astragalus* (Fabaceae)? *American Journal of Botany* 83: 1488–1502.

- Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S. (2002). A large scale genome analysis of *Lotus japonicus* — I. Genome sequencing and mapping. *Plant and Cell Physiology* 43: S124.
- Savolainen, V., Chase, M.W., Hoot, S.B., Morton, C.M., Soltis, D.E., Bayer, C., Fay, M.F., de Bruijn, A.Y., Sullivan, S. and Qiu, Y.-L. (2000). Phylogenetics of flowering plants based on combined analysis of plastid *atpB* and *rbcl* gene sequences. *Systematic Biology* 49: 306–362.
- Schnabel, A. and Wendel, J.F. (1998). Cladistic biogeography of *Gleditsia* (Leguminosae) based on *ndhF* and *rpl16* chloroplast gene sequences. *American Journal of Botany* 85: 1753–1765.
- Schutte, A.L. and van Wyk, B.-E. (1998). The tribal position of *Hypocalyptus* Thunberg (Fabaceae). *Novon* 8: 178–182.
- Shoemaker, R., Keim, P., Vodkin, L., Retzel, E., Clifton, S.W., Waterston, R., Smoller, D., Coryell, V., Khanna, A., Erpelding, J., Gai, X.W., Brendel, V., Raph-Schmidt, C., Shoop, E.G., Vielweber, C.J., Schmatz, M., Pape, D., Bowers, Y., Theising, B., Martin, J., Dante, M., Wylie, T. and Granger, C. (2002). A compilation of soybean ESTs: generation and analysis. *Genome* 45: 329–338.
- Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C. and Farris, J.S. (2000). Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *Botanical Journal of the Linnean Society* 133: 381–461.
- Soltis, D.E., Soltis, P.S., Morgan, D.R., Swensen, S.M., Mullin, B.C., Dowd, J.M. and Martin, P.G. (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.
- Soltis, D.E., Soltis, P.S., Nickrent, D.L., Johnson, L.A., Hahn, W.J., Hoot, S.B., Sweere, J.A., Kuzoff, R.K., Kron, K.A., Chase, M.W., Swensen, S.M., Zimmer, E.A., Chaw, S.-M., Gillespie, L.J., Kress, W.J. and Sytsma, K.J. (1997). Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Annals of the Missouri Botanical Garden* 84: 1–49.
- Sousa, M. and de Sousa, M.P. (1981). New World Lonchocarpaceae. In: R.M. Polhill and P.H. Raven (editors). *Advances in Legume Systematics, part 1*, pp. 261–281. Royal Botanic Gardens, Kew.
- Sousa, M. and Rudd, V.E. (1993). Revisión del genero *Styphnolobium* (Leguminosae: Papilionoideae: Sophoreae). *Annals of the Missouri Botanical Garden* 80: 270–283.
- Sprent, J.I. (2000). Nodulation as a taxonomic tool. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics, part 9*, pp. 21–43. Royal Botanic Gardens, Kew.
- Sprent, J.I. (2001). Nodulation in Legumes. Royal Botanic Gardens, Kew.
- Sprent, J.I. and McKey, D. (1994). *Advances in Legume Systematics, part 5, The Nitrogen Factor*. Royal Botanic Gardens, Kew.
- Steele, K.P., Tizon, E., Evan, R.C., Campbell, C.S. and Wojciechowski, M.F. (2000a). Sister group relationships of Fabaceae and Rosaceae: phylogenetic relationships of Eurosids I. *American Journal of Botany* 87: S160.
- Steele, K.P. and Wojciechowski, M.F. (2003). Phylogenetic analyses of tribes Trifolieae and Viciae, based on sequences of the plastid gene *matK* (Papilionoideae: Leguminosae). In: B.B. Klitgaard and A. Bruneau (editors). *Advances in Legume Systematics, part 10, Higher Level Systematics*, pp. 355–370. Royal Botanic Gardens, Kew.
- Steele, K.P., Yang, L., Sabir, M. and Wojciechowski, M.F. (2000b). Phylogenetic relationships in the tribes Trifolieae and Viciae (Fabaceae) using sequences of Mendel's stem length gene *Le*. *American Journal of Botany* 87: S159–S160.
- Takezaki, N., Rzhetsky, A. and Nei, M. (1995). Phylogenetic test of the molecular clock and linearised trees. *Molecular Biology and Evolution* 12: 823–833.
- Thompson, I.R., Ladiges, P.Y. and Ross, J.H. (2001). Phylogenetic studies of the tribe Brongniartieae (Fabaceae) using nuclear DNA (ITS-1) and morphological data.

- Systematic Botany* 26: 557–570.
- Thorne, R.F. (1992). Classification and geography of the flowering plants. *Botanical Review* 58: 255–348.
- Thorne, J.L., Kishino, H. and Painter, I.S. (1998). Estimating the rate of evolution of the rate of evolution. *Molecular Biology and Evolution* 15: 1647–1657.
- Travis, S. E., Maschinski, J. and Keim, P. (1996). An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically endangered plant, using AFLP markers. *Molecular Ecology* 5: 735–745.
- Tucker, S.C. (1987). Floral initiation and development in legumes. In: C.H. Stirton (editor). *Advances in Legume Systematics*, part 3, pp. 183–239. Royal Botanic Gardens, Kew.
- 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. and Douglas, A.W. (1994). Ontogenetic evidence and phylogenetic relationships among basal taxa of legumes. In: I.K. Ferguson and S.C. Tucker (editors). *Advances in Legume Systematics*, part 6, Structural Botany, pp. 11–32. Royal Botanic Gardens, Kew.
- van de Wouw, M., Maxted, N., Chabane, K. and Ford-Lloyd, B.V. (2001). Molecular taxonomy of *Vicia* ser. *Vicia* based on based on amplified fragment length polymorphisms. *Plant Systematics and Evolution* 229: 91–105.
- Wagstaff, S.J., Heenan, P.B. and Sanderson, M.J. (1999). Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). *American Journal of Botany* 86: 1346–1356.
- Watson, L.E., Sayed-Ahmed, H. and Badr, A. (2000). Molecular phylogeny of Old World *Trifolium* (Fabaceae), based on plastid and nuclear markers. *Plant Systematics and Evolution* 224: 153–171.
- Wernegreen, J.J. and Riley, M.A. (1999). Comparison of the evolutionary dynamics of symbiotic and housekeeping loci: a case for the genetic coherence of rhizobial lineages. *Molecular Biology and Evolution* 16: 98–113.
- Wheeler, E.A. (2001). Fossil dicotyledonous woods from the Florissant Fossil Beds National Monument, Colorado. In: E. Evanoff, K.M. Gregory-Wodzicki, K.R. Johnson (editors). *Proceedings of the Denver Museum of Science and Industry*, series 4, pp. 197–213.
- Wojciechowski, M.F., Sanderson, M.J. and Hu, J.-M. (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., Sanderson, M.J., Steele, K.P. and Liston, A. (2000). Molecular phylogeny of the “temperate herbaceous tribes” of papilionoid legumes: a supertree approach. In: P.S. Herendeen and A. Bruneau (editors). *Advances in Legume Systematics*, part 9, pp. 277–298. Royal Botanic Gardens, Kew.
- Wolfe, A.D. and Liston, A. (1998). Contributions of PCR-based methods to plant systematics and evolutionary biology. In: D.E. Soltis, P.S. Soltis and J.J. Doyle (editors). *Molecular Systematics of Plants II, DNA Sequencing*, pp. 43–86. Kluwer Academic Publishers, Boston.
- Wray, G.A., Levinton, S. and Shapiro, L.H. (1996). Molecular evidence for deep PreCambrian divergences among metazoan phyla. *Science* 274: 568–573.
- Yang, L., Steele, K. and Lavin, M. (2001). Phylogenetic analyses of selected genera of subfamily Papilionoideae, using nucleotide sequence data of the nuclear-encoded gibberellic acid 20-oxidase genes. Abstract, Botany 2001, American Society of Plant Taxonomists, Albuquerque, New Mexico.
- Zink, D., Schumann, K. and Nagl, W. (1994). Restriction fragment length polymorphisms of the photohemagglutinin genes in *Phaseolus* and *Vigna* (Leguminosae). *Plant Systematics and Evolution* 191: 131–146.